HANDBOOK OF CHEMICAL AND BIOLOGICAL PLANT ANALYTICAL METHODS

Editor-in-Chief | Kurt Hostettmann

WILEY

Handbook of Chemical and Biological Plant Analytical Methods

Handbook of Chemical and Biological Plant Analytical Methods

VOLUME I

Part One: Sample Preparation and Identification
Part Two: Instrumentation for Chemical Analysis

Editor-in-Chief

Kurt Hostettmann

Honorary Professor at the Universities of Geneva and Lausanne, Switzerland

WILEY

Handbook of Chemical and Biological Plant Analytical Methods

VOLUME II

Part Three: Strategies for Selective Classes of Compounds

Editor-in-Chief

Kurt Hostettmann

Honorary Professor at the Universities of Geneva and Lausanne, Switzerland

WILEY

Handbook of Chemical and Biological Plant Analytical Methods

VOLUME III

Part Four: Biological Analysis
Part Five: Drugs from Plants
Part Six: Conclusion and Perspectives

Editor-in-Chief

Kurt Hostettmann

Honorary Professor at the Universities of Geneva and Lausanne, Switzerland

WILEY

Contents

Contributors

Iqbal Ahmad *Department of Agricultural Microbiology, Aligarh Muslim University, Aligarh, India*

Nadja Arens *Department of Physiology and Cell Biology, Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany*

Yoshinori Asakawa *Faculty of Pharmaceutical Sciences, Tokushima Bunri University, Yamashiro-cho, Tokushima, Japan*

Frédéric Ausseil *Institut de Recherche Pierre Fabre, Toulouse, France*

Carlo Bicchi *Dipartimento di Scienza e Tecnologia del Farmaco, Università di Torino, Torino, Italy*

Anna R. Bilia *Department of Chemistry, University of Florence, Sesto Fiorentino (Firenze), Italy*

Wolfgang Blaschek *Department of Pharmaceutical Biology, University of Kiel, Kiel, Germany*

Nadine Bohni *School of Pharmaceutical Sciences, University of Geneva, University of Lausanne, Geneva, Switzerland*

Cecilia Cagliero *Dipartimento di Scienza e Tecnologia del Farmaco, Università di Torino, Torino, Italy*

Valérie Camel *AgroParisTech, Paris, France*

Swaranjit Singh Cameotra *Institute of Microbial Technology, Chandigarh, India*

Raquel A. P. Castanheiro *Centro de Química Medicinal da Universidade do Porto (CEQUIMED-UP), CIIMAR, Departamento de Química, Faculdade de Farmácia, Universidade do Porto, Porto, Portugal*

Haimei Chen *Institute of Medicinal Plant Development, Chinese Academy of Medical Science, Beijing, PR China*

Shilin Chen *The National Engineering Laboratory for Breeding of Endangered Medicinal Materials, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, PR China*

Philippe Christen *School of Pharmaceutical Sciences, University of Geneva, University of Lausanne, Geneva, Switzerland*

Chiara Cordero *Dipartimento di Scienza e Tecnologia del Farmaco, Università di Torino, Torino, Italy*

David J. Craik *The University of Queensland, Brisbane, Queensland, Australia*

Charlotte D'Souza *The University of Queensland, Brisbane, Queensland, Australia*

Denise Dagnino *Grupo Metabolômica, Universidade Estadual do Norte Fluminense, Campos dos Goytacazes, Brazil*

Bruno David *Institut de Recherche Pierre Fabre, Toulouse, France*

Bieke Dejaegher *Department of Analytical Chemistry and Pharmaceutical Technology (FABI), Vrije Universiteit Brussel (VUB), Brussels, Belgium*

Sagar Deshpande *Jacobs University Bremen, School of Engineering and Science, Chemistry, Bremen, Germany*

Stefanie Doell *Department of Physiology and Cell Biology, Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany*

Nahed El-Najjar *Department of Physiology and Biophysics, Weill Cornell Medical College, Doha, Qatar*

Hala Gali-Muhtasib *Department of Biology, American University of Beirut, Beirut, Lebanon*

Markus Ganzera *Institute of Pharmacy, Pharmacognosy, University of Innsbruck, Innsbruck, Austria*

Elisabeth Ginko *University of Vienna, Vienna, Austria*

Kazimierz Głowniak *Department of Pharmacognosy with Medicinal Plant Unit, Medical University of Lublin, Lublin, Poland*

Maria J. Gonçalves *Center for Neuroscience and Cell Biology, Department of Zoology, University of Coimbra, Coimbra, Portugal and Center of Pharmaceutical Studies, Faculty of Pharmacy, University of Coimbra, Coimbra, Portugal*

Ana M. González-Paramás *Facultad de Farmacia, Universidad de Salamanca, Salamanca, Spain*

Sarina Grutsch *University of Innsbruck and Center for Molecular Biosciences Innsbruck (CMBI), Institute of Organic Chemistry, Innsbruck, Austria*

G. M. Kamal B. Gunaherath *The Open University of Sri Lanka, Nugegoda, Sri Lanka*

A. A. Leslie Gunatilaka *The University of Arizona, Tucson, AZ, USA*

Qiaosheng Guo *Institute of Chinese Medicinal Materials, Nanjing Agricultural University, Nanjing, Jiangsu, PR China*

Irina A. Guschina *School of Biosciences, Cardiff University, Cardiff, UK*

Maria Halabalaki *Division of Pharmacognosy and Natural Products Chemistry, Department of Pharmacy, University of Athens, Athens, Greece*

John L. Harwood *School of Biosciences, Cardiff University, Cardiff, UK*

Christian W. Huck *Leopold-Franzens University, Innsbruck, Austria*

M. Rafiq Islam *Department of Natural Sciences, Northwest Missouri State University, Maryville, MO, USA*

Rakesh Jaiswal *Jacobs University Bremen, School of Engineering and Science, Chemistry, Bremen, Germany*

Mohamed G. E. Karar *Jacobs University Bremen, School of Engineering and Science, Chemistry, Bremen, Germany*

Beatrice Kaufmann *School of Pharmaceutical Sciences, University of Geneva, University of Lausanne, Geneva, Switzerland*

Mohd Sajjad Ahmad Khan *Institute of Microbial Technology, Chandigarh, India*

Anake Kijjoa *ICBAS-Instituto de Ciências Biomèdicas de Abel Salazar and CIIMAR, Universidade do Porto, Porto, Portugal*

Eirini Kouloura *Division of Pharmacognosy and Natural Products Chemistry, Department of Pharmacy, University of Athens, Athens, Greece*

Mariusz Kowalczyk *Institute of Soil Science and Plant Cultivation, State Research Institute, Pulawy, Poland*

Liselotte Krenn *Department of Pharmacognosy, University of Vienna, Vienna, Austria*

Christoph Kreutz *University of Innsbruck and Center for Molecular Biosciences Innsbruck (CMBI), Institute of Organic Chemistry, Innsbruck, Austria*

Anja Krüger *Institute of Pharmacy, Pharmacognosy, University of Innsbruck, Innsbruck, Austria*

Justyna Krzyzanowska *Institute of Soil Science and Plant Cultivation, State Research Institute, Pulawy, Poland*

Nikolai Kuhnert *Jacobs University Bremen, School of Engineering and Science, Chemistry, Bremen, Germany*

Jianqin Li *Institute of Medicinal Plant Development, Chinese Academy of Medical Science, Beijing, PR China*

Erica Liberto *Dipartimento di Scienza e Tecnologia del Farmaco, Università di Torino, Torino, Italy*

Chang Liu *Institute of Medicinal Plant Development, Chinese Academy of Medical Science, Beijing, PR China*

Shanfa Lu *Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, PR China*

Agnieszka Ludwiczuk *Department of Pharmacognosy with Medicinal Plant Unit, Medical University of Lublin, Lublin, Poland*

Hongmei Luo *The National Engineering Laboratory for Breeding of Endangered Medicinal Materials, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, PR China*

Ana P. Martins *Center of Pharmaceutical Studies, Faculty of Pharmacy, Health Science Campus, University of Coimbra, Coimbra, Portugal*

Marius F. Matei *Jacobs University Bremen, School of Engineering and Science, Chemistry, Bremen, Germany*

Veronika R. Meyer *Swiss Federal Laboratories for Materials Science and Technology, Laboratory for Protection and Physiology, St. Gallen, Switzerland*

Mohammed J. Meziani *Department of Natural Sciences, Northwest Missouri State University, Maryville, MO, USA*

Somayeh Mirsadeghi *Department of Chemistry, Islamic Azad University, Varamin Pishva Branch, Varamin, Iran*

Hans-Peter Mock *Department of Physiology and Cell Biology, Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany*

M. Abdul Mottaleb *Center for Innovation and Entrepreneurship, Northwest Missouri State University, Maryville, MO, USA and Department of Natural Sciences, Northwest Missouri State University, Maryville, MO, USA*

Susan E. Northfield *The University of Queensland, Brisbane, Queensland, Australia*

Wieslaw Oleszek *Institute of Soil Science and Plant Cultivation, State Research Institute, Pulawy, Poland*

Brás Heleno de Oliveira *Federal University of Paraná, Curitiba, Brazil*

Xiaohui Pang *The National Engineering Laboratory for Breeding of Endangered Medicinal Materials, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, PR China*

Petra H. Pfisterer *Institute of Pharmacy/ Pharmacognosy and Pharmaceutical Chemistry, and Center for Molecular Biosciences Innsbruck, University of Innsbruck, Innsbruck, Austria*

Madalena M. M. Pinto *Centro de Química Medicinal da Universidade do Porto (CEQUIMED-UP), CIIMAR, Departamento de Química, Faculdade de Farmácia, Universidade do Porto, Porto, Portugal*

Aaron G. Poth *The University of Queensland, Brisbane, Queensland, Australia*

Seied Mahdi Pourmortazavi *Faculty of Material and Manufacturing Technologies, Malek Ashtar University of Technology, Tehran, Iran*

Emerson F. Queiroz *School of Pharmaceutical Sciences, University of Geneva, University of Lausanne, Geneva, Switzerland*

Mehdi Rahimi-Nasrabadi *Department of Chemistry, Imam Hossein University, Tehran, Iran*

Pascal Richomme *University of Angers, UFR des Sciences Pharmaceutiques, Angers, France*

Judith M. Rollinger *Institute of Pharmacy/ Pharmacognosy and Pharmaceutical Chemistry, and Center for Molecular Biosciences Innsbruck, University of Innsbruck, Innsbruck, Austria*

Patrizia Rubiolo *Dipartimento di Scienza e Tecnologia del Farmaco, Università di Torino, Torino, Italy*

Lígia R. Salgueiro *Center for Neuroscience and Cell Biology, Department of Zoology, University of Coimbra, Coimbra, Portugal and Center of Pharmaceutical Studies, Faculty of Pharmacy, University of Coimbra, Coimbra, Portugal*

Celestino Santos-Buelga *Facultad de Farmacia, Universidad de Salamanca, Salamanca, Spain*

Johannes Saukel *University of Vienna, Vienna, Austria*

Andreas Schinkovitz *University of Angers, UFR des Sciences Pharmaceutiques, Angers, France*

Jan Schripsema *Grupo Metabolômica, Universidade Estadual do Norte Fluminense, Campos dos Goytacazes, Brazil*

Daniela Schuster *Institute of Pharmacy/ Pharmacognosy and Pharmaceutical Chemistry, and Center for Molecular Biosciences Innsbruck, University of Innsbruck, Innsbruck, Austria*

Christoph Seger *University of Innsbruck, Institute of Pharmacy/Pharmacognosy, CCB—Centrum of Chemistry and Biomedicine, Innsbruck, Austria and Institute of Medical and Chemical Laboratory Diagnostics (ZIMCL), University Hospital/Landeskrankenhaus Innsbruck, Innsbruck, Austria*

Denis Séraphin *University of Angers, UFR des Sciences Pharmaceutiques, Angers, France*

Barbara Sgorbini *Dipartimento di Scienza e Tecnologia del Farmaco, Università di Torino, Torino, Italy*

Claudia Avello Simões-Pires *School of Pharmaceutical Sciences, University of Geneva, University of Lausanne, Geneva, Switzerland*

Monique S. J. Simmonds *Royal Botanic Gardens, Kew, Richmond, Surrey, UK*

Krystyna Skalicka-Woźniak *Department of Pharmacognosy with Medicinal Plant Unit, Medical University of Lublin, Lublin, Poland*

Alexios-Leandros Skaltsounis *Division of Pharmacognosy and Natural Products Chemistry, Department of Pharmacy, University of Athens, Athens, Greece*

Hermann Stuppner *Institute of Pharmacy/ Pharmacognosy and Pharmaceutical Chemistry, and Center for Molecular Biosciences Innsbruck, University of Innsbruck, Innsbruck, Austria*

Job Tchoumtchoua *Division of Pharmacognosy and Natural Products Chemistry, Department of Pharmacy, University of Athens, Athens, Greece*

Martin Tollinger *University of Innsbruck and Center for Molecular Biosciences Innsbruck (CMBI), Institute of Organic Chemistry, Innsbruck, Austria*

Masao Toyota *Faculty of Pharmaceutical Sciences, Tokushima Bunri University, Yamashiro-cho, Tokushima, Japan*

Aurélie Urbain *Pharmacognosy and Bioactive Natural Products, University of Strasbourg, Illkirch, France*

Arto Urtti *Centre for Drug Research, University of Helsinki, Helsinki, Finland*

Heikki Vuorela *Division of Pharmaceutical Biology, University of Helsinki, Helsinki, Finland*

Pia Vuorela *Division of Pharmaceutical Biosciences, University of Helsinki, Helsinki, Finland*

Jean-Luc Wolfender *School of Pharmaceutical Sciences, University of Geneva, University of Lausanne, Geneva, Switzerland*

Christoph H. Wunderlich *University of Innsbruck and Center for Molecular Biosciences Innsbruck (CMBI), Institute of Organic Chemistry, Innsbruck, Austria*

Haibin Xu *The Key Laboratory of Bioactive Substances and Resources Utilization of Chinese Herbal Medicine, Ministry of Education, Institute of Medicinal Plant Development, Beijing, PR China*

Xiaolan Xu *Institute of Medicinal Plant Development, Chinese Academy of Medical Science, Beijing, PR China*

Liming Yang *Huaiyin Normal University, Huaian, Jiangsu, PR China*

Zaibiao Zhu *Institute of Chinese Medicinal Materials, Nanjing Agricultural University, Nanjing, Jiangsu, PR China*

Monica R. Zuzarte *Center for Neuroscience and Cell Biology, Department of Zoology, University of Coimbra, Coimbra, Portugal*

Preface

There are many books and book series dealing with chemical and biological methodologies for plant analysis. Thus, was it necessary to publish another book? That was my thought when Martin Röthlisberger from John Wiley & Sons approached me to become the Editor-in-Chief of the present Handbook. I was in the first instance not very enthusiastic and had some hesitations before finally accepting the task. Plants and plant-derived compounds and drugs are becoming more and more popular and also more and more researchers are involved in plant analysis. Quality control of herbal drugs is becoming essential to avoid severe health problems. In addition, in the future, many new drugs will be developed from plant sources. The present Handbook is quite unique as it deals with chemical and biological methodologies for plant analysis. It is a handbook and not an encyclopedia. Thus, it does not present all methods that are available for plant analysis, but a selection of the most important and most accurate ones. Before any analysis, there is an important step involving plant selection and collection, followed by extraction and sample preparation. Several instrumentations for chemical plant analysis are presented with an emphasis on hyphenated techniques such as the coupling between HPLC and mass spectroscopy and HPLC and NMR. A section of this Handbook is devoted to strategies for selective classes of compounds. However, not all classes of plant constituents are reviewed but the most interesting ones such as polysaccharides, saponins, cardiotonic glycosides, alkaloids, terpenoids, lipids, volatile compounds, and polyphenols (flavonoids, xanthones, coumarins, naphthoquinones, anthraquinones, proanthocyanidins, etc.). An interesting section deals with biological analysis including phenotyping, DNA barcoding techniques, transcriptome analysis, microarray, metabolomics, and proteomics. The fifth section is devoted to the screening of plant extracts and to strategies for the quick discovery of novel bioactive natural products. Safety assessment of herbal drugs is highly dependent on outstanding chromatographic and spectroscopic methods, which are highlighted here.

The aim of this Handbook is to introduce scientists involved in plant studies and current knowledge of methodologies to various fields of chemically and biochemically related topics in plant research. Emphasis is put on the rapid identification of constituents that could become drugs in the future. When we started work on this Handbook, I had three co-editors to assist me in this task. Unfortunately, one of them passed away, namely Professor Andrew Marston, before the book was completed. In order to honor his memory, this Handbook is dedicated to him, and you will find a short text related to him.

I would like to express my thanks to the two co-editors for their great help in the elaboration of this Handbook and to all the contributors for their collaboration by providing excellent manuscripts.

Dr Kurt Hostettmann
Champex-Lac, Switzerland
July 2014

In Memoriam, Andrew Marston, November 16, 1953 to March 26, 2013

It is a very sad moment for a retired professor to write an obituary for a younger colleague and friend. Andrew studied chemistry at the University College, London, and obtained his BSc degree in 1975. I met him for the first time in the same year when he joined the University of Neuchâtel, Switzerland, as a British Council award holder. He was involved in the research on phytochemistry of gentians and published his first paper with me on flavonoids of *Gentiana pyrenaica*. He not only liked to work in the laboratory but also enjoyed to work in the fields, as he had an excellent knowledge in taxonomy. In fact, we made a beautiful journey together to the French Pyrenees in order to collect the first plant he was working on. This trip was followed later by numerous scientific expeditions all over the world. After Neuchâtel, Andrew went back to England to write his PhD thesis at The University of Liverpool in the field of peptide synthesis, followed by a postdoctoral stay at the German Cancer Research Centre, Heidelberg, Germany, from 1979 to 1983. In October 1983, he joined the Institute of Pharmacognosy and Phytochemistry, University of Lausanne, Switzerland, to work with me on a Swiss National Science Foundation research project for one year. He was a brilliant young scientist, and the initially planned one year stay became a stay of 26 years! Andrew was involved in the isolation of biologically active compounds from plants used in traditional medicine and in the application of new chromatographic techniques for the separation and isolation of plant constituents. He has done pioneering work in the field of centrifugal partition chromatography, which resulted in the publication of research papers and a couple of review articles. He also achieved original work in the development of enzyme inhibition tests on TLC plates (TLC bioautography), which is useful for the search of acetylcholinesterase inhibitors from plants (Treatment of Alzheimer's disease). For his important contribution in various fields of phytochemistry, Andrew received, in 1994, the prestigious Rhône – Poulenc Rorer Award of the Phytochemical Society of Europe. We published together a book on preparative chromatography techniques, which was translated into Japanese, Chinese, Indonesian, Farsi, and Spanish. He is also co-author of a very complete monography on saponins. In 1994, my institute was transferred from Lausanne to Geneva University where Andrew held the position of *Maître d'enseignement et de recherche* (which corresponds to Senior Lecturer) until my retirement in 2009. When I retired, Andrew decided to look for another job and became Professor of Chemistry at the University of the Free State, Bloemfontein, South Africa. He was conducting in his new job phytochemical investigation on indigenous plants and teaching organic chemistry and natural product chemistry.

Andrew was an outstanding phytochemist, and his work resulted in the publication of more than 150 research papers and 35 review articles and chapters in books. He presented lectures and oral communications in numerous international symposia. He was also teaching in workshops held in Uruguay, Panama, Mexico, Peru, Brazil, Thailand, China, Indonesia,

Zimbabwe, Botswana, and Mali! He passed away on March 26, 2013 in Bloemfontein after a surgery of the brain to control his Parkinson's disease, which resulted in cerebral hemorrhages. He was born in Africa (Northern Rhodesia that became Zambia after independence in 1964) and died in Africa. Moreover, the scientific community has lost a great phytochemist. Everybody will miss Andrew because he was always modest, friendly, and helpful. I shall miss a friend whom I considered as my younger brother.

Dr. K. Hostettmann

Part One

Sample Preparation and Identification

1

Selection, Identification, and Collection of Plants

Monique S. J. Simmonds

Royal Botanic Gardens, Kew, Richmond, Surrey, UK

1   SELECTION OF PLANTS

An appropriate, well-researched strategy for the selection of plants for a study on natural product or drug discovery is often the key to a successful project. Most studies on natural product provide very little information as to why the specific plants were selected, other than to indicate that the species were selected because they were known to have medicinal or pesticidal properties. In these studies, the authors rarely provide any evidence that the plants being extracted in the laboratory have the medicinal properties attributed to that species in the literature. For example, have the plants been obtained from those that are using them? There are many pitfalls that are common to this field of research, but the proof that these pitfalls have been addressed is not evident from the Material and Methods section of a scientific paper. For example, few papers provide information that enables the researcher to evaluate the robustness of the historical information about the traditional uses of a species, or evidence that the authors read the original papers that describe the traditional uses or collected the plants from those who have traditionally used them. The confusion in the identification of the species was highlighted by Hsu (2006, 2010) who studied the earlier Materia Medica in China and reported that Shen Gua (1031–1095) back in 1086 documented the difference between the species as they are prepared in different ways. The literature is full of examples of assumptions being made about the history and the identification of plants. This short review aims to help highlight the importance of documenting the uses of the plants along with information about the plants being studied and the importance of placing the work on natural product into the context of policies, especially those that support the conservation of the natural resources we study. The majority of examples are from medicinal plant research, but the issues are often common to other natural resources.

Recent work on the development of the antimalarial compound, artemisinin, from *Artemisia annua* L. is now questioning whether the plant used over 2000 years ago in traditional Chinese medicine (TCM) to treat fevers was in fact *A. annua* (cao hao) or another species *Artemisia apiacea* Hance (quing hao)? It would appear that the species that was historically used in TCM to treat conditions now known to be associated with malaria was actually *A. apiacea*. The confusion in the identification of the species was highlighted by Hsu (2006), who studied the earlier Materia Medica in China and reported that Shen Gua (1031–1095), back in 1086, documented the difference between the species as they are prepared in different ways. Despite this finding, antimalarial research still focuses on artemisinin isolated from *A. annua* and the cultivation of chemotypes that have the potential to yield high amounts of artemisinin. Although over 600 compounds have been identified in *A. annua*, there is to date very little information

about the chemical potential of *A. apiacea*. The questions can be raised as to why more emphasis has not been placed on the traditionally used species *A. apiacea*, rather than *A. annua*. However, there is no doubt that the isolation of artemisinin from *A. annua* has made a major contribution to the treatment of malaria. Further advances in the use of these species might occur if a comparative study is undertaken of the two species as prepared traditionally. Such a study could include a standard way of extraction the plants (e.g. ethanol extraction) as well as the traditional methods used to make the extractions as well as the traditional formulae (van der Kooy and Sullivan, 2013). The information could assist in highlighting differences in the profile of compounds extracted from the plant that could impact the efficacy of the extracts. The different extracts could be tested through a system biology or pharmacometabonomics approach (Everett, Loo, and Pullen, 2013), in which blood and urine samples of animals or volunteers are analyzed along with the plant material. The information coming from these studies could increase our knowledge about the importance of the complexity of the diversity of compounds in the extracts and how some of these compounds could modulate the enzymes in the different parts of the alimentary channel that influence bioavailability of active compounds (Magalhaes *et al.*, 2012).

A key then to the start of a project is that the researcher should be confident that the plant they propose to study is the correct species, especially if they are going to evaluate the traditional uses of that plant. They should also think about how they are going to check and collate information about the species and how confident are they that the literature they are citing relates to the species they propose to study. This means that they should check with a botanist the identity of the material they propose to work with, as well as the scientific Latin binomial name of the plant. It is also suggested that help is sought to bring together all the relevant names together of the selected species before starting the search. This should include not only the Latin scientific botanical name but also the pharmaceutical names as well as common names. The fact that, currently, in this age of advanced internet searching systems, there is no one central resource that brings all plant names together reflects the complexity of the task (Chan *et al.*, 2012).

The following example illustrates how complex the botanical aspects of a literature review can be. One of the popular Chinese medicinal plant formulation used in China is Liu Wei Di Huang Wan "Six Flavor Rehmanni," which contains material from five plants and a fungus. Researchers wanting to collate information about what the activity of the formulation is could undertake a search of the literature using the Pin Yin name of the formulae Liu Wei Di Huang Wan or they could use the names of the plants and fungus used in the formulation. Table 1 illustrates the complexity in undertaking this task. Currently, a medicinal plant could have many scientific names as well as the accepted Latin binomial name and authority given to a species when it was first described. If researchers want to undertake a thorough review of the literature, they will need to have not only all the scientific names but also the common or trivial names of the species. It is also suggested that they include synonyms in the search terms. For example, a quick search of the Web of Science will show that there are more papers about the medicinal uses of the species *Paeonia ostii* T. Hong and J. X. Zhang using the name *Paeonia suffruticosa* Andr. (over 100 papers since the revision of the genus in 1999) than the accepted name *P. ostii* (14 papers since 1999, of which none are about its medicinal uses). It was initially thought that *P. siffruticosa* was a synonym for *P. ostii*, but the recent revision of the genus indicates that *P. suffruticosa* is a different species and not the one cited in the majority of papers (see Table 1 for details).

Collating the names together will enable the researchers to search for information that could relate to the species they propose to study, but how reliable is this information? The identification of plants is a common problem associated with research on the biological activity and chemistry of plants and fungi. The extent of the problem is difficult to establish but it will most likely increase with fewer students being taught traditional taxonomy and plant identification. The development of new mobile applications might assist but the technology is not there yet. The simple question to ask is did someone check the identification of the species being studied? Not all specimens will be easy to identify, especially if when the plant was collected, it was not in flower. One way to help deal with potential problems with the identification of plant is for researchers to keep numbered vouchers of the material they study, these vouchers be kept by their institute, be identified by their number in scientific papers, and be made available to others to check. This is an old tradition used in herbaria around the world and allows others

Table 1 Matrix of plant and fungus species names for guiding literature search and analysis.

Latin scientific name(s) as stated in the Pharmacopoeia of the PRC (English editions)[1,2]	Accepted* Latin scientific name[3,4,5]	Plant part used in TCM[1,2]	Latin pharmaceutical/ pharmacopoeia name(s)[1,2]	Chinese and Pin Yin names[1,2]	Latin scientific synonyms[3,4,5]	Official substitute species in Japanese formulations of Liu Wei Di Huang Wan	Unofficial substitutes and adulterants (using Latin scientific names[note a], ordered alphabetically by family)
The following six species are those used in the TCM formula: "Liu Wei Di Huang Wan"[1]
Alisma orientale (Sam.) Juzep[2] (note different and incorrect spelling in the 2005 edition of the Chinese Pharmacopoeia "*Alisma orientalis* (Sam.) Juzep.")	*Alisma orientale* (Samuelsson) Juzepczuk[3]	Tuber/rhizome	Rhizoma Alismatis[1]; Alismatis Rhizoma[2]	澤瀉 "Ze Xie" or "Zexie"	*Alisma plantago-aquatica* Linnaeus var. *orientale* Samuelsson[3]; *A. plantago-aquatica* subsp. *orientale* (Samuelsson) Samuelsson[4]	—	None known
Cornus officinalis Sieb. et Zucc.[1,2]	*Cornus officinalis* Siebold and Zuccarini[3]	Ripe fruit (processed with yellow rice wine)	Fructus Corni[1]; Corni Fructus (processed with wine)[note b]	山茱萸 "Shan Zhu Yu" or "Shanzhuyu"	*Macrocarpium officinale* (Siebold and Zuccarini) Nakai[3]; *Cornus officinalis* var. *koreana* Kitam.[4]	—	Many species in same or other families reported as local and/or historical unofficial substitutes but international trade uncertain: Berberidaceae: *Berberis amurensis.*, *Berberis Poiretii*; Caprifoliaceae: *Viburnum schensianum*; Cornaceae: *Cornus oblonga*; Rhamnaceae:

(continued overleaf)

Table 1 *(Continued)*

Latin scientific name(s) as stated in the Pharmacopoeia of the PRC (English editions)[1,2]	Accepted* Latin scientific name[3,4,5]	Plant part used in TCM[1,2]	Latin pharmaceutical/ pharmacopoeia name(s)[1,2]	Chinese and Pin Yin names[1,2]	Latin scientific synonyms[3,4,5]	*Official* substitute species in Japanese formulations of Liu Wei Di Huang Wan	*Unofficial* substitutes and adulterants (using Latin scientific names[(note a)], ordered alphabetically by family)
							1. Ziziphus mauritiana, Z. jujuba var. spinosa; Rosaceae: *Cerasus pleiocerasus, Malus baccata, Crataegus pinnatifida;* Rubiaceae: *Rubia cordifolia* Vitaceae: *Vitis vinifera, Vitis amurensis*
Dioscorea opposita Thunb.[1,2]	*Dioscorea polystachya* Turczaninow[3;(note c)]	Tuber/rhizome	Rhizoma Dioscoreae[1]; Dioscoreae Rhizoma[2]	山药 "Shan Yao" or "Shanyao"	*Dioscorea batatas* Decaisne[3]; *Dioscorea decaisneana* Carrière[3]; *Dioscorea doryphora* Hance[3]; *Dioscorea potaninii* Prain and Burkill[3]; *Dioscorea rosthornii* Diels[1]; *Dioscorea swinhoei* Rolfe[3]; *Dioscorea batatas* f. *clavata* Makino[4]; *D. batatas* f. *daikok* Makino[4]; *D. batatas* f. *flabellata* Makino[4]; *D. batatas* f. *rakuda* Makino[4]; *D. batatas* f. *tsukune* Makino[4];	

					Dioscorea japonica Thunb.[4,6]	*Dioscorea cayennensis* var. *pseudobatatas* Hauman[4]; *D. pseudobatatas* (Hauman) Herter[4]	2. *Dioscorea* spp. plus those of species in other families are reported as substitutes including *Dioscorea persimilis*, *Dioscorea fordii*, and *Dioscorea alata* plus: Convolvulaceae: *Ipomoea batatas* Euphorbiaceae: *Manihot esculenta*
Paeonia suffruticosa Andr.[1,2] / *Paeonia ostii* T. Hong and J. X. Zhang[3;(note d)]	Rootbark	Cortex Moutan[1]; Moutan Cortex[2]	牡丹皮 "Mu Dan Pi" or "Mudanpi"	*Paeonia ostii* var. *lishizhenii* B. A. Shen[3]; *P. ostii* subsp. *lishizhenii* (B. A. Shen) B. A. Shen[3]; *P. suffruticosa* Andrews subsp. *ostii* (T. Hong and J. X. Zhang) Halda[3]	—	Paeoniaceae: *Paeonia anomala* subsp. *veitchii* (synonym: *Paeonia veitchii*); *Paeonia delavayi*, and *Paeonia decomposita*
Poria cocos (Schw.) Wolf.[1,2] / *Wolfiporia extensa* (Peck) Ginns[5]	Sclerotium (storage organ and non-reproductive part)	Poria[1,2]	茯苓 "Fu Ling" or "Fuling"	*Daedalea extensa* Peck; *Macrohyporia cocos* (Schwein.) I. Johans. and Ryvarden[5]; *Macrohyporia extensa* (Peck) Ginns and J. Lowe[5]; *Pachyma cocos* (Schwein.) Fr.[5]; *Poria cocos* F.A. Wolf[5]; *Sclerotium cocos* Schwein.[5]; *Wolfiporia cocos* (F.A. Wolf) Ryvarden and Gilb.[5]	—	None known

Table 1　(*Continued*)

Latin scientific name(s) as stated in the Pharmacopoeia of the PRC (English editions)[1,2]	Accepted* Latin scientific name[3,4,5]	Plant part used in TCM[1,2]	Latin pharmaceutical/pharmacopoeia name(s)[1,2]	Chinese and Pin Yin names[1,2]	Latin scientific synonyms[3,4,5]	Official substitute species in Japanese formulations of Liu Wei Di Huang Wan	Unofficial substitutes and adulterants (using Latin scientific names[(note a)], ordered alphabetically by family)
Rehmannia glutinosa Libosch.[1,2]	*Rehmannia glutinosa* (Gaertner) Liboschitz ex Fischer and C. A. Meyer[3]	Processed tuber/rhizome[(note e)]	Radix Rehmanniae Praeparata[1]; Rehmanniae Radix Praeparata[2]	熟地黄 "Shu Di Huang," or "Shudihuang"	*Digitalis glutinosa* Gaertner[3]; *Rehmannia chinensis* Liboschitz ex Fischer and C. A. Meyer[3]; *R. glutinosa* var. *hemsleyana* Diels[3]; *R. glutinosa* var. *huechingensis* Chao and Shih[3]; *R. glutinosa* f. *huechingensis* (Chao and Shih) P. G. Hsiao[3]; *R. glutinosa* f. *purpurea* Matsuda[3]	—	None known

References

1. State Pharmacopoeia Commission of the P.R. China (2005, English ed.) *Pharmacopoeia of the People's Republic of China*, vol 1, 791 pp.
2. State Pharmacopoeia Commission of the P.R. China (2010, English ed.) *Pharmacopoeia of the People's Republic of China*, vol 1,1584 pp.
3. *Flora of China* (http://flora.huh.harvard.edu/china).
4. *World Checklist of Selected Plant Families* (http://apps.kew.org/wcsp).
5. *Index Fungorum* (http://www.indexfungorum.org/Names/Names.asp).
6. Anon. (2006, 15th edition, English) *Japanese Pharmacopoeia* (http://jpdb.nihs.go.jp/jp15e/).

Notes

* an "accepted" Latin scientific name is a taxonomically preferred name (i.e., it should be used in preference to any other Latin scientific name used for the same species).

a Authors (e.g., "L.") of Latin scientific names are not included in column 8, because these were not available in the source texts consulted; they cannot be added retrospectively because of the potential for creating confusion with their homonyms (i.e., different species that share the same Latin name).

b The Chinese Pharmacopoeia 2010 definition of the formula "Liuwei Dihuang Wan" states Fructus Corni is "*processed with wine*" (i.e., steamed or stewed in yellow rice wine).

c The name "*Dioscorea opposita* Thunb." is considered taxonomically[3] "illegitimate" and therefore should not be used in a scientific context even as a scientific synonym. Owing to its continued common use in TCM, however, it is a useful name for literature searching. Little has yet been published under the "accepted" Latin scientific name of *Dioscorea polystachya* Turczaninow.

d A recent and authoritative taxonomic revision (Hong De-Yuan Hong and Pan Kai-Yu, 1999, *Nord. J. Bot.*, 19(3): 289–300) concludes that the name "*Paeonia suffruticosa* Andr." has been misapplied for many years as the source of the TCM herb "Cortex Moutan" or "Moutan"; instead, the correct identity of this TCM herb is "*Paeonia ostii* T. Hong and J. X. Zhang." *P. suffruticosa* Andr. is considered to be a separate species (not a synonym) and includes a huge array of popular cultivars grown for ornamental purposes only. Owing to the widespread use of the name *P. suffruticosa* in TCM, however, it continues to be a useful name for literature searching.

e "Processed" defined here means "stewing"[2] [i.e., steaming or stewing with yellow rice wine (20–30 kg wine/100 kg of clean crude dried rhizomes/tubers/roots) until wine is totally absorbed].

to compare and check specimens. Currently, there is a drive to store all herbaria samples in a digital form; this is understandable as space becomes expensive, but with a classical herbarium specimen, there is the ability to take samples for DNA and chemical studies; thus, a case can be made for keeping the original voucher samples of key species.

Trends in science also modulate the focus and the methods used to select the plants. Between the 1960s and early 1990s, many companies had natural product screening projects in which they obtained samples from all over the world to screen against a portfolio of different targets. Research councils also supported the funding of related projects in universities with the aim of increasing the number of leads as well as the speed with which they could be screened. The high-throughput screening of natural resources then made way for the screening of computational libraries of compounds. By the mid-1990s, companies used their molecular discovery budgets for the purchase of these libraries rather than the collection of plants and fungi for natural resource screening.

Although there has been a decrease in companies randomly screening plants in high-throughput assays for potential new drugs, many university departments still have some capacity to undertake natural product research. Recently, there is an increased interest in the pharmaceutical industry in randomly screening natural compound libraries; however, very few of these libraries exist in the public sector.

The random screening programmes have clearly added to our knowledge about the diversity of compounds in plants. However, very little information has been released from the companies to enable a review of the success of these screening programs and the diversity of plants actually screened and against what targets. The National Cancer Institute in the United States screened over 35,000 species between 1960 and 1980, resulting in the development of paclitaxel and camptothecin (Fabricant and Farnsworth, 2001). Another report indicated that until 1992, the National Cancer Institute had found three plant extracts active against HIV (human immunodeficiency virus) out of 50,000 tested, and three out of 33,000 plant extracts tested were found to have antitumor activity (Williamson, Okpako, and Evans, 1996). Many companies closed their random screens because of the expense and high numbers of extracts needed and issues in being able to follow up leads when they occurred.

So, trends in natural product research have evolved to reflect changes in other aspects of science. In the late 1960s, advances in analytical chemistry assisted chemists to evaluate the complexity of the extracts used in medicinal plants, then advances in bioassays enabled the development of large random screening programs, and now the advances in both the understanding of diseases and the speed of hyphenated analytical chemical techniques are enabling more scientific rigor to be placed on the evaluation of medicinal plants. This is also happening at a time when the health services in many countries are looking at the value of integrated health systems.

Currently, the interest in understanding traditional health care systems that involve medicinal plants is increasing; yet, in many cases, changes in land use are having a negative impact on the supply of the plants to support these systems. This impact is greatest on those poorer populations often in developing countries that rely on the plants for their primary health care (Cordell, 2012). Measuring this impact is steeped in guess work. For example, how robust is the frequently quoted figure that 80% of the world's population relies on plants for their main source of medicine (Farnsworth *et al.*, 1985). Providing more accurate information about the plants being used in traditional and other forms of medicine along with information about how many have been studied would not only enable us to evaluate the robustness of this figure but also establish with greater clarity the importance of plants in our health systems, the habitats that supply the plants, as well as the proportion of these plants that have been subject to some form of scientific validation. The only way that this can be done is through linking information about the traditional medicinal uses of plants with studies undertaken on the plants along with information about the habitats they grow in. The common link among the different information sets is plant names, and as indicated earlier, currently, it is not possible to make these links with any degree of confidence.

Having a greater understanding and thus confidence of which plants have been studied and against which diseases would enable scientists to focus their research activities and thus help further our understanding of medicinal plants as well as support the habitats that support their growth. There is clearly a need to bring together more information about what is known about medicinal plants so that there can be a greater focus on scientifically evaluating those that show most promise. This could be as sources

of new lead molecules for new drugs or to provide information that would inform how standardized extracts with a defined profile of compounds could be obtained.

2 IDENTIFICATION

In many scientific papers dealing with medicinal or pesticidal plants collected from the wild, there is no indication of who helped the authors identify the plant that they worked on. Similarly, research based on samples obtained from the trade including market stalls does not provide information that confirms the identification of the plant or the part of the plant used. It is presumed that the identification of the material was correct. However, there are increased concerns that this might be a weak point in medicinal plant research publications, and thus, there is a real need to improve the way the samples are validated. This is especially true if the reputation of medicinal plant research is to be improved. Some journals, such as the Journal of Ethnopharmacology, now insist that the authors provide information that would enable others to check the status of the plant used in the published paper. For example, authors are asked to provide a voucher sample of the material studied along with information about the location of where the voucher is kept. They also recommend that the authors provide a chemical fingerprint of the sample worked on. This will enable others to check the identification should that be required as well as provide an estimate of quality of the chemistry. This will enable other researchers to compare material from different areas and help evaluate why in some cases plants are active, whereas in others they are not. Currently, the majority of plants are identified using classical morphological methods, but there is an increased interest in the use of DNA bar coding. The consortium for the Barcode of Life recommended that the two-locus combination of *rbcL* + *matK* could be used to discriminate among plants, but a study of 907 samples from 550 species had a success rate of only 72%. It has been suggested that the second internal transcribed spacer (ITS2) of nuclear ribosomal DNA region is a more robust marker (Chen *et al*., 2010). If this marker is to be used, then plants need to be further studied to validate the method to make sure it can be used to differentiate among closely related species as well as substitutes and adulterants. It is clear that DNA/genomic methods will play an important role in the future not only for the identification of the species but also for checking the proportion of different plant material present in a multi-ingredient product. However, although these methods can assist with identification of species, the chemical data are still required to evaluate quality.

The method of identification will in part reflect the question being asked. The development of "onomic" techniques could enable DNA-based profiling to be linked to metabolomic profiling that will link identification to quality. However, to do this, more research is needed to evaluate what exactly is the chemical profile of a quality plant. In some cases, the active ingredients in a plant are known but not the chemical profile or chemical fingerprint associated with quality (i.e., content of the active compound and other compounds in the plant extract) (Buriani *et al*., 2012).

A recent European-Union-funded project on "good practise in traditional Chinese medicine (GP-TCM)" provided a forum to discuss many of the issues associated with the validation of the vast amount of literature, especially, in China, on the plants used in TCM (www.GP-TCM.org). In TCM and in some other forms of traditional medicines (TMs), plants are used in mixtures and there were very few publications in English evaluating the medicinal properties of these formulae that follow the standard experimental methods used in the western literature to compare the formulae with a positive control. This project made a series of recommendations as to the scope of information that should ideally be known about the plant material being studied. The scope of information is more than most journals would accept in a Material and Methods section, but the information is very relevant for many medicinal plant projects. Thus, when undertaking a review of the literature on a species, the designing of a simple database to collate relevant information when reading the papers can help identify the gaps in the knowledge about a species. It might sound a lot of work, but it usually saves time as the information is recorded when reading the paper, and a record is kept of which papers contained the information.

The scope and amount of information to be collected will depend on the project as well as the stage within a project. However, if material is going to be used for a large-scale study, then it is important that as much of the information is obtained as this will assist in the later stages of the project. These are some suggestions as to what information should be recorded:

2.1 Species Name

Authors should include the Latin binomial name, authority and family of the plant being studied (with a reference to the source of the name: Flora, The Plant List, World Checklist) and if possible any synonyms and the name used in the local pharmacopoeia. It is also useful to have information about the local vernacular names of the species if available. When reading the paper note whether the authors had their plant verified and if there is a voucher. This information can be useful if the results from the study differ from others.

2.2 Part of Plant Used and Processing

Information should be obtained about the part of the plants used, when it is harvested, how it is harvested, and how it is prepared for use. For example, is it processed in any way to remove toxins?

2.3 Plant Source

Information should be obtained on where (country, area) does the plant come from, whether it is wild or cultivated material, and whether it is known to be harvested from an area that is traditionally associated with quality. If the material is wild harvested, what procedures are in place to avoid over harvesting and is the material grown under a specific organic certification system. Has the material been sourced from a supplier that follows the World Health Organization (WHO) good agricultural and collection practise and has the material got a certificate to confirm this? Has the material been tested for pesticides, heavy metals, and bacteria? If you are obtaining plant material from a commercial trade, then you should have the paperwork to indicate how it has been grown and whether it has been tested for pesticides and heavy metals.

2.4 Medicinal (or Other) Uses

Information must be obtained on what is known about the medicinal uses of the plant, what evidence is there that the literature relates to the plant you want to study, what part of the plant is used, whether it is used alone, how is it prepared, and how often and how much is to be taken (dose). What is known about the mode of action of the plant associated with the known uses? Are there any known adverse responses associated with the plant and is the plant known to cause herb–drug interactions?

2.5 Chemistry

Information should be obtained on what is known about the chemistry of the species, what is known about the chemistry of the plant used in the study, how was the extract of the plant made, and whether the method reflects the traditional use of the plant. Is there information available on the chemistry of the extract being tested? Does the plant contain compounds that other researchers have shown are active in screens and thus could be associated with the traditional uses of the plant? If a new compound is identified or a known compound is isolated and shown for the first time to explain the traditional use, do the authors indicate that they have protected the information, for example, by a patent? Having this knowledge at the start of the project can assist in evaluating the next stages of the project. Does the paper contain information about the identification of compounds that although not associated with the focus of the study might be useful as markers for the identification of the plant? Note should also be made of any toxic or potentially toxic compounds in the plant. These toxins could be removed by appropriate extraction methods.

2.6 Supply of Plants

At the start of a study, few researchers pay much attention to the supply of the plant material they are working on. This can be critical to the success of a project and should be addressed as soon as it becomes clear that more material will be needed. Researchers should look at the supply chain and establish whether the plant is cultivated or wild harvested and if cultivated whether it meets WHO best good agricultural practices for medicinal plants. Information should also be obtained about the growing conditions of the plant. This could influence the levels of activities in the plant.

3   COLLECTION OF PLANTS

Obtaining plant material from botanic gardens or from collaborators abroad can take time. There can also be a lot a paperwork, and researchers need to be prepared for this. Bureaucratic processes associated with obtaining plants for research can vary from country to country and from institute to institute. Researchers need to consider the costs of these procedures in their grant applications. Time spent communicating with potential collaborators at the early stages of a project can usually assist with these procedures. In the 1980s and early 1990, many pharmaceutical companies had large natural product/plant screening programs with academic institutes, and they would fund plant collecting trips in areas of high biodiversity, but very few of these projects produced leads of commercial value. However, there is a legacy that high biological diversity is linked to high "natural" wealth. Realizing this potential is a challenge. Currently, there is a very narrow range of about 3000 medicinal plants being traded internationally, and in most cases, their source does not link to mega-diverse countries. This is not to say that high diversity does not have the potential to produce more leads. In fact, the diversity of plants being traded in local communities is most likely very high, but there is no system in place to monitor the trade. In area of high diversity, most species will have a narrow geographical distribution, and thus, the uses of the plants will also be restricted. Few academic studies on these medicinal plants go beyond reporting their activity in an ethnobotanical survey. Therefore, the real potential of these plants is difficult to evaluate, especially when access to these plants for evaluation in university laboratories, even in the country in which they occur, is often complex. Despite these limitations, the results of these surveys are collated. Trends in the use of closely related species become clearer and can highlight species that justify further study. The development of ecosystem services is also likely to help evaluate the real link between biodiversity and social use of different species.

The implementation of the Convention on Biodiversity (CBD), agreed at the World Sumit at Rio de Janerio in December 1992, has had significant impacts on the exchange of natural resources for screening projects. This was especially true in the late 1990s when it was very unclear as to how to implement the legislation and policies associated with not only the sovereign rights of countries on their genetic resources but also the ownership these countries have on the traditional knowledge about the uses of these resources. These issues are very current today and information on the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their utilization can be found on the CBD website (www.cbd.int/abs/). This website also provides information about how to implement the CBD.

The setting up of regional officers around the world has provided a point of contact for scientists and others to find out about local CBD requirements. However, it is still complex and scientists are encouraged to work with the appropriate authorities to ensure that have the relevant permits before they undertake studies on the chemistry of plants, especially if their study could be targeted at drug discovery. To have a better understanding of what is needed before starting to obtain plants from different countries, it is suggested that reference be made to the national biodiversity strategies and action plans for the country it is planned to work in (www.cbd.int/nbsap/). Working through the protocols is complex and it is suggested that use is made of the many guidelines that have been produced (www.cbd.int/nbsap/guidance-tools/guidelines.shtml) as well as seeking professional help.

These guidelines provide an insight into the different procedures used by countries, and information is available about the points of contact within the countries that have endorsed the CBD (www.cbd.int). Since these procedures are in place, the potential for international and national teams to work together on evaluating medicinal plants has increased. Such studies should further increase the understanding of the traditional uses of plants as well as support the conservation of these species. Placing the proposed work into the context of an international strategy can assist to gain access. For example, the findings of many studies on medicinal plants could support some of the objectives of international strategies on health (WHO) and plant conservation (Global Strategy for Plant Conservation, GSPC). Medicinal plant research project is often driven by laboratory-based scientists wanting to explain the traditional uses of plants rather by local users wanting to validate the plants they use. It would greatly assist the funding of these projects if the drive for undertaking the work was reversed so that the local communities became the key stakeholders.

Historically, the majority of the plants that have been the source of some of our prescription drugs, such as the poppy, *Papaver somniferum*, and the foxglove *Digitalis* spp., are those with a long tradition of use. Thus, there is a clear value in collating information about the uses of plants and being able to evaluate their chemistry and pharmacology.

The interest in understanding traditional health systems goes in parallel with the increased emphasis by many companies in expanding their range of over-the-counter (OTC) medicines. For example, smaller pharmaceutical companies are interested in evaluating the potential of getting more products based on plants with a known history of traditional use to market as herbal medicines, functional foods, or novel foods. There is also an increased interest in looking at plants for sources of new pest control agents to replace the decreasing range that are currently available for use.

This is an interesting and timely shift that could help support both medicinal and pesticidal plant research, especially research that involves gathering information about the traditional uses of plants and then scientifically evaluating these uses. At a time when there are increasing demands on the land available to us to grow plants to feed, clothe, and house increasing populations, this interest could help identify the importance of a greater range of plants and if well thought out could support the conservation of these plants. However, there are many challenges. For example, climate changes are resulting in the loss of habitats and thus of plant diversity; hence, activities that could result in the increased demand of a species need to planned to ensure that there is a sustainable supply of the species to meet demand. Currently, the supply and value-added chains for plants are not usually linked in ways that support the funding of the conservation of the habitats that support the supply chain. However, organizations, such as Fair Trade, are making advances in the monitoring of the value chain. In fact, in many cases, the companies involved in the development of the end product are investing more in the supply chain, because of the importance they place in ensuring that they are supplied with quality plants. These are issues that go beyond the initial criteria used by a scientist in selecting their plants but as part of a holistic study need to be considered.

Thus, an emphasis does not always need to be on the study of poorly documented medicinal species. A strong case can be made for collating information about widely used plants that have not yet been commercialized but could be suitable for development as nonprescription "OTC" medicine. The regulations for defining what is classed as an OTC varies from country to country but the plants to be used in an OTC must have a long tradition of use and no known toxicity. Currently, is it not only the smaller companies that are interested in the OTC market but the larger pharmaceutical companies are also showing an increased interest in having a greater portfolio of OTC, as they are less expensive to get to market than a new drug.

The WHO in their 2002–2005 strategy for TM (WHO, 2002) provided a framework by which different countries could develop their strategies for TM systems, which includes different aspects of medicinal plant research. The five main strategies proposed by WHO were:

- *Facilitating integration of traditional medicine/ complimentary alternative medicine TM/CAM into national health care systems* by helping Member States to develop their own national policies on TM/CAM.
- *Producing guidelines for TM/CAM* by developing and providing international standards, technical guidelines, and methodologies for research into TM/CAM therapies and products, and for use during manufacture of TM/CAM products.
- *Stimulating strategic research into TM/CAM* by providing support for clinical research projects on the safety and efficacy of TM/CAM, particularly with reference to diseases such as malaria and HIV/AIDS (acquired immunodeficiency syndrome).
- *Advocating the rational use of TM/CAM* by promoting evidence-based use of TM/CAM.
- *Managing information on TM/CAM* by acting as a clearing house to facilitate information exchange on TM/CAM.

These strategies were developed to meet four objectives with an emphasis on ensuring that appropriate policies and regulations are in place to support patient safety.

The strategy incorporates four objectives:

1. *Policy*. Integrate TM/CAM with national health care systems, as appropriate, by developing and implementing national TM/CAM policies and programs.

2. *Safety, Efficacy, and Quality*. Promote the safety, efficacy, and quality of TM/CAM by expanding the knowledgebase on TM/CAM, and by providing guidance on regulatory and quality assurance standards.
3. *Access*. Increase the availability and affordability of TM/CAM, as appropriate, with an emphasis on access to poor populations.
4. *Rational Use*. Promote therapeutically sound use of appropriate TM/CAM by providers and consumers.

Codrell (2012) provides a very good overview of the progress being made toward the delivery of these objectives. Although over 10 years old, all of the objectives are still relevant today and can provide a focus for collecting and studying those species that have been identified as having impact and have a wide use. Currently, there are about 3000 species of medicinal plants that are globally traded and tracked by the International Trade Centre. In 2008, the value of this market was $19.5 billion and it was expected to be at least $32.9 billion by 2013. With an increased international interest in OTC and investment from the pharmaceutical sector, this market will increase. A key to the magnitude of this increase will be dependent on advances in our knowledge of the uses of the plants so that more can be registered as OTC as well as the sustainable supply of good quality material. Thus, a clear case can be made for selecting these species that have been shown to have efficacy but require more research before they get to market.

If a species is to be included in an OTC product, then it is important that a portfolio of information be gathered that shows the traditional use of the plants and the proposed method of making the product link directly with the traditional method of preparation. The portfolio for registering an OTC will require information on the chemistry of the plants. Therefore, it is important that information is gathered about the chemical profile of a quality plant and the methods to be used to grow plants that provide good quality material. To gather this scope of information and provide a robust case requires access to samples of plants from different areas of the world it grows in as well as the expertise of pharmacologists, specialists in the target disease, natural product chemists, and horticulturalists. If plant-based medicine research are going to attract greater support from funders than increasing our knowledge about those already traded and the volume of trade will assist increase

interest and funding. The use of "onomics" does provide a new approach to furthering the scientific understanding of how and which compounds in the plants are having a beneficial effect on the body (Buriani *et al.*, 2012).

As indicated above, a strong case can be made for furthering our understanding of medicinal plants that are already in the international trade, but a case can also be made to study less well-known species, especially if they are used by local communities. The GSPC has 16 targets that support the conservation of plants that are relevant for medicinal plant research including targets to conserve the knowledge about the traditional uses of these plants. There are three GSPC targets that are highly relevant to the selection of plants and could provide a focus for research projects:

3.1 Objective IV

Objective C: Using plant diversity sustainably (www.cbd.int/gspc/targets.shtml)

1. Target 11: No Species of Wild Flora is Endangered by International Trade

Focus is placed on those species that are actually threatened by international trade and are listed by CITES appendix 1, but the aim would be to avoid increasing pressure on other species, so that if a species enters the trade, there is a sustainable source to meet the needs of the market.

2. Target 12: 30% of Plant-Based Products Derived from Sources that are Sustainably Managed

This has been classified to include

1. Plant-based products include food products, timber, paper, and other wood-based products; other fiber products; and ornamental, medicinal, and other plants for direct use.
2. Sources that are sustainably managed and include: Natural or seminatural ecosystems that are sustainably managed (by avoiding overharvesting of products, or damage to other components of the ecosystem), excepting that commercial extraction of resources from some primary forests and near-pristine ecosystems of important conservation value might be excluded. Sustainably managed plantation forests and agricultural lands.

3. Target 13: The Decline of Plant Resources, and Associated Indigenous and Local Knowledge Innovations and Practices, that Support Sustainable Livelihoods, Local Food Security, and Health Care, Halted

Plant diversity underpins livelihoods, food security, and health care. This target supports the international development targets, to "ensure that current trends in the loss of environmental resources are effectively reversed at both global and national levels by 2015." The target covers both plant resources and the associated ethnobotanical knowledge about the plants.

The majority of medicinal plants, especially those used by local communities, are wild harvested. If one of these plants was developed commercially as a product, then unless there was a cultivation program in place, there would be a real danger that the plants would be overharvested and thus become endangered. There are many examples of this happening. The most recent is *Hoodia gordonii*, a plant that belongs to the family Apocynaceae and is distributed in arid areas of South Africa and Namibia (van Heerden, 2008). It is a slow growing plant and the natural populations have mostly been lost due to the commercialization of products containing *H. gordonii* as an appetite-suppressant. It is difficult to follow the literature on the traditional uses of this species. It was first collected by Robert Gordon in 1779 and published by Masson as *Stapelia gordonii* Masson and was reclassified by Decaisne in 1844 as *H. gordonii* (Masson) Sweet ex Decne. Because of the complexity of linking the local names with the Latin names, it is difficult to validate the literature on the appetite-suppressing activity of *Hoodia*, although the use of *Hoodia* in quenching thirst is clearer (van Heerden, 2008). Owing to the increased interest in plants with antiobesity properties, the demand for products that contain plants such as *Hoodia* has increased, resulting in overexploitation of the limited amount of material been grown and an increased number of adulterants entering the market.

Thus, when a plant is being selected for commercial development, it is of critical importance that plans are quickly put in place for it to be grown sustainably. With *Hoodia*, emphasis was placed on getting material for cultivation in parts of South Africa and Namibia, but not enough effort was placed on making sure that the plants being selected had the potential to produce higher levels of the active compound, the triglycoside of 12β-tigloyloxy-14β-hydroxypregn-5-en-20-one, known as *P57*. Projects such as this example on *Hoodia* link directly with the delivery of Targets 12 and 13. Targets 11 and 12 also link with the WHO's declaration that identified the following six articles adopted by the WHO Congress on TM, Beijing, China, 8 November 2008 (www.who.int/medicines/areas/traditional/congress/beijing_declaration/en/):

- The knowledge of TMs, treatments, and practices should be respected, preserved, promoted, and communicated widely and appropriately based on the circumstances in each country.
- Governments have a responsibility for the health of their people and should formulate national policies, regulations, and standards, as part of comprehensive national health systems to ensure appropriate, safe, and effective use of TM.
- Recognizing the progress of many governments to date in integrating TM into their national health system, we call on those who have not yet done so to take action.
- TM should be further developed based on research and innovation in line with the "Global Strategy and Plan of Action on Public Health, Innovation and Intellectual Property" adopted at the 61st World Health Assembly in 2008.
- Governments, international organizations, and other stakeholders should collaborate in implementing the global strategy and plan of action.
- Governments should establish systems for the qualification, accreditation, or licensing of TM practitioners. TM practitioners should upgrade their knowledge and skills based on national requirements.
- The communication between conventional and TM providers should be strengthened, and appropriate training programs should be established for health professionals, medical students, and relevant researchers.

International strategies such as these can support the selection of a species of which some work has been done on their activity but more is needed if the species is going to enter the trade. The quality of material entering the trade and the regulations associated with monitoring the quality will vary from country to country. The standards of using are based on monographs that provide information about the

species, its known uses, chemistry, and safety. A key to getting more plants being traded will be to not only increase the diversity of species covered by monographs but also improve the quality of the monographs. Within China, there has been an increase in government funding for the improvement of their pharmacopoeia and the same is happening in India. In Africa, there are attempts to provide monographs for the rich diversity of African medicinal plants, but projects such as this need funds. The production of monographs is one way to help share data about developing best practice; this can drive research to address areas of work that need to be evaluated further and to improve the quality of plant entering the market. It can also provide a platform for innovation, as greater knowledge about the active compounds in the plants and their mode of action will enhance not only the use of the plants themselves but also the potential for the plants to drive innovative drug development programs as has occurred with Taxus, Narcissus, and Artemisia.

ACKNOWLEDGMENTS

REFERENCES

Buriani, A., Garcia-Bermejo, B. M. L., Bosisio, E., *et al.* (2012) *J. Ethnopharmacol.*, **140**, 535–544.

Chan, K., Shaw, D., Simmonds, M. S. J., *et al.* (2012) *J. Ethnopharmacol.*, **140**, 469–475.

Chen, S., Yao, H., Han, J., *et al.* (2010) *PLoS ONE*, **5**(1), e8613, DOI: 10.1371/journal.pone.0008613.

Cordell, G. A. (2012) New strategies for traditional medicine 1–45, in *Medicinal Plants: Biodiversity and Drugs*, eds. M. Rai, G. A. Cordell, J. L. Martinez, M. Marinoff and L. Rastrelli, CRC Press, Boca Raton, FL, p. 688.

Everett, J. R., Loo, R. L., and Pullen, F. (2013) *Ann. Clin. Biochem.*, **50**, 523–545.

Fabricant, D. S., and Farnsworth, N. R. (2001) *Environ. Health Perspect.*, **109**(Suppl. 1), 69–75.

Farnsworth, N.R., Akerele, O., Soejarto, D.D., and Guo, Z. G., (1985) *Bulletin of the World Haealth Organisation*, **63**, 965–981.

van Heerden, F. R. (2008) *J. Ethnopharmacol.*, **119**, 434–437.

Hsu, E. (2006) *Trans. R. Soc. Trop. Med. Hyg.*, **100**, 505–508.

Hsu, S. (2010) *Plants Health and Healing*, S.Harris (Eds.), Berghahn Books, Oxford, pp. 83–114.

van der Kooy, F. and Sullivan, S. E. (2013) *J. Ethnopharmacol.*, **150**, 1–13.

Magalhaes, P. M., Dupont, I., Hendrickx, A., *et al.* (2012) *Food Chem.*, **134**, 864–871.

WHO (2002) WHO Traditional medicine Strategy 2002–2005, www. Who.int/medicines/library/trm/trm_stat_eng.pdf.

Williamson, E., Okpako, D. T., and Evans, F. J. (1996) *Selection, Preparation and Pharmacological Evaluation of Plant Material*, John Wiley & Sons, Ltd, Chichester.

2

Extraction Methodologies: General Introduction

Valérie Camel

AgroParisTech, Paris, France

1 INTRODUCTION

Plant materials are of great interest in food, cosmetic, and pharmaceutical industries, as they generally contain numerous compounds that may afford interesting properties, being either bioactive or providing physical properties to the final product. Hence, plant materials contain numerous active compounds with high nutritional or therapeutic values (e.g., lipids, flavors, antioxidants, phytochemicals, pharmaceutics, and pigments).

Exploiting plant materials requires a multi-step procedure: sampling, pre-treatment of the sample, extraction of compounds of interest, further possible concentration or cleanup of the obtained extract, and final analysis of the compounds in the extract to ensure the quality of the final product before using, to authenticate the plant material or to identify new compounds. In addition, biological activity or sensory quality of the extract may be assessed. In this multi-step procedure, extraction can be considered to be a major key step as it will condition the quality and quantity of the compounds recovered, and will also greatly contribute to the time spent for the overall procedure.

Therefore, this chapter details the principle of extraction, enhancing also the several problems frequently encountered with this step. The different parameters to consider for choosing an extraction technique are discussed. The main extraction techniques available for plant materials are briefly exposed, with a view of helping the reader to choose the best technique appropriate to its application. Finally, future trends in extraction techniques for plant materials are discussed.

2 PRINCIPLE OF EXTRACTION

The basic principle of extraction refers to the distribution of a compound between two immiscible phases, enabling their further separation and recovery of the extracted compound. This process could appear simple and easily to achieve; yet, as entropy is gained only in mixing, not in separation, extraction can only be achieved by applying external work and allowing diffusion to be consistent with thermodynamics.

2.1 Basic Terminology Related to Extraction

Commonly, the compounds of interest are called the *solutes* or *analytes*; other compounds present with no interest are named *interferences* or *interfering compounds* (Figure 1). The initial plant material containing the compound to be extracted is named the *sample* or *sample matrix*. The nonmiscible phase is called the *extractant* or *the extraction solvent*, being generally
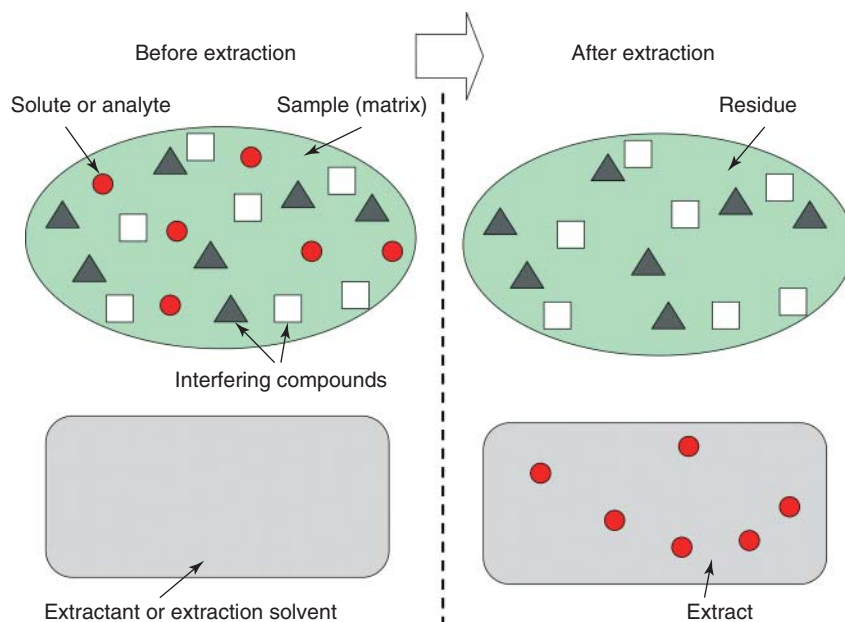
Figure 1   Scheme of the extraction process and its basic terminology.

a liquid solvent or a supercritical fluid; however, in some cases, it can be a solid phase (sorbent). Once the extraction step is performed, the sample is discarded from the extraction solvent containing the extracted solute (named the *extract*); the remaining sample matrix discarded is often referred to as the *residue*. In some cases, this residue is submitted again to one or more successive extractions (possibly under more stringent conditions), to recover part of the solute not extracted during the first step.

The main challenge of the extraction step is to selectively (i.e., without those interfering compounds) and possibly quantitatively (i.e., all the analyte molecules initially in the sample) extract the analytes. To achieve this goal, efficient contact between the sample and the extractant is required; as discussed later, this can be achieved by mixing, but other ways are also used. The extraction *technique* refers to the technological system used to perform the extraction; depending on the experimental conditions applied (e.g., solvent nature, extraction time, and temperature), the same technique will lead to several extraction *methods*.

The theoretical progress of an extraction is illustrated in Figure 2. As the extraction lasts (by increasing the time, increasing the solvent volume,
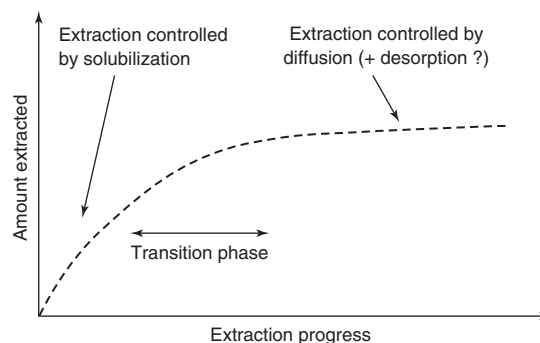
Figure 2   General extraction curve giving amount of extracted solute as a function of extraction progress (e.g., time).

or performing a new cycle), the extracted amount of the solute is increasing. The gain in amount extracted is quite fast at the beginning, where the process is governed by solubilization of the compound by the solvent (i.e., partitioning of the solute between the solvent and the sample); then as the extraction proceeds, the extracted amount levels off, as the step is limited by diffusion (and desorption from the matrix also in some solid samples).

2.2 Extraction from Liquids: A Simple Process

Extraction from liquids is rather simple, because of availability of solutes. The process is driven initially by partitioning of the solute between the liquid sample and the extractant (mostly a liquid in classical liquid–liquid extraction (LLE), or a sorbent in solid-phase extraction (SPE)), and the only requirement being that the extractant must not be miscible with the sample to ensure isolation of the extract. Then, as the extraction goes on, diffusion of the solute across the liquid–liquid boundary layer becomes limiting (in case of stirring, diffusion in the bulk phases can be neglected); it mainly depends on solute shape and size, as well as on solvent viscosity.

As liquid samples are scarce when dealing with plant materials, extraction methods dedicated to solid samples are mostly required for such samples.

2.3 Extraction from Solids: A Multi-Step Process

While extraction of compounds from liquid matrices is quite easily achieved because of the partition of the solutes between the extracting liquid and the matrix, extraction from solid matrices is more difficult. In that case, the compounds of interest must be released from the solid substrate to achieve their extraction by the solvent; thereby, diffusion inside the solid matrix may be the limiting step. In addition, when solutes are strongly adsorbed onto the matrix, the energy of interaction between the compounds and the matrix must be overcome; besides when the solutes are retained in plant organs or cells, extraction conditions must be strong enough to disrupt the cells; otherwise, the solutes will not be available for the extracting solvent.

The extraction and recovery of a solute from a solid matrix can be regarded as a five-stage process (Figure 3): (1) desorption of the compound from the matrix active sites; (2) diffusion into the matrix itself; (3) solubilization of the analyte by the extractant; (4) diffusion of the compound in the extractant present within the pore; and (5) transportation to the bulk extractant before final collection of the extracted solute. The obtaining of quantitative and reproducible recoveries necessitates the careful control and optimization of each step; in particular, collection of the extract needs to be carefully controlled as it is often neglected as compared to the extraction step, whereas it may lead to partial losses.

In practical environmental applications (e.g., extraction of pollutants from soils and sediments), the first step is usually the rate-limiting step, as solute–matrix interactions are very difficult to overcome and predict. However, for other matrices (e.g., plant materials), the rate may be limited by either the solubilization or the diffusion steps; in addition, compounds may be trapped in plant cells, thus being difficult to extract unless the cell membranes are disrupted. Consequently, the optimization strategy
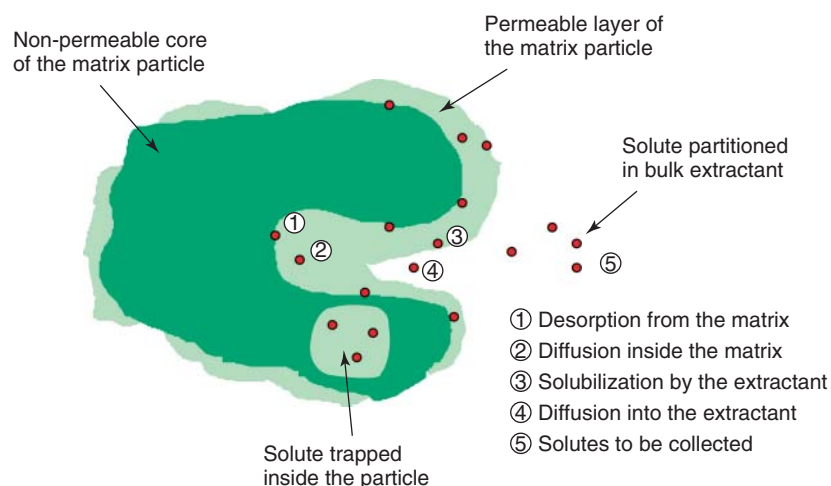
Figure 3 Illustration of the different successive processes involved during the extraction of solutes from a solid particle.

will strongly depend on the nature of the matrix to be extracted.

Several models have been proposed to describe extraction from solid matrices (following the curve presented in Figure 2). The most famous is the "hot-ball diffusion" (HBD) model developed by Bartle *et al.*, and detailed below:

$$\frac{m}{m_0} = 1 - \left(\frac{6}{\pi^2}\right) \sum \left(\frac{1}{n^2}\right) \exp\left(\frac{-n^2\pi^2 Dt}{r^2}\right)$$

where m is the mass of extracted solute at time t, m_0 the initial mass of solute in the sample, n an integer, D the diffusion coefficient of the solute in the sample matrix, and r the radius of the spherical particle of the sample matrix.

According to this model, during the initial phase of extraction (the solubilization-controlled phase), a concentration gradient exists at the surface of the particle, so that diffusion from the particle is rapid. As the extraction proceeds, the concentration across the entire particle becomes even, and the diffusion inside the particle controls the process. The HBD model is suggested for a matrix that contains small quantities of extractable solutes, so that the extraction is not limited by solubility (Huang *et al.*, 2012).

Another model has been proposed by Sovová for plant materials: the "broken plus intact cells" (BIC) model (Huang *et al.*, 2012). Solid particles with intact and broken cells arise from previous grinding of the sample because of partial break of cell walls; solutes from broken cells are assumed to diffuse directly to the extractant, whereas solutes from intact cells diffuse only to the broken cells. This model has been recently adapted to consider the plant microstructure (Sovová, 2012a).

In practice, diffusion in the solid sample is very limiting; as a result, the development of extraction methods for solids has focused on improving the diffusion step. Ideally, the final amount extracted should be the sample content for the solute of interest; however, this is not always observed in practice, as part of the solute molecules may be hardly recovered by the solvent under the experimental conditions used. One of the big challenges in optimizing extraction conditions is to find suitable conditions for quantitative extraction without hindering selectivity.

2.4 Types of Interactions Governing the Extraction Process

Knowledge of the intermolecular interactions prone to occur during the extraction step is required for proper optimization of the process conditions. Both interactions between the solutes of interest and the sample matrix as well as with the extracting solvent must be considered (Figure 4). In addition, interactions between the extractant and the sample should be considered, as they are crucial for achieving efficient and selective extractions. Consequently, as much knowledge as possible must be
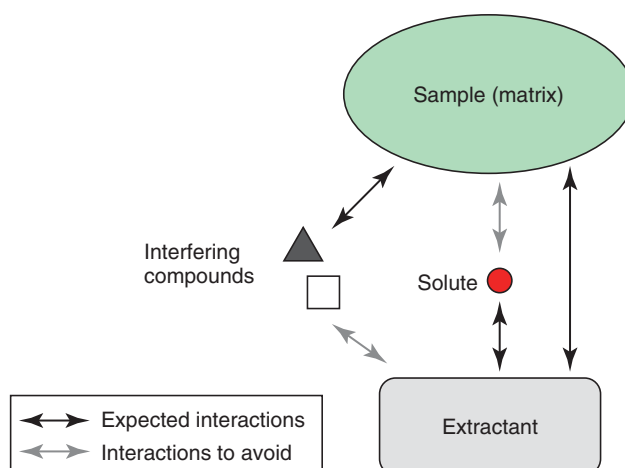
Figure 4   Scheme of the interactions to consider in extraction.
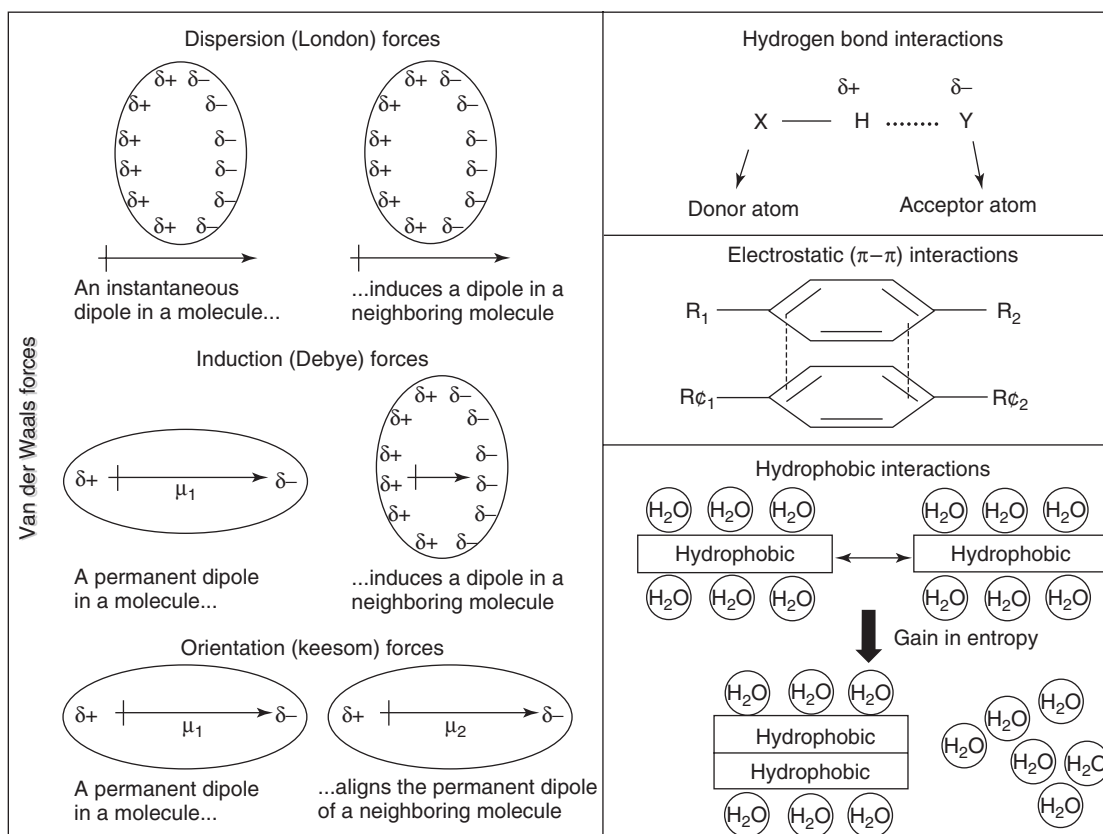
Figure 5   Main types of interaction encountered in extraction.

gained about the sample (chemical nature and content of major and minor constituents, chemical nature and content of solutes of interest along with their stability, and localization of the analytes inside the matrix).

The interactions occurring during extraction deal with noncovalent binding: mainly van der Waals interactions (orientation or Keesom interactions, induction or Debye interactions, and dispersion or London interactions depending on the dipolar moment of the solvent and the solutes), hydrogen bonding, electrostatic interactions (π–π electrons), and hydrophobic interaction. All these interactions are illustrated in Figure 5.

Ionic interactions are also possible for some ionic or ionizable solutes and some samples. Understanding the nature of interactions that will occur during the extraction process is very helpful in finding optimal experimental conditions.

3   FACTORS TO CONSIDER IN EXTRACTION

In order to develop adequate extraction conditions, several factors need to be considered before optimization of the experimental conditions; they are detailed below and gathered in Figure 6.

As for the same plant matrix the content of the obtained extracts can differ greatly depending on the extraction conditions, careful look at these factors before conceiving the extraction is advised.

3.1   What is the Objective of the Extraction Step?

A preliminary step is to clearly define the objectives dedicated to the extraction stage; indeed, different objectives may lead to different extraction techniques and/or methods.

Objective of the extraction step

Scale of the extraction process
Fate of the extract

Compounds to be extracted

Nature of the solutes
Levels of analytes in plant matrix
Target or non-target approach

Important factors to consider in extraction

Nature of the plant matrix

Localization of solutes
Nature and levels of other compounds
Water content

Pre-treatment of the plant matrix

Homogenization
Drying
Grinding

Figure 6   Important factors to consider for establishing a suitable extraction methodology.

The scope of extraction step with regards to plant materials is very large. One can consider different extraction strategies, depending on the fate of the extract, as well as on the type of extract considered. In particular, one key point is whether the plants to be extracted contain known compounds or unknown structures that are looked for, especially as thermal stability of the solutes may be important. Whatever the strategy conducted, one crucial point remains that the extraction step should avoid the degradation of initial compounds present as much as possible. Thus, the extraction temperature will be a major factor to control in order to avoid such degradations as recently reported for comprehensive plant metabolomics (t'Kindt *et al*., 2009).

3.1.1   Scale of the Extraction Process

The extraction stage can be performed at several scales: (i) analytical scale – in that case, the objective is to determine (both identification and quantification) the analytes extracted; (ii) preparative scale – here, the obtained extract will be further used, so it is required to get a sufficient amount (but quantitative extraction is not essential) of most of the time a class of extracted compounds, before further analysis of such compounds; and (iii) pilot scale – the

aim is to produce large amounts of the extracted compounds (again quantitative extraction is not a prerequisite, but it is obviously advantageous) for further applications (e.g., use as medicine). Considering the required scale is important as some techniques deserve major practical facilities or drawbacks at pilot scales. In addition, the production of plant extracted compounds may require the use of standard extraction procedures (e.g., the production of phyto-pharmaceuticals).

3.1.2   Fate of the Extract

Knowing the fate of the extract is important to choose the right extraction technique, as well as the proper experimental conditions. If the extraction is dedicated to isolate target solutes from the sample matrix, the extraction conditions will be highly selective and appropriate for such compounds. On the opposite, if the main goal is to characterize as much as possible all the compounds in the plant material (i.e., screening or profiling), the extraction conditions should be as mild as possible to cover numerous possible chemical structures (in some cases, several extraction steps should be performed). When the objective of the extraction step is to screen the biological activity of the plant material, the extraction solvent should be suitable for the biological

test to avoid evaporation; in addition, extraction conditions should be carefully optimized and controlled, as they greatly affect the composition of the extract and consequently its biological activity as recently reported for antioxidant or insecticidal activity from plant matrices (Michiels *et al*., 2012; Pavela *et al*., 2010). For example, herbal extracts, found in the pharmacopoeias of several countries, provide unique therapeutic properties because of possible synergistic effects between active molecules present (unlike synthetic drugs that contain a single active molecule); however, the multiple active constituents they contain are highly dependent on the extraction conditions.

Another point to consider is the possibility of performing a subsequent cleanup or filtration of the extract or not. As several techniques do require the extract to be filtered or centrifuged, this point needs to be considered when choosing an extraction technique.

3.1.3 Type of Information Investigated

The aim of the extraction may be to identify some compounds in the matrix (for characterization, authentication, etc.); that way, the information is only qualitative, and complete extraction of the solutes is not expected. On the opposite, when the extraction is aimed at giving indication about the analyte levels in the sample, extraction conditions that achieve complete recoveries are preferred. It must be pointed out that when dealing with analytes at trace levels, quantitative extraction will be expected to gain in sensitivity even for qualitative determinations.

3.2 Nature of the Plant Matrix

3.2.1 Structure and Composition of the Matrix

Getting knowledge about the composition of the sample matrix is a prerequisite before performing an extraction. Information about the nature and levels of major constituents of the sample will help in choosing an extractant that affords high selectivity. In particular, sample preparation methods are generally developed considering the matrix fat content,

the limit between "fatty" and "nonfatty" matrices being usually set at 2–5% (Gilbert-López *et al*., 2009).

Knowing the macroscopic structure of the matrix and the localization of the analytes inside is also very useful, especially if they are expected to be contained inside the cellular matrix of the plant material (e.g., essential oil). In that particular case, their extraction will require strong conditions capable of disrupting the cells to release the analytes into the extractant. Plant materials have been recently classified according to the behavior of their secretory structures during the extraction stage (Stamenić *et al*., 2008).

The water content of the sample is also interesting information. When dealing with hydrophobic analytes, water molecules may act as a barrier and prevent efficient extraction of solutes; therefore, reducing the water content of the sample, or even drying, may be advisable. In addition, when microwaves are used, sample heating will depend on its water content.

3.2.2 Pre-Treatment of the Matrix

Plant materials are mainly solid matrices that deserve most of the time pre-treatment before performing the extraction (Romanik *et al*., 2007). Indeed, recovery of solutes from solid matrices may be quite a challenge, depending on the initial matrix as discussed earlier, so that the pre-treatment step generally improves the extraction efficiency. Drying may be recommended as mentioned earlier. In addition, to enhance the solute availability, grinding finely the matrix is of prime importance to favor the extraction as this reduces the diffusion step inside the matrix and enhance specific surface area; in addition, grinding enables the breaking of some cells, and the subsequent release of solutes contained inside. However, too fine particles may lead to clogging, or to the formation of channels inside the sample that will prevent the solvent to penetrate through the sample matrix (Hyötyläinen, 2009; Pinelo *et al*., 2006; Zuknik *et al*., 2012). In some cases (e.g., extraction of proteins), breakage of the plant cells is required; this can be achieved by different means as discussed below (mainly sonication or enzymes) (Newton *et al*., 2004). On the other hand, this should be avoided in other cases (e.g.,

extraction of nonstructural carbohydrates) (Raessler, 2011).

3.3   Nature and Level of the Compounds of Interest

3.3.1   Nature of the Solutes

The knowledge of the chemical structure of the solutes of interest is of prime importance, to assess their physicochemical properties and to predict the type of interactions they may develop with both the sample matrix and the extraction solvent. In particular, assessment of their polarity is required to help in choosing the best solvent for their extraction. Besides, the knowledge of their volatility and stability (e.g., thermolability, susceptibility to be degraded in contact with air or when exposed to light) is also to be considered, as this may turn toward the right technique to use (e.g., volatile compounds should be preferably extracted using headspace (HS) techniques because of their high selectivity for such compounds as discussed later). In addition, when bioactive solutes are to be extracted, the nature of the solvent is crucial, as it may affect the bioactivity of the extracted solutes; in some cases, complete removal of the solvent will be necessary to avoid change in bioactivity.

In particular cases where the operator has no knowledge of the solutes to be extracted, then assumptions must be made according to the type of the sample matrix; then, extraction methods dedicated to several types of solutes depending on their properties (mainly polarity and volatility) must be put forward.

3.3.2   Levels of the Analytes in the Plant Matrix

Knowing the expected levels of the solutes in the sample is also helpful, as trace levels will require extreme cautious to minimize losses and contaminations through all the extraction step, and possible further concentration step. Whenever possible, extraction techniques that afford enrichment of the analytes will be preferred when dealing with trace levels solutes; that way, techniques that minimize the use of solvents (with subsequent extract dilution) should be preferred, as well as sorptive techniques as discussed below.

3.4   Assessing the Extraction Efficiency

3.4.1   Criteria to be Used

Frequently, the criterion used in evaluating the effectiveness of the extraction is extraction yield: this can be either the total yield of compounds extracted (typically the total mass extracted) or the yield for some target compounds (expressed in that case either as the mass extracted for these compounds or as the recovery in percentage if the initial contents of these compounds in the sample are known). As chromatographic techniques are mostly used for analyzing the extracts, yields are generally assessed based on chromatographic peak areas; it is important to point out that higher peak areas do not necessarily imply higher extraction efficiency, as interfering co-eluted substances may be present. For example, higher peak areas observed with methanol/water extraction as compared to dichloromethane for extracting xanthones and flavanones from the root bark of *Maclura pomifera* C.K. Schneid. were suspected to be due to unwanted compounds extracted with the former solvent (Teixeira and da Costa, 2005). Therefore, peak purity and solvent selectivity are critical to avoid bias in extraction efficiency assessment for target compounds.

When the extraction step aims at screening compounds, the qualitative composition of the extract is more appropriate than quantitative recovery of target solutes. In addition, when extracts are tested for their bioactivity, the response in the bioassay may be the preferred criterion.

3.4.2   Methods for Estimating Recoveries

Depending on the application, partial or quantitative recovery of the compounds of interest from the plant material is desirable. In addition, it is frequently of concern to assess the solute levels in the initial sample. Consequently, very often the extraction efficiency needs to be known. This may be quite a challenge when dealing with unknown samples.

Several ways may help in assessing the extraction recovery as illustrated in Figure 7 and detailed below.

Assessing recoveries of target compounds for samples spiked with known levels of these target compounds remains the simplest method; for that
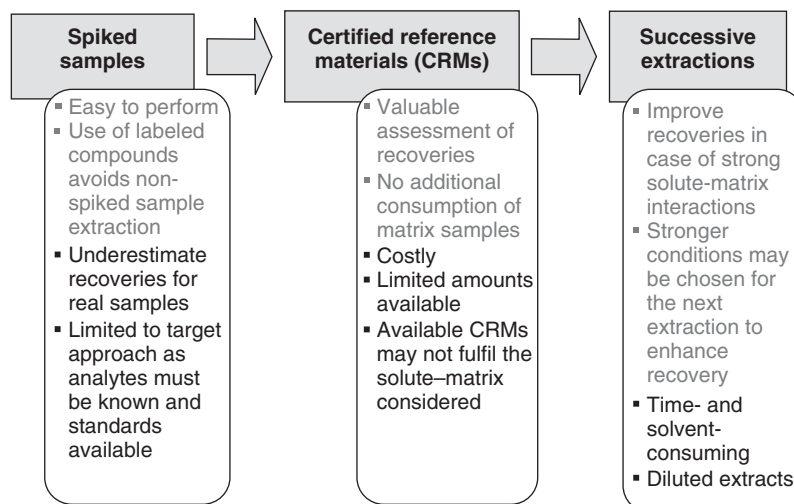
Figure 7 Successive steps to assess extraction efficiency. For each step, benefits are shown in *gray text*, whereas pitfalls are indicated in *black text*.

reason, it is undoubtedly the most used in practice. When nonlabeled compounds are used, additional extraction of nonspiked samples will give information about the background level of the matrix and should be considered to avoid overestimation of spike recoveries. This step is avoided in case of spiking with labeled compounds (using stable isotopes such as ^{2}H or ^{13}C), as native compounds are nonlabeled (because of the poor natural abundance of the labeled compounds). Yet, recoveries from spiked samples do not reflect recoveries from nonspiked samples when dealing with solid matrices, as solute–matrix interactions remain low for spiked solutes; in addition, as spiking is frequently performed a short time before performing the extraction, the spiked solutes do not have the time to deeply diffuse inside the matrix, making them more easily available by the extractant as compared to native solutes. Consequently, this method must be considered the first step, useful for choosing initial extraction conditions, but insufficient for optimizing conditions and assessing recoveries for real samples (Camel, 2001).

In order to better optimize extraction conditions, the use of certified reference materials (CRMs) is recommended. Such materials are either real samples or artificially spiked samples but that were aged, so that solute–matrix interactions in CRMs do mimic the interactions that will be encountered for real samples. That way, their extraction enables a viable assessment of recoveries for target compounds. Nevertheless, owing to the time and cost required for the preparation of CRMs, such materials are not available for any type of application; in addition, they are available in limited amounts and are highly expensive. Therefore, their use should be limited to the final optimization step of the extraction process, and/or to the validation of the developed extraction method. Examples of CRMs commercially available for plant materials are NIST-3251 (green tea leaves for catechins and xanthines), BCR-679 (white cabbage for trace elements), BCR-431 (Brussels sprout for vitamins), or B-MYC0852 (rye for ergot alkaloids).

In case no CRM is available for the application considered, the performance of successive extractions on the same unknown sample should give an indication of the completeness of the extraction. It may be advisable to progressively choose more drastic conditions for the successive extractions, to make sure that all the extractable solutes can be recovered. Such an approach is rather time-consuming and solvent-consuming; however, it has the advantage of being applied to the samples of interest. In addition, it can help in finding the optimal extraction conditions (e.g., temperature and time) for these samples.

4   CHOICE OF THE EXTRACTANT

The nature of the best solvent for a particular application is to be found; in practice, "like dissolves like." Thus, this is the first parameter to consider when developing an extraction method, as its choice will influence the qualitative composition of the extract. For convenience, the main properties to look for when choosing the solvent are depicted in Figure 8, with values indicated in Table 1 for most frequently used solvents and discussed later. It must be kept in mind that, once the solute is extracted by the solvent, most of the time, it must be further isolated from the solvent; in consequence, the solvent should also be chosen in order to facilitate this procedure.

4.1   Polarity and Selectivity of the Extractant

The solvent polarity is undoubtedly one of the main properties of the solvent to consider first. Its polarity should match that of the solute of interest. Low polarity solvents yield more lipophilic compounds, whereas medium-polarity solvents extract both apolar and polar compounds; water extracts mainly polar compounds. As an illustration, apolar solvents (e.g., hexane, cyclohexane, toluene, and diethyl ether) are recommended for extracting alkaloids, terpenoids, fatty acids, or flavonoids from plants; for alkaloids, flavones, polyphenols and, saponins, polar solvents (e.g., ethanol, acetone, and acetonitrile) are better used (Puri *et al*., 2012). When mixing solvents that are miscible, the polarity of the mixture can be calculated considering the volume fraction of each solvent in the mixture and their own polarity. Solvent polarity is also important to consider when microwave radiations are used to enhance the extraction, as it will have an effect on radiation absorption by the solvent.

Yet, besides polarity, selectivity of the solvent is also important in case of expected selective extraction of target compounds, to avoid further cleanup of the extract as much as possible. For that purpose, solvents have been classified on polarity scales, considering the type of interactions they may develop with solutes as reviewed (Barwick, 1997).

The Rohrschneider–Snyder scheme classifies solvents according to their chromatographic strength (P') also called "*polarity*." It is based on gas–liquid partition data for several test solutes in 81 liquids; values of partition coefficients were further used to calculate polarity P' and selectivity values x_i (Barwick, 1997). Selectivity here considers all types of solvent–solute interactions except dispersive interactions (as it lies on interaction with test solutes that are acidic, basic, or dipolar): x_e for the ability to behave as proton acceptor, x_d for the ability to behave as proton donor, and x_n for the ability to behave as a strong dipole. Values assessed for common solvents are indicated in Table 1. Solvents with the same functionality were classified in the same group; that way, all solvents and groups of solvents could be plotted on a "solvent selectivity triangle" (also called the *Snyder's triangle*). This is very useful for choosing the solvent with the best selectivity suited for an extraction (Nyiredy, 2004). In case of different functional groups within a molecule, the more polar is supposed to dominate the selectivity characteristics. Similarly, in case of a solvent mixture, the selectivity is determined largely by the selectivity of the more polar solvent.

Another classification was also developed, and it is now the most used in practice (Barwick, 1997). It is based on the Hildebrand solubility parameter
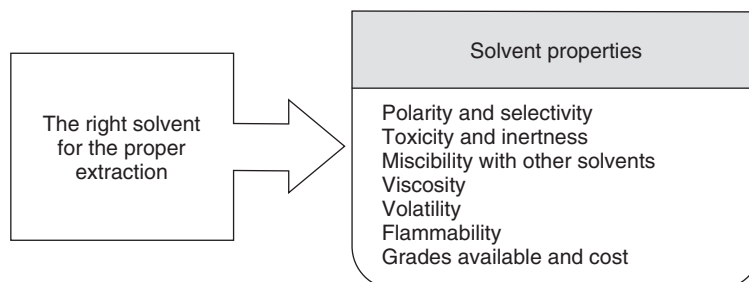
Figure 8   Solvent properties to consider in extraction.

Table 1 Properties of common liquid solvents used in extraction (at 20–25°C).

Solvent	Density ($g\,cm^{-3}$)	Boiling point (°C)	Viscosity ($10^{-3}\,Pa\,s$)	Dielectric constant	Dipolar moment (Debye)	Molar volume ($cm^3\,mol^{-1}$)	P'	x_e	x_d	x_n	δ_t ($MPa^{1/2}$)	δ_D ($MPa^{1/2}$)	δ_P ($MPa^{1/2}$)	δ_H ($MPa^{1/2}$)
n-Hexane	0.65	68.7	0.31	1.88	0.08	130.6	0.1	—	—	—	14.9	14.9	0.0	0.0
Isooctane	0.69	99.2	0.50	1.94	0	165.1	0.1	—	—	—	14.3	14.3	0.0	0.0
Cyclohexane	0.78	80.7	1.00	2.02	0	107.9	0.2	—	—	—	16.8	16.8	0.0	0.2
Toluene	0.86	110.6	0.59	2.38	0.36	105.9	2.4	0.25	0.28	0.47	18.2	18.0	1.4	2.0
Diethyl ether	0.71	34.5	0.24	4.27	1.15	104.4	2.8	0.53	0.13	0.34	15.5	14.5	2.9	4.6
Dichloromethane	1.33	39.8	0.44	8.93	1.14	63.9	3.1	0.29	0.18	0.53	20.2	18.2	6.3	6.1
Tetrahydrofuran	0.89	66.0	0.55	7.58	1.75	81.0	4.0	0.38	0.20	0.42	19.5	16.8	5.7	8.0
Ethanol	0.79	78.3	1.10	24.5	1.69	58.3	4.3	0.52	0.19	0.29	26.5	15.8	8.8	19.4
Ethyl acetate	0.90	77.1	0.45	6.08	1.88	97.9	4.4	0.34	0.23	0.43	18.2	15.8	5.3	7.2
Acetone	0.79	56.3	0.36	21.1	2.69	73.5	5.1	0.35	0.23	0.42	19.9	15.5	10.4	7.0
Methanol	0.79	64.7	0.59	33	2.87	40.6	5.1	0.48	0.22	0.31	29.4	14.7	12.3	22.3
Acetonitrile	0.78	81.6	0.38	36.6	3.44	52.6	5.8	0.31	0.27	0.42	24.4	15.3	18.0	6.1
Acetic acid	1.05	117.9	1.12	6.2	1.50	57.1	6.0	0.39	0.31	0.30	21.4	14.5	8.0	13.5
Aniline	1.02	184.0	3.77	6.9	1.55	91.3	6.3	0.32	0.32	0.36	23.7	20.1	5.8	11.2
Dimethylformamide	0.94	153.0	0.92	37	3.86	77.0	6.4	0.39	0.21	0.40	24.9	17.4	13.7	11.3
Dimethyl sulfoxide	1.10	189.0	2.24	47	3.96	71.0	7.2	0.39	0.23	0.39	26.7	18.4	16.4	10.2
Water	1.00	100.0	1.00	78	1.85	18.0	10.2	0.37	0.37	0.25	47.8	15.5	16.0	42.3

Source: Data compiled from Barwick (1997), Lü *et al.* (2008), and Villmow *et al.* (2012).

δ_t (introduced by Hildebrand and Scott in 1950), defined initially for volatile compounds, as indicated below (Hansen, 2004):

$$\delta_t = \left(\frac{E^v}{V_m} \right)^{1/2}$$

where E^v represents the cohesion energy (or the latent energy of vaporization for volatile compounds) of the solvent, and V_m its molar volume.

This parameter considers the effect of molecular size on relative solubility. In particular, the larger the molecule volume is, the greater will be the effect of a change in solvent polarity on the solubility of the compound. This solubility parameter can be decomposed into several factors, which provide information about the type of interactions the solvent will be able to establish with the solutes. Four selectivity parameters were first introduced; Hansen (2004) later simplified to three parameters as indicated below:

$$(\delta_t)^2 = (\delta_D)^2 + (\delta_P)^2 + (\delta_H)^2$$

where δ_D accounts for atomic nonpolar interactions (or dispersion interactions), δ_P for molecular dipolar interactions, and δ_H for molecular hydrogen bonding interactions.

That way, solvents can be plotted in 3D graphs (called the *Hansen* "*room*"), similarly as for the Snyder's triangle. Values of Hildebrand solubility and its selectivity terms are given in Table 1 for common solvents; they are established at 25 °C but correlations with temperature have been proposed as higher temperatures decrease their values (Villmow *et al.*, 2012). Solute solubility in a solvent is maximal when the solute and the solvent have the same δ_t value. Thus, it is possible to mix two solvents, with respective higher and lower δ_t values than that of a given solute, to give a solvent mixture with a δ_t value equal to that of the solute. To assess the Hildebrand solubility of a solute, the molecule can be divided into different chemical groups, and the contribution of each functional group can be estimated for each selectivity term according to existing tables (Lü *et al.*, 2008; Stefanis and Panayiotou, 2008).

The Hildebrand solubility parameter has also been determined for supercritical fluids (δ_{SF}), so that its use is very convenient whatever the extractant. The following relation is well established for supercritical fluids:

$$\delta_{SF} = 1.25 P_c^{1/2} \left(\frac{\rho_{SF}}{\rho_{liq}} \right)$$

where P_c is the critical pressure of the supercritical fluid, ρ_{SF} its density under supercritical state, and ρ_{liq} the density of the fluid under liquid state.

This clearly explains the great interest that deserves supercritical fluids: their polarity can be adjusted according to the pressure and temperature conditions, as they affect the density value. Thus, such extractants offer greater capability for selective extractions than do liquid solvents. For further details, see **Supercritical Fluid Extraction**.

To conclude, the Hildebrand solubility parameter is very useful for a preliminary prediction of the best solvent suited for a particular extraction as observed in several applications (Lang *et al.*, 2005; Srinivas *et al.*, 2009). However, the solvent efficiency must be tested experimentally, as strong solute–matrix interactions may limit the predictability of the Hildebrand parameter.

4.2 Toxicity and Inertness

Solvent toxicity is of prime importance to avoid exposure to highly toxic solvents for the operator who will manipulate the solvents for the extraction. In addition, toxicity of the solvent of the final extract needs to be considered when the final extract deserves some applications; in particular, plant extracts that are further used for food or medicine are submitted to restrictions with regards to the possible solvents used. In addition, when the final extract needs to be tested for bioactivity or sensory attributes, the solvent must be as low toxic as possible.

The inertness of the extractant is also crucial, to avoid changes in the analytes (especially degradation, dehydration, isomerization, or derivation) during the extraction process, as well as to limit damages to the extraction system. Consequently, very reactive extractants should be banned; this is the case for some organic solvents, and also with some supercritical fluids such as water (being highly corrosive under supercritical conditions). In addition, several solvents should be avoided for particular applications; for example, methanol may induce methylation of some solutes, and acetone the formation of acetonides. In a similar way, the pH of aqueous media should be carefully controlled to avoid degradation of compounds sensitive to pH changes; as an illustration, epimerization and hydrolysis of glycosides under acidic conditions have been observed.

4.3 Other Properties

4.3.1 Miscibility

Solvent miscibility is important to consider not only to choose a solvent nonmiscible to the sample in case the latter is liquid but also to investigate solvent mixtures that are miscible in all proportions because, as discussed earlier, solvent mixtures are frequently needed to get the best extractant for a particular application, based on polarity and selectivity.

4.3.2 Viscosity

Solvent viscosity η is of great importance, and should be as low as possible. The higher the viscosity is, the lower will be the diffusion coefficients (D) of the solutes for a given temperature (T) as expressed by the Stokes–Einstein equation:

$$D = \frac{10^{-7}T}{\eta V_{m}^{1/3}}$$

In addition, viscous solvents are less efficient to penetrate inside pore matrix, so that extraction of solid matrix is hindered in that case.

4.3.3 Volatility

Solvents with extreme volatility should be avoided, because of practical difficulties and also because of their flammability. On the other hand, solvents should be volatile enough to ensure easy and efficient removal of the solvent for further isolation of the extracted solutes (or extract enrichment in case of trace compounds); this is generally performed by evaporation or distillation.

4.3.4 Flammability

Nonflammable solvents should be used as much as possible to avoid hazards for the operator and extra safety measures. Flammable solvents have flash points below 37.8 °C; this is the case of some common solvents, such as acetone, diethyl ether, or toluene.

4.4 Grades Available and Cost

It is generally required to use solvents of high grade quality, to minimize the presence of impurities in the final extract, especially as they will be concentrated in case of further concentration step. Of course, the higher the quality standard of the solvent is, the higher the cost will be.

5 EXTRACTION TECHNIQUES FOR PLANT MATERIALS

Numerous techniques can be used to extract compounds from plant materials as illustrated in Figure 9. Most of them operate with batch systems, where the entire sample material is loaded into the extraction device. Then, depending on the system, the sample and extracting phase are brought into contact under either static or dynamic conditions (in the latter case, the extractant generally flushes continuously through the sample with a fixed flow rate) (Hyötyläinen, 2009).

Basically, one can distinguish the so-called *traditional techniques*, which have been used for years; they afford simple and low cost equipment; however, on the other hand, they are generally time- and solvent-consuming, with sometimes lower extraction efficiency than recent extraction techniques. The latter refers to techniques that have appeared in recent 20 years, with a view of reducing both the time and the solvent volume required for the extraction, along with providing similar or even higher efficiency (Kaufmann and Christen, 2002). A gain in capability to automation has also been observed with these recent techniques developed as discussed later. For illustration, see the reviews related to extraction methods for botanicals and herbal preparations (Ong, 2004) as well as advances in extraction of nutraceuticals from plants (Wang and Weller, 2006), and techniques dedicated to pesticide extraction from cereals and derivatives or isoflavones from soybeans and soy foods (González-Curbelo *et al.*, 2012; Pareja *et al.*, 2011; Rostagno *et al.*, 2009).

One major point lies also in the need to cleanup complex extracts recovered. For that purpose, several techniques may be used, such as the traditional LLE or the more recent SPE (Romanik *et al.*, 2007). As the aim of this chapter is to focus on the extraction stage for plant materials, techniques dedicated to the cleanup are not presented here.
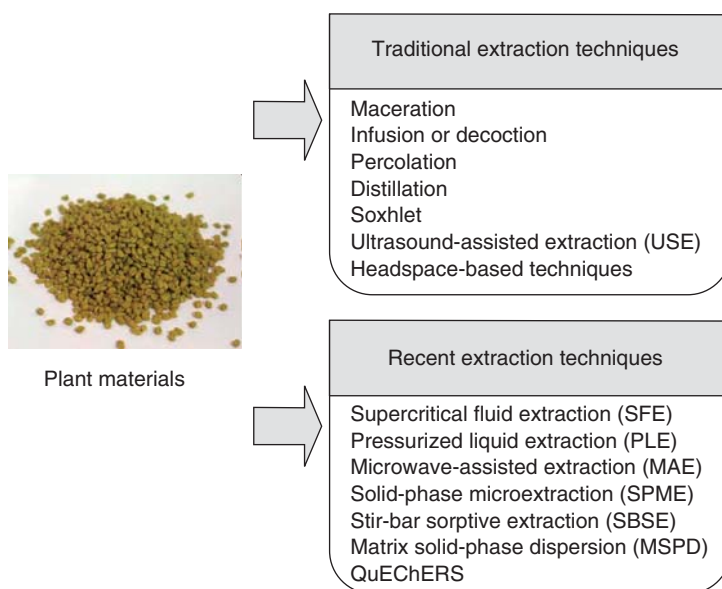
Figure 9   Traditional and recent extraction techniques available for plant materials.

5.1   Traditional Extraction Techniques

5.1.1   Maceration

Maceration is the simplest extraction technique dedicated to plant materials. After grinding, the plant sample is put in contact with the extracting solvent, and allowed to stand for hours or days for equilibration under ambient temperature. Shaking the vessel can be performed a few times over the extraction stage, to facilitate the process; placing the vessel on a heat source (such as a heating plate or steam bath) can also improve the extraction as observed for extracting phenolics from buckwheat herb (Hinneburg and Neubert, 2005), or pesticides (i.e., paraquat and diquat) from cereals (Kolberg *et al*., 2012).

This technique remains long and is not very efficient for extracting compounds that are tightly bound to the matrix or those that are contained in tissue cells; besides, it requires a further filtration or centrifugation step to separate the extract from the sample matrix. Consequently, its use remains today limited to particular applications; its mild conditions are suitable for volatile and thermally labile compounds. It remains particularly used for extracting phenolic compounds or antioxidants from plant materials, using, for example, dichloromethane, ethanol, methanol, or methanol/water mixture as the solvent

(Li *et al*., 2007; Teixeira and da Costa, 2005). Its application has been also recently reported for extracting cyclotides from plants with hydroalcoholic solutions (Yeshak *et al*., 2012).

5.1.2   Infusion or Decoction

Extraction with hot water is called *infusion* or *decoction* depending on the implementation. In infusion, water (boiling or cold) is added to the milled sample; in decoction, the sample is boiled in water for 15–30 min. These are the traditional extraction methods for medicinal herbs or tea, for example; the extracts are further directly consumed by humans. As an example, decoction was found efficient in extracting antioxidants from several medicinal plants (Li *et al*., 2007).

5.1.3   Percolation

In percolation, the sample matrix is poured into a conic flask; the extraction solvent is further percolated through the matrix bed. A daily domestic illustration is the filter coffee preparation. In practice, fine powders and plants may swell, thereby resulting in clogging of the system.

5.1.4 Distillation

Distillation can be conducted either directly on the plant material or after water addition; it is dedicated to essential oil extraction. The sample, mixed or not with water, is submitted to direct heating or to water steam. The resulting vapors are cooled and collected in a separator, where essential oil separates from water. Most of the time, the crude oil obtained requires further distillation for cleanup. Hydrodistillation is the most usual method for the determination of essential oil (indeed, the essential oil of a plant is internationally defined as the volatile fraction of the plant obtained by hydrodistillation, steam distillation, dry distillation, or a suitable mechanical process without heating plant materials); its application to phytochemicals has been recently reviewed (Zhao *et al*., 2011). This technique presents several drawbacks. Firstly, it is time consuming (e.g., 3 h). In addition, its use for the determination of aroma compounds composition has been criticized because of transformation of aroma-active compounds due to heating, steam, or pH. Besides, nonvolatile aroma compounds are not extracted, and highly volatile and water-soluble compounds can get lost (Ormeňo *et al*., 2011; Richter and Schellenberg, 2007).

5.1.5 Soxhlet

The Soxhlet extractor is one of the oldest systems; it was conceived in 1879 by Franz von Soxhlet for the extraction of fats from milk (Luque de Castro and Priego-Capote, 2010). Basically, it consists of a porous cellulose cell (also called *a thimble*) where the solid sample is placed (after being previously ground); this cell being further put into the extractor made of glass (i.e., the thimble holder). The organic solvent, placed in a flask at the bottom of the extraction chamber, is heated up to its boiling point; its vapors flow in a lateral tube, before being condensed at the top of the extractor. That way, the liquid solvent progressively fills the thimble, being in contact with the solid sample for extraction. Once the extractor is almost filled by the condensed solvent, a siphon aspirates all the liquid present in the thimble-holder, allowing its return to the flask at the bottom. Indeed, the main feature of this system is the recycle of the extracting solvent over time, avoiding possible saturation of the solvent during the maceration step as

classically observed upon simple contact between the sample matrix and the solvent, and displacing transfer equilibrium so that high extraction yield is expected. That way, the extraction proceeds as a batch process with discontinuous solvent infusion: static extraction when the liquid fills the thimble, along with recirculation of the solvent through all the process.

Despite severe drawbacks (no stirring, long extraction times – several hours, large solvent volumes, lack of selectivity, and possible degradation of extracted solutes due to long heating at the boiling point of the solvent), Soxhlet extraction is still frequently used in many applications (such as the extraction of natural antioxidants from plant materials) because of its simplicity. One key point is that the system can be left unattended, and multiple extractions can be simultaneous performed using several systems.

Developments have been proposed in recent years to overcome some of the main drawbacks of the classical Soxhlet; they are deeply discussed in a recent review (Luque de Castro and Priego-Capote, 2010). Briefly commercial systems have appeared that enable heating of the extraction chamber in order to speed the extraction process, either by conventional heating (in the Soxtec system) or by microwave heating (in open-vessels microwave-assisted extraction systems). However, such systems increase the equipment cost and reduce the flexibility of the Soxhlet, so that the classical system remains frequently used.

5.1.6 Ultrasound-Assisted Extraction

Ultrasounds are acoustic vibrations with frequencies above 20 kHz; their use for assisting extraction started in the 1950s. Extraction assisted by ultrasounds (also called *sonication*) consists basically of performing maceration under energetic conditions that enhance diffusion and disruption of solute–matrix interactions. Vibrations are transmitted through the liquid and bubbles with negative pressures are formed (phenomenon called *cavitation*); their implosion creates high pressures and temperatures locally in the system (Bendicho *et al*., 2012). In practice, two types of systems can be used: either the ultrasonic bath or the ultrasonic probe. Both systems are efficient, with their own drawbacks (Ramos,

2012). Baths are more widely used, but they face a lack of uniformity of the distribution of ultrasound energy along with a decline of power over time. Probes ensure energy to be focused on a localized sample zone, so that cavitation is more efficient; on the other hand, the large amount of heat generated requires cooling of the vessel, and volatiles may be lost under such conditions. Reactors designed for industrial-scale applications have also appeared (Vilkhu *et al*., 2008; Vinatoru, 2001).

Ultrasounds may exert several effects on plant cell walls as observed using microscopic examination (Huie, 2002; Shirsath *et al*., 2012; Toma *et al*., 2001; Vilkhu *et al*., 2008; Vinatoru, 2001): (i) they can destroy the thin skin of external glands, thus facilitating release of their content (e.g., essential oil) into the solvent and (ii) they cause enlargement of the cell wall because of a better swelling and hydration of plant materials, so that mass transfer is improved.

Ultrasound-assisted extraction (USE) has been reported for several compounds from plant materials, such as phenolic compounds, antioxidants, oils, proteins, fats, polysaccharides, dyes and pigments, saponins, and medicines (Shirsath *et al*., 2012; Vilkhu *et al*., 2008). For example, USE has been recently reported to offer a significant (13–100%) improvement in dye extraction efficiency from plant materials as compared to magnetic stirring at 45 °C (Sivakumar *et al*., 2011). However, as temperature increases can be high locally, especially with probe sonication, USE may cause decomposition or oxidation of compounds (Bendicho *et al*., 2012). In addition, a single extraction is generally insufficient for achieving quantitative extractions, so that successive extractions are generally performed; this induces consumption of large solvent volumes and diluted extracts. In practice, variations in extraction yield have been observed as a result of differences in plant structure, rheology, or composition.

In recent years, several developments have been reported, and a gain in interest toward the use of USE can be noted as recently reviewed for natural products (Shirsath *et al*., 2012).

5.1.7 *Headspace-Based Techniques*

When volatile compounds are the sole compounds of interest (e.g., essential oil), the use of HS techniques is highly recommended because of their selectivity toward such compounds. Several systems exist. The most common are mentioned below; for better detail about strategies for volatile compounds, see **Analytical Strategies for Multipurpose studies of a plant volatile fraction**.

The simplest operates under static conditions and is called "*HS*"; it is most of the time directly coupled to a gas chromatograph for analytical purposes, to avoid delicate handling of gaseous phases. HS is nonquantitative in nature, as volatile extraction is limited by solute partitioning between the sample and the HS. Other systems operate under dynamic conditions that allow a better extraction of volatiles; in that case, the sample HS is continuously flushed with an inert gas and entrained out of the sample vial through a sorbent bed where the extracted volatiles are retained (sometimes cooling the trap is necessary to retain highly volatile compounds). Then, volatiles are recovered from the sorbent bed by thermal desorption. Continuous flushing of the solutes favors the volatile partitioning in the gaseous phase, so that quantitative extraction may be obtained in some cases.

5.2 Recent Extraction Techniques

The traditional techniques face now severe competition because of the development of new extraction techniques over recent 20 years, such as supercritical fluid extraction (SFE), pressurized liquid extraction (PLE), and microwave-assisted extraction (MAE) (Camel, 2001, 2002; Dean and Xiong, 2000; Huie, 2002). As extraction of solids is mainly limited by diffusion processes, such techniques have merged with a view of speeding up the diffusion step, along with reducing the solvent volumes used because of their cost and toxicity. As the extraction temperature is a crucial parameter in diffusion, these recent techniques afford the opportunity of elevated temperatures. In that way, rapid and efficient extractions can be performed, because of enhanced solubilities, higher diffusion coefficients, lower viscosities, and better desorption of the solutes from the matrix (when solutes are stable under high temperatures).

5.2.1 *Supercritical Fluid Extraction*

The first technique that merged is SFE; its development occurred in the late 1980s. Its particularity lies

in the use of an extractant being in its supercritical state, which means that both pressure and temperature are above their critical values for this extractant. Supercritical fluids possess unique properties, intermediate between those of gas and liquids, which depend on the pressure, temperature, and composition of the fluid (Sairam *et al*., 2012). In particular, their viscosity is lower than those of liquids, and the diffusion coefficients are higher, enabling more efficient extractions. Besides, the density (and therefore the solvent power of the fluid) may be adjusted by varying both the pressure and the temperature, affording the opportunity of theoretically performing highly selective extractions as mentioned earlier. For deeper details about this technique, see **Supercritical Fluid Extraction**.

SFE proceeds with batch systems functioning in either static or dynamic modes (frequently mixed modes are used: an initial static phase is followed by a dynamic phase). Isolation of the extracted solutes is easily obtained upon depressurization of the fluid (but in some cases, restrictor clogging may occur, especially when fresh plant materials are used because of ice formation). With the use of carbon dioxide, the deleterious effects (i.e., cost, toxicity, and environmental concern) of organic solvents can be minimized or even avoided. SFE also affords the possibility of fractionating the extract.

Several reviews detail this technique and its applications, especially related to plant materials (Fornari *et al*., 2012; Lang and Wai, 2001; Sovová, 2012b). Despite unique selectivity of supercritical fluids and major advantages of this technique, SFE has faced severe competition of other recent techniques over years, mainly because of the high dependence of the extraction conditions with the sample, leading to fastidious optimization procedures and difficulty when using this technique routinely; the relatively high investment costs associated with the high pressure process are also responsible for this limited use of SFE. Nevertheless, SFE is still widely used for particular applications related to plant materials (such as extraction of essential oils or bioactive compounds) (Fornari *et al*., 2012; Huie, 2002; Pourmortazavi and Hajimirsadeghi, 2007), especially as this technique is suitable for industrial scale and as manufacturing costs may be competitive (Meireles, 2003). For example, nowadays industrial processes using supercritical CO_2 are used to produce decaffeinated coffee beans and tea leaves or hop extracts. In addition, SFE offers unique selectivity and avoids degradation of extracted solutes because of moderate temperature and absence of exposure to atmospheric oxygen.

5.2.2 Pressurized Liquid Extraction

The limits of SFE, especially the use of mainly carbon dioxide as the supercritical fluid, thereby hindering the simple transposition of classical extraction methods using organic solvents, has put forward the development of a new technique capable of performing extraction under high temperature but still using classical organic solvents: the PLE, also called *pressurized solvent extraction* (*PSE*). This technique has merged in 1995, and was developed by the Dionex Corporation under the trademarked name of its commercial systems: accelerated solvent extraction (ASE) (Camel, 2001; Kaufmann and Christen, 2002).

Commercial PLE systems operate under batch conditions and static extractions. The sample is placed into a sealed container (with frits at both ends, ensuring the solid matrix to stay in the cell), and the liquid solvent is pumped in the cell to fill the void volume. In such systems, classical organic solvents can be used under moderate to high temperatures (up to 200 °C); the liquid state is maintained under elevated temperatures upon application of a moderate pressure. With commercially available systems, extractions are performed in static mode only; yet, some SFE systems have been adapted to be used with solvents, enabling PLE to be performed in dynamic mode (Santos *et al*., 2011).

Optimization of extraction conditions is facilitated as compared to SFE as organic solvents recommended in traditional techniques may be most of the time used and pressure has a low influence. In addition, developed systems offer a high level of automation. Yet, only one extraction at a time can be conducted, as in the case for SFE.

Application of PLE to organic contaminants and bioactive and nutritional compounds from plant origin food has been recently reviewed (Mustafa and Turner, 2011; Sun *et al*., 2012). Dispersion of the sample matrix in a drying or inert sorbent (such as sodium sulfate or diatomaceous earth) is frequently performed before packing the resulting mixture in the extraction cell in order to prevent aggregation of particles or channel formation; for example,

diatomaceous earth has been recently reported for PLE of isoflavones from pulses (Delgado-Zamarreño *et al*., 2012), and glass beads for alkaloids from *Macleaya microcarpa* Fedde (Urbanová *et al*., 2012). Adding a sorbent layer just after the sample layer in the cell may ensure *in situ* cleanup as reported for the extraction of polyhalogenated pollutants from plants (Pérez *et al*., 2013).

Compared to traditional extraction techniques, PLE offers similar (or even higher) extraction yields, with savings in time (5–15 min classically) and solvent volumes as reported for medicinal plants (Benthin *et al*., 1999). In addition, the setup of PLE equipment provides protection for oxygen- and light-sensitive compounds; on the other hand, thermolabile compounds may suffer from elevated extraction temperature in some cases (Urbanová *et al*., 2012). For deeper details about this technique, see **New Trends in Extraction of Natural Products: Microwave-Assisted Extraction and Pressurized Liquid Extraction**.

5.2.3 Microwave-Assisted Extraction

MAE merged in the late 1980s (Kaufmann and Christen, 2002). Instead of heating the extracting solvent by traditional convection as performed in classical maceration, the novelty here lies in the use of microwaves. Microwaves are electromagnetic radiations with a frequency from 0.3 to 300 GHz; microwave sources for extraction systems typically operate at 2.45 GHz. For the microwave irradiation to be absorbed, polar compounds must be present; this can be either the solvent or the water content of the sample depending on the application. This results in a very fast and efficient heating of either the solvent or the sample depending on their nature, owing to the particular effects of microwaves on matter (namely dipole rotation and ionic conductance) (Camel, 2000).

Several MAE systems have been proposed (Camel, 2001; Dean and Xiong, 2000). With open vessels, a system was developed that was quite similar to a Soxhlet assisted by microwaves (Luque de Castro and Priego-Capote, 2010). When open vessels are used, solvents with low dielectric constants (i.e., essentially microwave-transparent) are generally preferred to avoid high temperatures, so that extraction conditions are rather mild; that way, extraction of thermally labile compounds may be possible. On the opposite, with closed vessels, solvents with high dielectric constants are preferred that therefore absorb microwave radiations (vessel material is mostly Teflon, which is transparent to microwaves); under such conditions, samples are submitted to elevated temperatures and pressures. Therefore, in such systems, the principle is quite similar to the one of PLE, as the pressure inside the vessels ensures that the solvent remains in its liquid state as the temperature is elevated. Commercially available closed-vessel systems (pressurized MAE) are interesting as they afford the opportunity of performing several simultaneous extractions; however, after extraction, a delay is necessary for the pressure and temperature to decrease in the vessel before opening it. Apart from increasing the total time, this step can result in readsorption of extracted analytes onto the sample matrix. Besides, further filtration is required to obtain the final extract.

MAE faces a great interest for extracting plant materials, especially as the water content of the sample ensures rapid efficient heating (Kaufmann and Christen, 2002). Then, superheating inside the plant cells causes liquid vaporization within the cells, which may disrupt the cell walls and/or plasma membranes. That way, the mass transfer of plant compounds (especially plant secondary metabolites) into solvent as well as of solvent into plant materials is facilitated, allowing efficient extraction. However, thermo- or oxygen-sensitive compounds can undergo degradation or oxidation under classical MAE conditions; vacuum MAE has been recently developed to ensure extraction of sensitive analytes as illustrated with auxin from plants (Hu *et al*., 2011).

The use of MAE for plant active ingredients, such as secondary metabolites (e.g., flavonoids, quinones, and terpenoids), has been recently reviewed (Chan *et al*., 2011; Zhang *et al*., 2011). MAE faces the advantage of being quite adaptable to pilot scale, even though industrial-scale applications remain limited. For more details about MAE, see **New Trends in Extraction of Natural Products: Microwave-Assisted Extraction and Pressurized Liquid Extraction**.

5.2.4 Solid-Phase Microextraction

Solid-phase microextraction (SPME) is a real nonsolvent technique, very simple and convenient to use, introduced by Pawliszyn in the early 1990s. A

capillary fiber, covered with a polymeric phase, is put into the sample or its HS, and the solutes partition between the sample and the phase. Desorption into a gas chromatograph injector enables analysis of the solutes. Hence, SPME is a real nonsolvent extraction technique, offering an elevated sensitivity as it avoids dilution of the extracted solutes by a solvent. However, as it lies on partitioning, this technique requires careful calibration if quantitative informations are expected.

As plant materials are mainly solids, SPME is mainly conducted under headspace mode (HS-SPME) (Stashenko and Martínez, 2007); several applications for medicinal plants have been reported (Huie, 2002). Efficient screening of essential oil from aromatic plants, as well as of biogenic volatile organic compounds from plants, were achieved using this technique (Ormeño et al., 2011; Richter and Schellenberg, 2007); odor-active compounds have been also efficiently extracted from fenugreek, this technique offering an extract with an odor very similar to that of the genuine fenugreek seeds (Mebazaa et al., 2009). Recent developments have also reported possible in vivo sampling using SPME, avoiding tedious sampling and pre-treatment steps (Bojko et al., 2012; Matamoros et al., 2012; Musteata and Pawliszyn, 2007). In that case, the fiber must be biocompatible to ensure that (i) fiber materials do not generate undesirable local or systemic reactions within the biological matrix under study and (ii) macromolecules do not adhere. As the SPME fiber extracts only a negligible fraction of the analyte, disturbance of the native chemical equilibrium is minimized; that way, it allows the measurement of analyte concentrations under real physiological conditions. The measurement of emerging organic contaminants within an onion bulb has been recently reported using C_{18} silica fibers (Matamoros et al., 2012). Such developments are promising for metabolomic studies as recently reviewed (Bojko et al., 2012).

Further details about SPME are presented in **Solid-Phase Microextraction (SPME) and Its Application to Natural Products**.

5.2.5 Stir-Bar Sorptive Extraction

Stir-bar sorptive extraction (SBSE) can be considered to be a variant of SPME as a stir-bar covered with a polymeric phase is put into the sample, and partitioning of the solutes enables their extraction; analysis is further performed after thermal desorption (or sometimes upon dissolution in a low solvent volume) (David and Sandra, 2007). As compared to SPME, this technique affords a higher phase volume, enabling higher extraction yields (Duan et al., 2011). This low cost and simple technique is mostly dedicated to liquid samples; hence, it has been used as a cleanup technique of solid extracts in recent years. However, headspace mode (HS-SBSE) has also been reported for solid samples, especially plant materials such as coffee, grapes, or tobacco as recently reviewed (Prieto et al., 2010).

5.2.6 Matrix Solid-Phase Dispersion

Matrix solid-phase dispersion (MSPD) is a patented process that was proposed in 1989 (Barker, 2007; Capriotti et al., 2010). This technique is based on the manual blending of the matrix (solid, semi-solid, or viscous samples) with an abrasive solid support material that allows the mechanical disruption of the matrix structure. The resulting homogenized mixture is then packed in an empty syringe barrel SPE column, and flushed with a solvent that will ensure elution of compounds of interest; as in SPE, a previous solvent washing can be carried out to remove interfering compounds less retained on the sorbent (Ramos, 2012).

The efficiency of MSPD depends on the type and quantity of the solid support, the sample quantity (a sample/solid support ratio of 1:4 is common; however, it should be optimized), and the eluotropic strength and volume of the elution solvent. Common solid supports used in MSPD are modified silica, mostly octadecyl-bonded (C_{18}) silica. Extraction of pesticides or phenolic compounds (i.e., phenolic acids and isoflavonoids) from plant materials have been reported in recent years with this support. Sea sand was reported to be more efficient for extracting phenolic compounds (flavanones and xanthones) from the root bark of M. pomifera C.K. Schneid., probably owing to a more effective disruption of the plant cells with this material; when compared to maceration, smaller solvent amounts and less sample preparation time were required, still enabling higher yields to be obtained (Teixeira and da Costa, 2005). Polar supports, such as Florisil or alumina, were also reported to be efficient for plant materials,

especially as they afford the preferential adsorption of polar plant components (such as pigments and chlorophylls), allowing the removal of these interfering compounds from the final extract (Abhilash *et al*., 2007).

This extraction technique affords simplicity, rapidity, flexibility, and low solvent consumption without any costly investment required; in addition, extraction and cleanup can be performed in one single step by placing an extra sorbent layer at the bottom of the MPSD cartridge. For all these reasons, this technique has faced growing applications in recent few years as recently reviewed (Barker, 2007; Bogialli and Di Corcia, 2007; Capriotti *et al*., 2010; Garcia-Lopez *et al*., 2008). Yet, it is performed under room temperature and atmospheric pressure conditions, which may be inappropriate in some applications for quantitative extraction of solutes. In addition, this technique cannot be completely automated as it requires the blending in a mortar with a pestle by an operator.

5.2.7   QuEChERS

The QuEChERS (Quick, Easy, Cheap, Rugged, Effective, and Safe) method was introduced in 2003 for the extraction of pesticides from fruits and vegetables (Matamoros *et al*., 2012). Practically, it is not a new technique, but rather a new detailed procedure involving several sample treatment stages based on micro-scale solvent extraction followed by cleanup using dispersive SPE (d-SPE) (Fenik *et al*., 2011; Wilkowska and Biziuk, 2011). After shaking the sample with the solvent (acetonitrile in the original method developed), magnesium sulfate is added to promote water separation from the organic solvent. A subsequent treatment with a primary and secondary amine sorbent ensures the removal of polar plant compounds (such as organic acids, sugars, and pigments); others protocols include sample shaking with graphitized carbon black to remove plant compounds such as sterols and pigments.

Owing to its simplicity and rapidity, the QuEChERS method has been adapted for numerous pesticides and other contaminants from fruits and vegetables as well as fatty plant matrices (e.g., olives and olive oil) (Camino-Sánchez *et al*., 2011; Gilbert-López *et al*., 2009; Wilkowska and Biziuk, 2011). It seems appropriate to be used for screening purposes in combination with chromatographic techniques coupled to high resolution mass spectrometry (Cervera *et al*., 2012). However, it suffers from a lack of automation.

5.3   Keys for Choosing the More Appropriate Technique

Ideally, an extraction method should be rapid, simple, and inexpensive (both low investment and low functioning costs), achieving quantitative recovery for solutes of interest (i.e., no loss nor degradation), being in addition highly selective when target approach is considered (i.e., no interfering compounds co-extracted) and providing a final extract sufficiently concentrated to enable solutes to be further measured without any concentration step; finally, the method should generate little or no laboratory wastes, and be integrated to automation. In reality, all techniques available present advantages and drawbacks as summarized in Table 2.

Table 2   Comparison of main extraction techniques for compounds from solid plant materials.

|	Soxhlet	USE	SFE	PLE	MAE	HS-SPME	MSPD
Sample size (g)	5–100	1–30	1–10	5–50	1–20	0.5–10	0.5–10
Extraction time	6–48 h	3–45 min	10–60 min	5–40 min	5–40 min	15–90 min	2–20 min
Solvent consumption (mL)	150–500	10–100	2–30	15–60	10–60	0–1	1–5
Method development	Simple	Simple	Complex	Simple	Simple	Simple	Simple
Selectivity	Very low	Low	High	Low	Low	Moderate	Moderate
In situ cleanup available	No	No	Yes	Yes	No	No	Yes
Filtration or centrifugation required	No	Yes	No	No	Yes	No	No
Operator skill	Low	Moderate	High	Moderate	Moderate	Moderate	Moderate
Level of automation	Low	Moderate	High	High	Moderate	Moderate	Very low
Cost of instrumentation	Very low	Low	High	High	Moderate	Low	Very low
Cost per analysis	Moderate	Low	Moderate	Moderate	Low	Low	Moderate

Source: Data adapted from Camel (2001), Dean and Xiong (2000), Hyötyläinen (2009), and Nyiredy (2004).

Some guiding principles can be given to help choosing a technique. As far as volatile compounds are concerned, HS-based extractions with HS-SPME or HS-SBSE should be recommended as they avoid or minimize the use of solvent, and afford high sensitivity and selectivity toward such compounds. In case of non- or semi-volatile compounds tightly bound to the matrix, PLE or Soxhlet techniques should be recommended, especially the former owing to its rapidity and lower solvent consumption; in addition, elevated pressure in PLE may force the solvent inside the sample, being beneficial to the extraction. SFE can also be used, but organic solvent along with elevated temperature will be required to ensure efficient extraction of such retained compounds. SFE is best suited to selective extraction of thermolabile compounds (Camel, 2001). For solutes trapped in plant cells, USE and MAE are very convenient, as the strong energy delivered disrupts cells and releases the solutes. Partial disruption can also be achieved using MSPD; this technique offers simple and rapid methods, with low solvent volumes. On the other hand, it requires much manual handling of the sample, preventing any automation.

The choice of the extraction technique is of prime importance as it will affect not only the time and cost required to obtain the extract but also the quality of the extract. In recent 20 years, traditional techniques have been replaced by recent techniques in a great part for extracting plant materials as illustrated by Figure 10. Among the latter, SFE, MAE, PLE, and SPME are widely used.

It clearly appears that no universal technique for all types of samples and analytes exists. In practice, the best extraction technique will depend on the application considered, especially the nature of
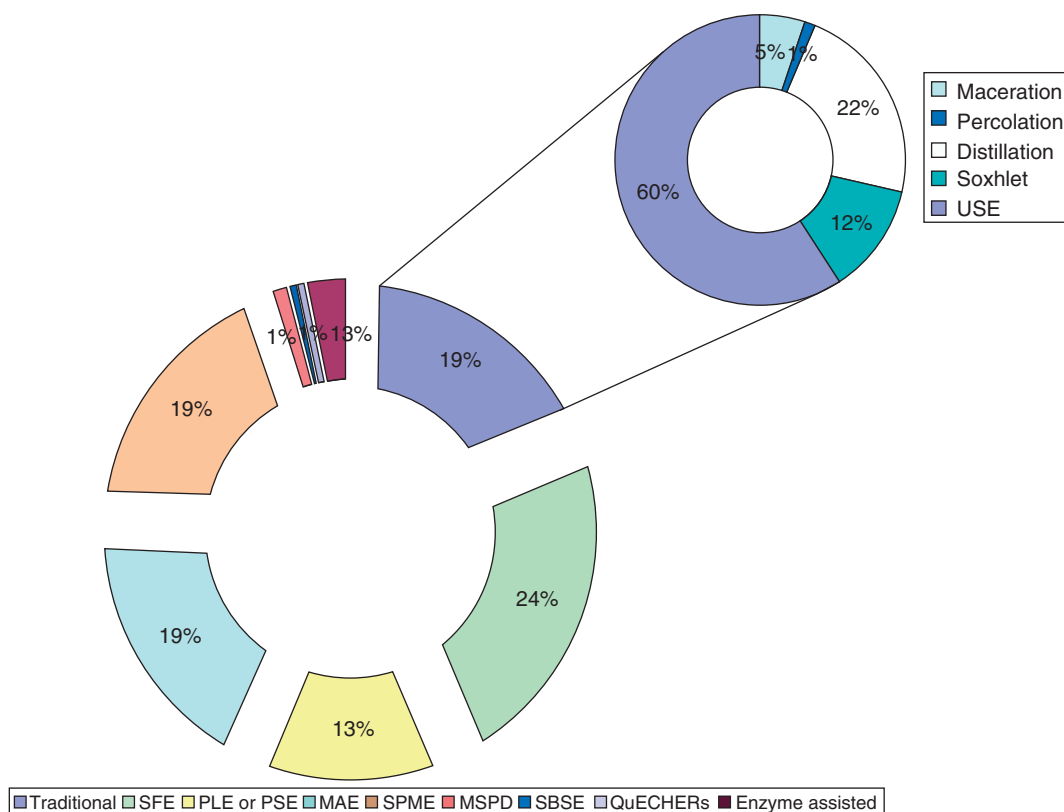
Figure 10 Use of extraction techniques to plant materials for recent 20 years according to the literature indexed in the Web of Science (using the Science Citation Index Expanded Database over the period 1 January 1992–27 November 2012). Name of the technique looked for as keyword in the title of the publication, and "Plant" as keyword in the topic; manual selection according to the application to plant materials (to avoid applications related to water treatment plants, for example).
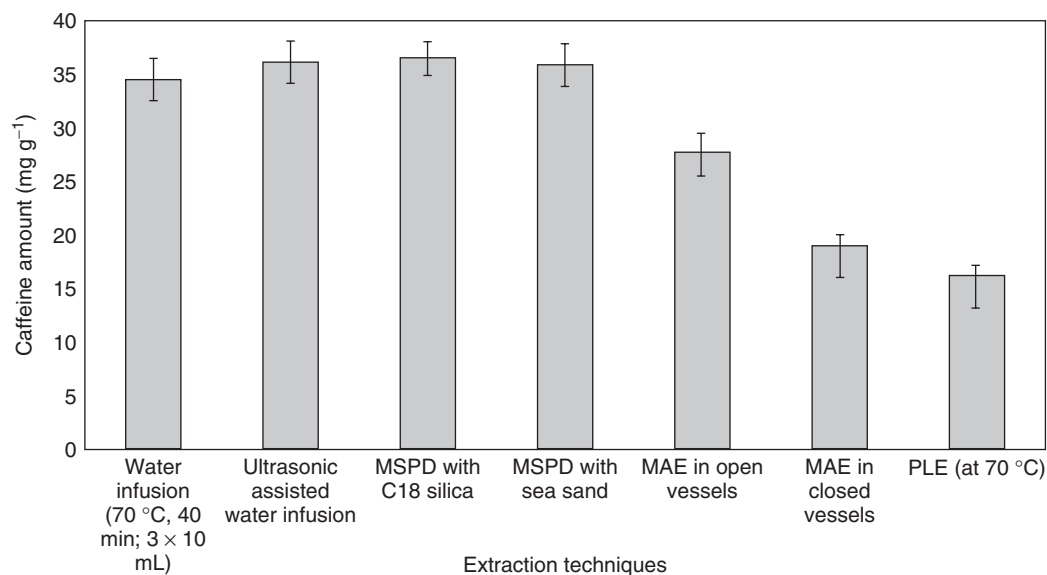
Figure 11　Efficiency (expressed in amount) of caffeine extraction from powdered green tea leaves depending on the techniques used. (*Source*: Data from Dawidowicz and Wianowska (2005).)

compounds of interest as well the nature of the sample matrix. As an example, several techniques were compared in the case of caffeine from powdered green tea leaves (Dawidowicz and Wianowska, 2005). As illustrated in Figure 11, caffeine amounts extracted varied greatly depending on the technique used. Even though the extraction of caffeine from powdered tea leaves is expected to be easy (owing to low solute–matrix interactions), elevated pressure in PLE as well in MAE hindered the extraction, possibly by squeezing the soft tea matrix, thereby making the diffusion of caffeine from the inside to the outside of the matrix difficult or preventing penetration of the inner matrix by the extractant.

6　FUTURE TRENDS: TOWARD GREEN CHEMISTRY

Several developments have been made in recent years to minimize or completely avoid the use of solvents in the extraction step, to fulfill the requirements of green chemistry (Tobiszewski *et al*., 2009). In addition, consumption of energy has been considerably reduced. Besides, there is a growing interest in the use of "generally recognize as safe" (GRAS) solvents, such as carbon dioxide, water, or ethanol, especially

owing to severe regulation of solvent residues in compounds or products extracted from plant materials, especially for medical and food applications. The use of enzymes to enhance extraction is also more and more considered. Finally, the combination of techniques is increasing to benefit from all advantages of those techniques.

6.1　Solvent-Free Microwave Extraction

Solvent-free microwave extraction (SFME) has been proposed to recover essential oil from aromatic plants (Chan *et al*., 2011; Navarrete *et al*., 2012). Its principle is very simple: the plant material (sometimes water pre-treated) is placed in an open microwave reactor, without adding any water or solvent. The internal heating of water present in the plant materials allows heating of the sample; this distends the plant cells and leads to rupture of the glands and oleiferous receptacles, thereby releasing essential oil that is evaporated by the *in situ* water of the plant material. In cases of dried plant materials, a microwave-absorbing medium (e.g., iron carbonyl powder, graphite powder, or active carbon powder) can be added to the sample (Wang *et al*., 2006). A cooling system placed outside the microwave oven

ensures condensation of the distillate continuously. As with hydrodistillation, the extraction temperature is equal to the boiling point of the water (100 °C). However, reaching this temperature is much more rapid with SFME (only 5 min) than with hydrodistillation (near 90 min); consequently, the energy consumption is lowered (0.25 kWh instead of 4.5 kWh) (Lucchesi *et al.*, 2004). In addition, composition of extract obtained differs between these two techniques; higher amounts of oxygenated compounds (highly odoriferous, thus most valuable) and lower amounts of monoterpenes hydrocarbons (less valuable) were reported in essential oils obtained with SFME as compared to hydrodistillation. It was suspected to be related to the much lower water content and time required for SFME, thereby reducing thermal and hydrolytic effects that are commonly observed with polar water as the solvent (Lucchesi *et al.*, 2004).

6.2   Use of GRAS Solvents

Owing to the increasing concern about the quality and safety of plant-derived products and the stricter regulations about the residual level of solvents, the use of supercritical fluids (especially CO_2) has merged in recent years. As an illustration, SFE has been recently reported to allow the extraction of free amino acids from broccoli leaves using CO_2 modified with 35% methanol (i.e., operating in subcritical state), with reduced solvent consumption and time as compared to classical solvent extraction (Arnáiz *et al.*, 2012). Supercritical CO_2 offers a green alternative to classical organic solvents for extracting lycopene as recently reviewed (Zuknik *et al.*, 2012). There is also great interest in using water as an extractant in what is called *pressurized hot water extraction* (*PHWE*) or *subcritical water extraction* (*SWE*) (Mustafa and Turner, 2011); indeed, as water is highly corrosive under supercritical state, subcritical state is preferred for extraction purposes (Kritzer, 2004). SWE is carried out using hot water (100–374 °C, i.e., below its critical temperature) under high pressure (from 10 to 60 bar classically) to maintain water in the liquid state. Interestingly, while water at room temperature is a highly polar solvent, its polarity decreases as the temperature is elevated, so that medium and low polarity compounds may also be extracted under such conditions (Carr *et al.*, 2011). This technique

has been recently reviewed (Kronholm *et al.*, 2007; Teo *et al.*, 2010). Application of SFE and SWE to functional ingredients (especially antioxidants) from different natural sources (such as plants) has also been recently reviewed (Carr *et al.*, 2011; Herrero *et al.*, 2006; Ong *et al.*, 2006). For example, SWE was found efficient for extracting stevioside sweetener from *Stevia rebaudiana* Bertoni (Pól *et al.*, 2007). SWE has been also reported as a useful alternative technique for the extraction of essential oils as compared to traditional techniques (steam distillation and Soxhlet), as it avoids loss and degradation of volatile and thermolabile compounds because of rapid extraction times and low working temperatures (Ozel and Kaymaz, 2004). However, owing to the elevated temperatures used, undesirable compounds may be produced de novo when using SWE (such as Maillard reaction products).

Other environmentally compatible solvents have been recently used for plant material extraction, such as surfactants (e.g., nonionic Genapol X-080 or anionic Triton X-100) in USE, MAE, or PLE of medicinal herbs (Mustafa and Turner, 2011; Ong, 2004; Tang *et al.*, 2009); the use of surfactants in extraction has been recently reviewed (Ballesteros-Gómez *et al.*, 2010; Rodríguez *et al.*, 2008). Ionic liquids also face a growing interest because of their interesting and unique properties (Tobiszewski *et al.*, 2009); as an example, their use in MAE efficiently extracted alkaloids from medicinal herbs, owing to their excellent microwave-absorbing ability (Tang *et al.*, 2009). Polyethylene glycol has also been reported as a valuable green solvent in MAE from medicinal plants (Zhou *et al.*, 2011).

6.3   Enzyme-Assisted Extraction

Use of enzymes in extracting plant materials is of great interest as enzymes are capable of degrading or disrupting cell walls and membranes, thus enabling better release and efficient extraction of solutes. Hence, enzymes are most of the time used as a matrix treatment before the extraction stage as reported for phenols from grape skins (Pinelo *et al.*, 2006); ultrasounds may improve the enzyme action (Bermejo *et al.*, 2004; Chen *et al.*, 2012). Faster extraction, higher efficiency, lower solvent volumes, and lower energy consumption have been reported using enzyme treatment.

For a proper use of enzymes, it is crucial to know their catalytic property and mode of action, their optimal conditions, and also the nature and structure of the plant material. Indeed, various enzymes (e.g., cellulases and pectinases) have been used to enhance extraction from plants as recently reviewed (Puri *et al*., 2012). Yet, at present time, their use for industrial scale is precluded because of their expensive cost and their susceptibility to environmental conditions that may be different on pilot scale as compared to analytical scale.

6.4 Hyphenation or Combination of Extraction Techniques

In recent years, a strategy of combining different extraction techniques has merged, to take benefits from each technique, with a view of reducing solvent consumption, time, and energy dedicated to the extraction stage in the context of green chemistry (Rostagno *et al*., 2010). As an example, pre-treating *Aloe barbadensis* Mill. fresh leaves with SFE (using CO_2) before USE lowered the energy loss as compared to USE (Ivanović *et al*., 2009). Such strategy should also be considered on an industrial scale as recently discussed (Wijngaard *et al*., 2012). Two approaches can be considered: hyphenation of techniques (i.e., they are sequentially performed) or combination of techniques (i.e., they are integrated into a single process). It must be pointed out that, when techniques are combined, the extraction process is more complex to optimize because of several interacting variables.

In particular, ultrasounds have faced a gain in interest (Bendicho *et al*., 2012; Shirsath *et al*., 2012). Hence, ultrasound-assisted SFE (USFE) has been reported to enhance extraction as mechanical stirring is impossible in SFE; a beneficial effect of ultrasounds was reported and was attributed to disruption of the cell structures (Fornari *et al*., 2012). Similarly, ultrasound-assisted MAE (UMAE) has been reported (Chan *et al*., 2011).

Another combination lies in SFME of plant material, for oil extraction, followed by solvent (water or ethanol) extraction enhanced by either microwaves or ultrasounds, as recently reported to recover antioxidants from rosemary (Rodríguez-Rojo *et al*., 2012). The combination of USE with enzyme treatment has also been reported, with a beneficial effect of ultrasounds toward the enzyme efficiency (Shirsath *et al*., 2012).

Finally, combination of MSPD with either PLE or USE has been reported in recent years, and was found to enhance the extraction efficiency and rapidity (Ramos, 2012).

7   CONCLUSION

The extraction stage remains frequently the critical part of the analytical scheme, whatever the considered approach (targeting or profiling). Hence, this step deserves particular attention. As numerous extraction techniques can be used, several factors need to be considered to choose the best technique devoted to a specific application as detailed in this chapter. The degree of selectivity expected and the stability of solutes are particularly important to assess; the nature of the plant material and the localization of the analytes are also key factors. Extraction methods should be developed in agreement with green chemistry principles as far as possible; in addition, offering a reproducible quality of the extract is a prerequisite. There is no doubt that hyphenation or combination of techniques will increase in the near future as such an approach can afford beneficial aspects of the different techniques.

REFERENCES

Abhilash, P. C., Jamil, S. and Singh, N. (2007) *J. Chromatogr. A*, **1176**, 43–47.

Arnáiz, E., Bernal, J., Martín, M. T., *et al*. (2012) *J. Chromatogr. A*, **1250**, 49–53.

Ballesteros-Gómez, A., Sicilia, M. D. and Rubio, S. (2010) *Anal. Chim. Acta*, **677**, 108–130.

Barker, S. A. (2007) *J. Biochem. Biophys. Methods*, **70**, 151–162.

Barwick, V. (1997) *Trends Anal. Chem.*, **16**, 293–309.

Bendicho, C., De La Calle, I., Pena, F., *et al*. (2012) *Trends Anal. Chem.*, **31**, 50–60.

Benthin, B., Danz, H. and Hamburger, M. (1999) *J. Chromatogr. A*, **837**, 211–219.

Bermejo, P., Capelo, J. L., Mota, A., *et al*. (2004) *Trends Anal. Chem.*, **23**, 654–663.

Bogialli, S. and Di Corcia, A. (2007) *J. Biochem. Biophys. Methods*, **70**, 163–179.

Bojko, B., Cudjoe, E., Gómez-Ríos, G. A., *et al*. (2012) *Anal. Chim. Acta*, **750**, 132–151.

Camel, V. (2000) *Trends Anal. Chem.*, **19**, 229–248.

Camel, V. (2001) *Analyst*, **126**, 1182–1193.

Camel, V. (2002) *Anal. Bioanal. Chem.*, **372**, 39–40.

Camino-Sánchez, F. J., Zafra-Gómez, A., Ruiz-García, J., *et al.* (2011) *J. Food Compos. Anal.*, **24**, 427–440.

Capriotti, A. L., Cavaliere, C., Giansanti, P., *et al.* (2010) *J. Chromatogr. A*, **1217**, 2521–2532.

Carr, A. G., Mammucari, R. and Foster, N. R. (2011) *Chem. Eng. J.*, **172**, 1–17.

Cervera, M. I., Portolés, T., Pitarch, E., *et al.* (2012) *J. Chromatogr. A*, **1244**, 168–177.

Chan, C. H., Yusoff, R., Ngoh, G. C., *et al.* (2011) *J. Chromatogr. A*, **1218**, 6213–6225.

Chen, R., Li, S., Liu, C., *et al.* (2012) *Process Biochem.*, **47**, 2040–2050.

David, F. and Sandra, P. (2007) *J. Chromatogr. A*, **1152**, 54–69.

Dawidowicz, A. L. and Wianowska, D. (2005) *J. Pharm. Biomed. Anal.*, **37**, 1155–1159.

Dean, J. R. and Xiong, G. (2000) *Trends Anal. Chem.*, **19**, 553–564.

Delgado-Zamarreño, M. M., Pérez-Martin, L., Bustamante-Rangel, M., *et al.* (2012) *Anal. Bioanal. Chem.*, **404**, 361–366.

Duan, C., Shen, Z., Wu, D., *et al.* (2011) *Trends Anal. Chem.*, **30**, 1568–1574.

Fenik, J., Tankiewicz, M. and Biziuk, M. (2011) *Trends Anal. Chem.*, **30**, 814–826.

Fornari, T., Vicente, G., Vazquez, E., *et al.* (2012) *J. Chromatogr. A*, **1250**, 34–48.

Garcia-Lopez, M., Canosa, P. and Rodriguez, I. (2008) *Anal. Bioanal. Chem.*, **391**, 963–974.

Gilbert-López, B., García-Reyes, J. F. and Molina-Díaz, A. (2009) *Talanta*, **79**, 109–128.

González-Curbelo, M. A., Herrera-Herrera, A. V., Ravelo-Pérez, L. M., *et al.* (2012) *Trends Anal. Chem.*, **38**, 32–51.

Hansen, C. M. (2004) *Prog. Org. Coat.*, **51**, 77–84.

Herrero, M., Cifuentes, A. and Ibañez, E. (2006) *Food Chem.*, **98**, 136–148.

Hinneburg, I. and Neubert, R. H. H. (2005) *J. Agric. Food Chem.*, **53**, 3–7.

Hu, Y., Li, Y., Zhang, Y., *et al.* (2011) *Anal. Bioanal. Chem.*, **399**, 3367–3374.

Huang, Z., Shi, X. H. and Jiang, W. J. (2012) *J. Chromatogr. A*, **1250**, 2–26.

Huie, C. W. (2002) *Anal. Bioanal. Chem.*, **373**, 23–30.

Hyötyläinen, T. (2009) *Anal. Bioanal. Chem.*, **394**, 743–758.

Ivanović, J., Žižović, I., Petrović, S., *et al.* (2009) *Chem. Ind. Chem. Eng. Quart.*, **15**, 271–278.

Kaufmann, B. and Christen, P. (2002) *Phytochem. Anal.*, **13**, 105–113.

Kolberg, D. I. S., Mack, D., Anastassiades, M., *et al.* (2012) *Anal. Bioanal. Chem.*, **404**, 2465–2474.

Kritzer, P. (2004) *J. Supercrit. Fluids*, **29**, 1–29.

Kronholm, J., Hartonen, K. and Riekkola, M. L. (2007) *Trends Anal. Chem.*, **26**, 396–412.

Lang, Y. H., Cao, Z. M. and Jiang, X. (2005) *Talanta*, **66**, 249–252.

Lang, Q. and Wai, C. M. (2001) *Talanta*, **53**, 771–782.

Li, H.-B., Jiang, Y., Wong, C.-C., *et al.* (2007) *Anal. Bioanal. Chem.*, **388**, 483–488.

Lü, Y., Shi, J. J. and Lei, S. (2008) *J. Fuel Chem. Technol.*, **36**, 297–301.

Lucchesi, M. E., Chemat, F. and Smadja, J. (2004) *J. Chromatogr. A*, **1043**, 323–327.

Luque de Castro, M. D. and Priego-Capote, F. (2010) *J. Chromatogr. A*, **1217**, 2383–2389.

Matamoros, V., Calderon-Preciado, D., Dominguez, C., *et al.* (2012) *Anal. Chim. Acta*, **722**, 8–20.

Mebazaa, R., Mahmoudi, A., Fouchet, M., *et al.* (2009) *Food Chem.*, **115**, 1326–1336.

Meireles, M. A. A. (2003) *Curr. Opin. Solid State Mat. Sci.*, **7**, 321–330.

Michiels, J. A., Kevers, C., Pincemail, J., *et al.* (2012) *Food Chem.*, **130**, 986–993.

Mustafa, A. and Turner, C. (2011) *Anal. Chim. Acta*, **703**, 8–18.

Musteata, F. M. and Pawliszyn, J. (2007) *J. Biochem. Biophys. Methods*, **70**, 181–193.

Navarrete, A., Mato, R. B. and Cocero, M. J. (2012) *Chem. Eng. Sci.*, **68**, 192–201.

Newton, R. P., Brenton, A. G., Smith, C. J., *et al.* (2004) *Phytochemistry*, **65**, 1449–1485.

Nyiredy, S. (2004) *J. Chromatogr. B*, **812**, 35–51.

Ong, E. S. (2004) *J. Chromatogr. B*, **812**, 23–33.

Ong, E. S., Cheong, J. S. H. and Goh, D. (2006) *J. Chromatogr. A*, **1112**, 92–102.

Ormeño, E., Goldstein, A. and Niinemets, Ü. (2011) *Trends Anal. Chem.*, **30**, 978–989.

Ozel, M. Z. and Kaymaz, H. (2004) *Anal. Bioanal. Chem.*, **379**, 1127–1133.

Pareja, L., Fernández-Alba, A. R., Cesio, V., *et al.* (2011) *Trends Anal. Chem.*, **30**, 270–291.

Pavela, R., Sajfrtová, M., Sovová, H., *et al.* (2010) *Ind. Crop Prod.*, **31**, 449–454.

Pérez, R. A., Tadeo, J. L., Albero, B., *et al.* (2013) *Anal. Bioanal. Chem.*, **405**, 389–400.

Pinelo, M., Arnous, A. and Meyer, A. S. (2006) *Trends Food Sci. Technol.*, **17**, 579–590.

Pól, J., Ostrá, E. V., Karásek, P., *et al.* (2007) *Anal. Bioanal. Chem.*, **388**, 1847–1857.

Pourmortazavi, S. M. and Hajimirsadeghi, S. S. (2007) *J. Chromatogr. A*, **1163**, 2–24.

Prieto, A., Basauri, O., Rodil, R., *et al.* (2010) *J. Chromatogr. A*, **1217**, 2642–2666.

Puri, M., Sharam, D. and Barrow, C. J. (2012) *Trends Biotechnol.*, **30**, 37–44.

Raessler, M. (2011) *Trends Anal. Chem.*, **30**, 1833–1843.

Ramos, L. (2012) *J. Chromatogr. A*, **1221**, 84–98.

Richter, J. and Schellenberg, I. (2007) *Anal. Bioanal. Chem.*, **387**, 2207–2217.

Rodríguez, J. J. S., Ferrera, Z. S., Moreno, D. V., *et al.* (2008) *Anal. Bioanal. Chem.*, **391**, 725–733.

Rodríguez-Rojo, S., Visentin, A., Maestri, D., *et al.* (2012) *J. Food Eng.*, **109**, 98–103.

Romanik, G., Gilgenast, E., Przyjazny, A., *et al.* (2007) *J. Biochem. Biophys. Methods*, **70**, 253–261.

Rostagno, M. A., Villares, A., Guillamón, E., *et al.* (2009) *J. Chromatogr. A*, **1216**, 2–29.

Rostagno, M. A., D'Arrigo, M. and Martínez, J. A. (2010) *Trends Anal. Chem.*, **29**, 553–561.

Sairam, P., Ghosh, S., Jena, S., *et al.* (2012) *Asian J. Res. Pharm. Sci.*, **2**, 112–120.

Santos, D. T., Albuquerque, C. L. C. and Meireles, M. A. A. (2011) *Proc. Food Sci.*, **1**, 1581–1588.

Shirsath, S. R., Sonawane, S. H. and Gogate, P. R. (2012) *Chem. Eng. Proc.*, **53**, 10–23.

Sivakumar, V., Vijaeeswarri, J. and Anna, J. L. (2011) *Ind. Crops Prod.*, **33**, 116–122.

Sovová, H. (2012a) *J. Chromatogr. A*, **1250**, 27–33.

Sovová, H. (2012b) *J. Supercrit. Fluids*, **66**, 73–79.

Srinivas, K., King, J. W., Monrad, J. K., *et al.* (2009) *J. Food Sci.*, **74**, E342–E354.

Stamenić, M., Zizovic, I., Orlović, A., *et al.* (2008) *J. Supercrit. Fluids*, **46**, 285–292.

Stashenko, E. E. and Martínez, J. R. (2007) *J. Biochem. Biophys. Methods*, **70**, 235–242.

Stefanis, E. and Panayiotou, C. (2008) *Int. J. Thermophys.*, **29**, 568–585.

Sun, H., Ge, X., Lv, Y., *et al.* (2012) *J. Chromatogr. A*, **1237**, 1–23.

Tang, F., Zhang, Q., Nie, Z., *et al.* (2009) *Trends Anal. Chem.*, **28**, 1253–1262.

Teixeira, D. M. and Costa, C. T. da (2005) *J. Chromatogr. A*, **1062**, 175–181.

Teo, C. C., Tan, S. N., Yong, J. W. H., *et al.* (2010) *J. Chromatogr. A*, **1217**, 2484–2494.

t'Kindt, R., Morreel, K., Deforce, D., *et al.* (2009) *J. Chromatogr. B*, **877**, 3572–3580.

Tobiszewski, M., Mechlińska, A., Zygmunt, B., *et al.* (2009) *Trends Anal. Chem.*, **28**, 943–951.

Toma, M., Vinatoru, M., Paniwnyk, L., *et al.* (2001) *Ultrason. Sonochem.*, **8**, 137–142.

Urbanová, J., Pěnčíková, K., Gregorová, J., *et al.* (2012) *Phytochem. Anal.*, **23**, 477–482.

Vilkhu, K., Mawson, R., Simons, L., *et al.* (2008) *Innov. Food Sci. Technol.*, **9**, 161–169.

Villmow, T., John, A., Pötschke, P., *et al.* (2012) *Polymer*, **53**, 2908–2918.

Vinatoru, M. (2001) *Ultrason. Sonochem.*, **8**, 303–313.

Wang, L. and Weller, C. L. (2006) *Trends Food Sci. Technol.*, **17**, 300–312.

Wang, Z., Wang, L., Li, T., *et al.* (2006) *Anal. Bioanal. Chem.*, **386**, 1863–1868.

Wijngaard, H., Hossain, M. B., Rai, D. K., *et al.* (2012) *Food Res. Int.*, **46**, 505–513.

Wilkowska, A. and Biziuk, M. (2011) *Food Chem.*, **125**, 803–812.

Yeshak, M. Y., Burman, R., Eriksson, C., *et al.* (2012) *Phytochem. Lett.*, **5**, 776–781.

Zhang, H. F., Yang, X. H. and Wang, Y. (2011) *Trends Food Sci. Technol.*, **22**, 672–688.

Zhao, J., Lv, G.-P., Chen, Y.-W., *et al.* (2011) *J. Chromatogr. A*, **1218**, 7453–7475.

Zhou, T., Xiao, X., Li, G., *et al.* (2011) *J. Chromatogr. A*, **1218**, 3608–3615.

Zuknik, M. H., Norulaini, N. A. N. and Omar, A. K. M. (2012) *J. Food Eng.*, **112**, 253–262.

3

Supercritical Fluid Extraction

Seied Mahdi Pourmortazavi[1], Mehdi Rahimi-Nasrabadi[2] and Somayeh Mirsadeghi[3]

[1]*Faculty of Material and Manufacturing Technologies, Malek Ashtar University of Technology, Tehran, Iran,* [2]*Department of Chemistry, Imam Hossein University, Tehran, Iran* and [3]*Department of Chemistry, Islamic Azad University, Varamin Pishva Branch, Varamin, Iran*

1 INTRODUCTION

Traditionally, the essential oils and volatile components extracted from various plants and herbal products have been used in the preparation of cosmetics, foodstuffs, cleaning solutions, fragrance compounds, herbicides, and insecticide products. Moreover, essential oils have been used in traditional medicine since ancient to present time as diuretic, digestive, expectorant, and sedatives (Rahimi-Nasrabadi *et al.*, 2009). Furthermore, during recent decades, several scientific studies have been carried out on biological properties (such as antioxidant, antiviral, anti-inflammatory, and antibacterial) of essential oils extracted from different plants and herbs (Gholivand *et al.*, 2011, 2012; Rahimi-Nasrabadi *et al.*, 2009, 2012a, 2012b, 2013a, 2013b). Today, essential oils are widely used in aromatherapic science (as a branch of medicine that employs the essential oils and other aromatic compounds for curative effects). These widespread applications of plant oils caused introducing various techniques and methods for effective extraction and isolation of oils from plant matrices.

Different traditional techniques have been used for the extraction of essential oils from plant matrix, that is, steam-distillation, HD, and liquid-solvent extraction (Kerrola *et al.*, 1994; Luque de Castro and Garcia-Ayuso, 1998). However, some other new and effective technologies obtained much attention in the research and field developments for extraction and isolation of these compounds from raw materials (Grevenstuk *et al.*, 2012; Dandekar and Gaikar*et al.*, 2002). The results of researches for the efficient extraction of the analytes led to the development of modern, fast, and more powerful extraction techniques (Pourmortazavi and Ghadiri, 2005; Pourmortazavi and Hajimirsadeghi, 2007). Moreover, essential oil extraction techniques with the goal of green extraction and separation science are now available (Pourmortazavi *et al.*, 2004a). SFE using carbon dioxide is a promising alternative method that is widely used for the extraction and isolation of essential oil from plant matrices (Ebrahimzadeh *et al.*, 2003).

2 PROPERTIES OF SUPERCRITICAL FLUIDS

Supercritical fluids (SCFs) are described as the fluids that are above their critical values of temperature and pressure. In the supercritical state, only one fluid phase exists. It has both gas and liquid-like properties. Under supercritical state, the gases will not be condensed by increasing pressure. The phase
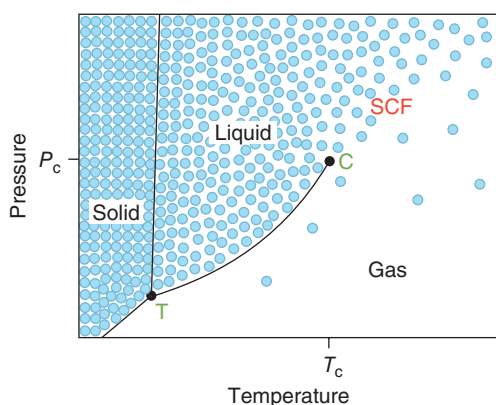
Figure 1 A schematic phase diagram for a pure component including its supercritical fluid (SCF) region; the triple point (T); and critical point (C) in the phase diagram are marked. The circles in the figure represent the variation in density of the substance in the different regions of the phase diagram. (*Source*: Cooper (2000). Reproduced by permission of Royal Society of Chemistry.)

Table 1   Critical parameters for selected substances.

Substance	T_c (°C)[a]	P_c (bar)[b]	ρ_c (g cm^{-3})[c]
Ar	−122.3	48.5	0.53
CH_4	−82.5	46.4	0.16
C_2H_6	32.4	48.8	0.20
C_2F_6	19.9	30.6	0.62
CHF_3	26.2	48.5	0.62
CO_2	31.1	73.8	0.47
C_2H_4	10.0	51.2	0.22
SF_6	45.6	37.2	0.73
NH_3	132.5	112.8	0.24
MeOH	240.6	79.9	0.27
EtOH	243.5	63.8	0.28
C_6H_6	289.0	48.9	0.30
H_2O	374.2	220.5	0.32

[a] T_c is critical temperature.
[b] P_c is the critical pressure.
[c] ρ_c is the corresponded density at the critical pressure and critical temperature.

Source: Reprinted with permission from Darr J.A., Poliakoff M. 1999. New directions in inorganic and metal-organic coordination chemistry in supercritical fluids. Chem Rev 99: 495–541. Copyright 1999 American Chemical Society.

diagram for a pure compound at supercritical conditions is shown in Figure 1.

SCFs possess unique properties. Their solvating effectiveness could be controlled by small changes in pressure and temperature. The densities of SCFs are much greater than those of typical gases and slightly less than those of the organic liquids. Meanwhile, the values of viscosity for the SCFs are almost similar to those for typical gases and considerably less than those of liquids. In other words, the SCFs possess densities and dissolving capacities similar to those of certain liquids, whereas their viscosities are lower and so the SCFs have better diffusion properties. These properties cause high fluid phase capacity along with favorable transport properties, which make SCFs appropriate solvents for the extraction purposes. Table 1 presents a clearer comparison among density, viscosity, and diffusivity of the gases, SCFs, and organic liquids.

Various solvents and gases are available for use as SCFs, that is, carbon dioxide, ethane, nitrous oxide, propane, ammonia, fluoroform, *n*-pentane, sulfur hexafluoride, and water. Literature review shows that among the wide variety of the above-mentioned compounds, carbon dioxide is the mainly used supercritical solvent. CO_2 with T_c of 31.1 °C and P_c of 7.38 MPa is a cheap, low toxic, inflammable, environmentally friendly, and safe fluid (Bayat *et al.*, 2012). Supercritical carbon dioxide (SC-CO_2) is also attractive because of its high diffusivity combined with its easily tunable solvating power. In

addition, CO_2 at room temperature and pressure is in gaseous form, which makes the analyte recovery very simple and after extraction process provides solvent-free analytes. This is an important advantage for an extraction solvent especially in preparation of food and natural product samples. In addition, SFE using CO_2, which is performed at low temperatures using a nonoxidant medium, allows the extraction of thermally labile or easily oxidized analytes (Pourmortazavi and Hajimirsadeghi, 2007). The high solvating power of SCFs is the major argument for laboratories to be interested in developing SFE methods for routine analyses in their innovative research. During recent decade, a number of laboratories replaced their conventional extraction methodologies with SFE-based technologies to minimize the consumption of organic solvents.

Unfortunately, the application of CO_2 as SCF is somewhat restricted because of its inadequate solvating power especially for highly polar analytes; however, this limitation could be paled by the addition of an appropriate modifier (Pourmortazavi and Hajimirsadeghi, 2005). The addition of relatively small amounts (less than 10%) of methanol (MeOH) as modifier to carbon dioxide considerably expands its extraction range to include more polar analytes. Furthermore, the modifiers could reduce the interactions between analyte and matrix and hence improve the

quantitative extraction (Yamini *et al*., 2002). There are many literature reviews available that deal with the basic principles of SFE; some of them are: (Pourmortazavi *et al*., 2014; Brunner, 2005; Pereira and Meireles, 2010; Zuknik *et al*., 2012; Sahena *et al*., 2009; Mukhopadhyay, 2000; Oliveira *et al*., 2011; Erkey, 2000; Lehotay, 1997; Zougagh *et al*., 2004).

## 3 SFE SYSTEM AND EXTRACTION PROCEDURE

A schematic illustrating the main components of a typical SFE is shown in Figure 2. A gas cylinder provides a source of CO_2, which is pumped into the system with either a piston or a syringe-type pump. When using modified fluids, the piston pumps are more convenient. Because using these pumps, changing of modifier is easier, and the system is not contaminated by previous solvent. Meanwhile, the piston pumps allow continuous addition of fluid into the system; however, a fixed amount of fluid is provided via a syringe pump. Consequently, more samples could be extracted consecutively by SFE when a piston pump is used in the system.

As shown in Figure 2, for the extraction of analytes polar than pure carbon dioxide, a supplementary modifier pump is used in the system. The sample is placed into the extraction vessel. Then, the vessel is attached to the fluid pass inside a temperature-controlled oven and heated to desired temperature, which is above the T_c of the extraction fluid. Another important part of the SFE system is the restrictor. The restrictor maintains back pressure

in the system by limiting the amount of passing fluid that caused the supercritical conditions be maintained. In fact, the restrictor acts as the interface between the high pressure cell and the laboratory atmosphere.

The extracts collection system is placed next to the restrictor. Various collection methods could be used in SFE systems. Two of most common of analytes collection systems are solid phase and liquid solvent (Pourmortazavi and Hajimirsadeghi, 2007). Usually, the chromatography column packing materials, that is, octadecyl silica, solid sorbents such as Tenax and graphite, or inert surfaces such as glass wool and glass or stainless steel beads are used in the solid phase collection systems. In order to effective trapping of extracts, the solid phase trapping systems are commonly cooled for the collection of volatile extracted compounds. In the liquid collection method, the restrictor outlet is immersed directly into a solvent such as MeOH, ethanol (EtOH), or dichloromethane. The used solvent for trapping should be solvated the extracted analytes.

SFE procedure could be carried out by two principal modes of static and dynamic. In the static extraction, sample matrix is exposed to a fixed amount of SC-CO_2 for a definite time. In contrast, during the dynamic extraction, fresh CO_2 continuously passes through the sample matrix. In the most of SFE experiments, a combination of both static and dynamic modes is employed. At first, a static mode is used to allow the penetration of SC-CO_2 to the plant matrix and solubilize the analytes, followed by a dynamic mode that sweeps the analytes from the extraction vessel, through the restrictor, and into
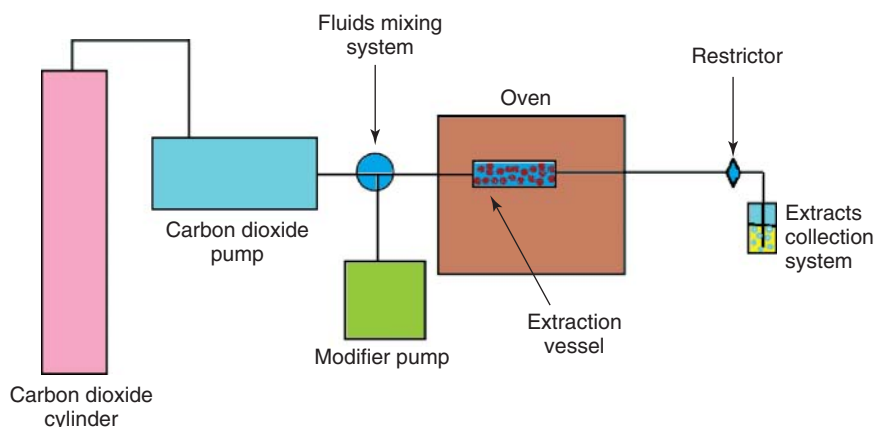
Figure 2 A typical scheme for supercritical fluid extraction (SFE) system.

the collection system. After passing CO_2/analytes through the restrictor, the CO_2 decompresses to the atmospheric pressure and loses its SCF properties such as solvating power. The analytes therefore are released into the trapping system.

4   EFFECT OF SFE CONDITIONS

Various parameters such as pressure, temperature, CO_2 flow rate, and extraction time could affect the solubility of an analyte molecule in the CO_2 phase (Aghel *et al*., 2004). The pressure and temperature of SCF influence the density of the fluid that determines the number of interactions between CO_2 and molecules of the analytes. Increasing these interactions provide enough cohesive forces between analyte molecules and SCF to conquer binding forces between analyte and sample matrix. In other words, the analyte molecules solubilized in the SCF. Therefore, the value of target molecules solubility in SC-CO_2 is depended on the molecular weight of analyte compound and the level of interactions between SC-CO_2 and analyte molecules (del Valle and Urrego, 2012; Lucien and Foster, 2000). Thus, optimization of the experimental parameters is a critical step in developing of SFE method for the efficient extraction and recovery of the plant oils and volatile components. On the other hand, kinetics of SFE from plants matrices is variable due to the effect of various parameters on the extraction process. Previous studies (Sovova, 2012) proposed that the kinetics of extraction process may be limited by the following steps: (1) the value of analytes solubility in the fluid phase, (2) the internal diffusion of fluid phase through the low permeable solid matrix, and (3) the effect of more complex samples by both fluid phase equilibrium and mass transfer process. Commonly, the extraction of the first fraction of analytes is limited by their solubilities in fluid phase and the extraction rate of the rest analytes is limited by diffusion process. Meanwhile, the phase equilibrium is dependent on the interactions between analytes and the solid matrix. The equilibrium solubility of the analytes without considerable interactions with the matrix (i.e., vegetable oil in seeds) could be determined easily (Huang *et al*., 2012). In plant samples, the oil content in solid matrix is relatively high and hence the concentration of oil in fluid phase equilibrium is independent upon the matrix and is proper to its solubility in SC-CO_2. In other words,

at the first period of the SFE process, the rate of extraction process is reasonably controlled by the equilibrium between fluid phase and solid phase at the extraction vessel. On the other hand, substantial analytes–matrix interactions caused the extraction rate be dependent on the bounding of analytes to the solid matrix such as either on the outer surface or inside the solid structure. Thus, mass transfer kinetics of analytes from the solid matrix to the SCF phase has important role in the extraction. Prediction of the mass transfer coefficient required accurate modeling and quality control on line. In fact, mass transfer factors are characterized by the axial dispersion, external mass transfer coefficient, and internal mass transfer coefficient (or solute diffusivity in the solid matrix) in terms of two dimensionless numbers. Usually, correlation equations using characteristic dimensionless numbers are used for this aim. Various correlation equations have been proposed for the prediction of mass transfer coefficient which combined the effects of various factors on mass transfer process (Sovova, 2012; del Valle and Urrego, 2012).

Among various parameters that affect SFE yield, pressure, temperature, modifier, extraction time, and SCF flow rate play important roles. Thus, the effects of these parameters on the SFE of plant sample are discussed in this chapter.

4.1   Pressure of SC-CO_2

Pressure of SCF is the most important physical parameter in the SFE procedure, and it has both theoretical and practical implications in the SC-CO_2 extraction. This parameter could be used to tune the selectivity of the SCF solvent in the extraction of various analytes. As a general belief among the analytical chemists, the maximum solubility of a solute in the SCFs is achieved by maximum fluid density. The increasing of pressure results in the higher recovery of the oil and extracts. This behavior is predictable because of raising the extraction pressure, which leads to a higher fluid density, and more solubility of the oil and extracts in the SC-CO_2 (Tonthubthimthong *et al*., 2001). Many reports could be found on the investigations performed for the effect of SC-CO_2 pressure on the yield of extraction and extract composition of various plants samples. Table 2 presents some of these papers and optimum pressure for the efficient extraction of plant extractable components.

Table 2 Some of investigations performed on effect of pressure on the SFE of plant.

Plant sample	Range of pressure (MPa)	Optimum pressure (MPa)	Extraction conditions (temperature, flow rate, etc.)	Result of pressure optimization	References
Alkanna tinctoria	5–35	17.5	80 °C; 5 g min^{-1}	Increasing total yield of alkannins extraction	Akgun *et al.* (2011)
Hypericum perforatum L.	10, 15, 20	20	40 °C	Maximum yield of extraction	Martín *et al.* (2011)
Mucuna seeds (*Mucuna aterrima*, *Mucuna cinerium*, and *Mucuna deeringiana*)	15–25	25	40 °C; 3 mL min^{-1}	Higher extraction yields and increasing concentration of l-Dopa in defatted meal	Augusto dos Santos Garcia *et al.* (2012)
Aerial parts of sage (*Salvia officinalis* L.)	15, 30	—	50 °C	15 MPa as optimal pressure for selective extraction of diterpenes; 30 MPa as optimal pressure for higher total extracts yield	Glisic *et al.* (2011)
Rosemary (*Rosmarinus officinalis* L.)	15–40	30	50 °C	Increasing carnosic acid recovery	Visentín *et al.* (2011)
Lavender (*Lavandula angustifolia*)	16–28	22	45 °C; 1 h extraction time	Increasing extraction yield	Zhi-ling *et al.* (2011)
Wasabi (*Wasabia japonica* Matsum)	15–25	20	35 °C	Increasing extraction yield of isothiocyanate	Li *et al.* (2010)
Zanthoxylum bungeanum seeds	15–30	29.28	41.19 °C; 10.94% added amount of modifier	Increasing extraction yield	Xia *et al.* (2011)
Aerial parts of *Bidens tripartita* L.	9.1, 15.1	15.1	50 °C; 1 mL min^{-1}; 2 min static and 15 min dynamic extractions time	Increasing extraction yield	Kaškonienė *et al.* (2011)
Sorghum bug (*Agonoscelis pubescens*)	20–40	30	60 °C; 15 g min^{-1} C; EtOH as modifier; extraction time 30 min	Increasing extraction yield	Mariod *et al.* (2010)
Alpinia oxyphylla seeds	20–40	28.5	67.5 °C; 2.7 h extraction time	Increasing extraction yield of oil	Ju *et al.* (2010)
Calendula officinalis L.	35–80	≥50	50–60 °C; ≥2 h extraction time	Increasing extraction yield of Faradiol Monoesters	Förg *et al.* (2002)
Salvia officinalis L.	25, 35	35	100 °C; 0.05 kg min^{-1}; EtOH as modifier	Increasing solubility of sage extractives	Daukšas *et al.* (2001)
Capsicum frutescens fruits	15–23	23	40 °C; 10 min extraction time	Increasing extraction yield and content of capsaicinoids	Gouveia *et al.* (2006)
Soybean oil (*Glycine max*)	20–28	25	55 °C; 60 min extraction time	Increasing phytosterol esters contents in raffinate up to 82.4% with the yield of 72%	Torres *et al.* (2009)
Onion oil (*Allium cepa* L.)	16–24	20.6	40.6 °C; 22 L h^{-1}; 260 min extraction time; EtOH as modifier	Highest yield of extracted oil	Ye and Lai (2012)
Polygonum cuspidatum	10–30	30	50 °C; 20 kg h^{-1}; 10% EtOH as modifier	Increasing yield of resveratrol by pressure increase until 15 MPa; very low yield of piceid before 15 MPa; significant increase yield of piceid after 15 MPa and leveled out after 25 MPa	Wenli *et al.* (2005)

Table 2   (*Continued*)

Plant sample	Range of pressure (MPa)	Optimum pressure (MPa)	Extraction conditions (temperature, flow rate, etc.)	Result of pressure optimization	References
Sea buckthorn (*Hippophae rhamnoides* L.)	30–46	46	60 °C	Increasing extraction yield and carotenoid content	Cossuta *et al.* (2007)
Atractylode lancea rhizome	15–25	20	50 °C; 15 L h^{-1}; 1.5 h dynamic extraction time	Increasing extraction yield	Chen *et al.* (2009)
Angelica dahurica	25–35	30	50 °C; 2 L min^{-1}; 75% EtOH as modifier	Increasing extraction yield	Chen *et al.* (2008)
Tara pods (*Caesalpinia spinosa*)	35–55	55	40 °C; 2.5 L min^{-1}; 30 min static time; 30 min dynamic extraction time	Increasing extraction yield of polyphenols	Romero *et al.* (2012)
Oregano (*Origanum virens* L.) bracts	5–30	10	37 °C; 0.5 kg h^{-1}	Increasing extraction yield of essential oil; avoiding co-extraction of undesirable cuticular waxes	Gaspar (2002)
Soybean oil (*Glycine max*)	10–30	30	50 °C; 1.629 L min^{-1}; 4 h extraction time	Highest yield of soybean oil; higher content of triacylglycerols with unsaturated fatty acids (linoleic and linolenic acid)	Jokić *et al.* (2012)
Satureja hortensis	30.3–40.5	35.0	72.6 °C, 8.6% (v/v) MeOH as modifier	Increasing yield of extraction	Khajeh (2011)
Persea indica	10–20	20	40 °C; 1.44 kg h^{-1} flow rate; 0.56 mm particle size	Increasing yield of extraction	Martín *et al.* (2011)
Mentha pulegium and *Satureja montana* L.	12.5, 20, and 30	30	40 °C; 0.04 kg h^{-1} flow rate	Increasing extraction yield of hydrocarbon fraction with enhancing pressure	Palavra *et al.* (2011)
Seabuckthorn or *Hippophae rhamnoids* L. seeds	10, 20, 30, and 40	40	35 °C; 10 g min^{-1} flow rate; 60 min extraction time	Increasing extraction yield of tocopherols and carotenes	Kagliwal *et al.* (2011)
Microula sikkimensis seed	21, 24, and 27	24	45 °C; 20 L h^{-1} flow rate; 3 h extraction time	Increasing extraction yield of oil	Lina *et al.* (2010)

Khajeh (2011) investigated the effect of pressure in the range 30.3–40.5 MPa and some other parameters on extracts of *Satureja hortensis* cultivated in Iran. The results showed that optimal conditions for SFE of the studied plant are a pressure of 35.0 MPa, temperature of 72.6 °C, and 8.6% (v/v) MeOH as modifier, whereas, under these extraction conditions, the main extracted compounds were γ-terpinene (35.5%), thymol (18.2%), and carvacrol (29.7%). In addition, the obtained response surface diagrams demonstrated that the pressure and modifier volume are highly significant factors in the SFE of this plant, whereas an increase in pressure and modifier volume increases the yield of extracts.

Martín *et al.* (2011) studied the capability of SC-CO$_2$ for the extraction of bioactive compounds from *Persea indica*. In this study, the effect of pressure was evaluated at three different levels ranging from 10 to 20 MPa on the yield of extraction. The results showed that increasing SCF pressure leads to the higher extraction yield, and 20 MPa is the best condition in the studied pressure range to obtain higher yields.

Palavra *et al.* (2011) applied SC-CO$_2$ for the extraction of volatile oils from aromatic plants including *Mentha pulegium* and *Satureja montana* L. SC-CO$_2$ was used at 40 °C and pressures of 12.5, 20, and 30 MPa for the volatile oil extractions. Their finding

showed that the hydrocarbon fraction in the extracted oils increased with the enhancing the pressure of SC-CO$_2$. This trend permits to obtain richer hydrocarbon extracts using higher pressures. Meanwhile, the highest yield of SFE at 30 MPa was achieved (72 g kg^{-1}).

Kagliwal *et al*. (2011) extracted bioactive components from *Seabuckthorn* or *Hippophae rhamnoids* L. (SBT) seed oil using SC-CO$_2$. In this study, the effect of SC-CO$_2$ pressure at a flow rate of 10 g min^{-1} on the extraction yield was evaluated at a constant temperature of 35 °C during 60 min. Different pressures (10, 20, 30, and 40 MPa) were used in this investigation. The results showed that the yield of recovery for SBT active components (tocopherols and carotenes) increases with enhancing the pressure up to 40 MPa (Table 3). These results were obtained because the solvating power of solvent at the SCF state is dependent on the fluid density. Therefore, increasing extraction pressure leads to the higher fluid density and hence the solubility of the analytes increases. As shown in Table 3, the extraction yields for tocopherols and carotenes at pressure of 40 MPa were 77.23% and 75.55%, respectively. Meanwhile, EC$_{50}$ (from DPPH (2,2-diphenyl-1-picrylhydrazyl) assay) of the extract at this pressure was about 49.9 mg mL^{-1}. These results confirm that 40 MPa is optimum pressure for the current SC-CO$_2$, whereas the extracted components at 10 MPa pressure showed no radical-scavenging activity. The analysis of residual SBT active components in the dried SBT seed after SC-CO$_2$ extraction showed that the residual tocopherols and carotenes in the SBT sample after SFE are very low (about 0.5% and 2.56%, respectively) and mass balance could not describe these data; thus, such behavior may be due to degradation of tocopherols and carotenes under SC-CO$_2$ (Cocero *et al*., 2000), which has been reported previously for other carotenoids such as lycopene (Topal *et al*., 2006). The P_c for degradation of carotene depending on the operating temperature is varied from 27 to 35 MPa (Chang and Randolph, 1989). However, the real mechanism for the degradation of the compound due to the SC-CO$_2$ pressure is controversial and may be due to either oxidation of carotenes and production of epoxides (Saldaña *et al*., 2006) or isomerization of β-carotene (Mattea *et al*., 2009).

Lina *et al*. (2010) extracted oil from the *Microula sikkimensis* seeds using SFE. Evaluating the effect of pressure on the yield of seed oil was carried out by SC-CO$_2$ extractions at three different levels of

Table 3 Variation in recovery percentage of SBT actives with pressure at a constant temperature of 35 °C.

Pressure (MPa)	DPPH · (EC50, mg mL^{-1})[a]	Tocopherol (%)[a]	Carotene (%)[a]
10	3706.59 ± 115.32	8.36 ± 0.21	3.58 ± 0.11
20	111.26 ± 4.53	38.14 ± 0.63	18.86 ± 0.57
30	45.09 ± 2.30	62.97 ± 0.77	44.32 ± 0.98
40	49.91 ± 2.49	77.23 ± 0.71	75.55 ± 1.52

[a] Results are mean ± SD of at least three determinations.
Source: Reprinted from Sep Purif Technol, 80, Kagliwal L.D., Patil S.C., Pol A.S., Singhal R.S., Patravale V.B., Separation of bioactives from seabuckthorn seeds by supercritical carbon dioxide extraction methodology through solubility parameter approach, 533–540, Copyright 2011, with permission from Elsevier.
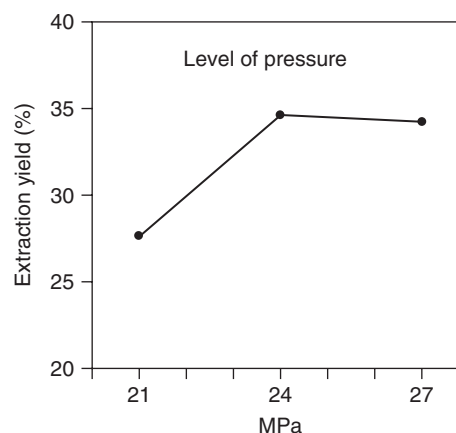
Figure 3 Effect of SC-CO$_2$ pressure on extraction yield of *Microula sikkimensis* seed oil. (*Source*: J Am Oil Chem Soc, 87, 2010, 1221–1226, Supercritical Carbon Dioxide Extraction of Microula sikkimensis Seed Oil, Lina S., Fei R., Xudong Z., Yangong D., Fa H. With kind permission from Springer Science and Business Media.)

pressure (21, 24, and 27 MPa). The results showed that the yield of extraction is influenced considerably with a mean yield from 27.6% at 21 MPa to 34.6% at 24 MPa. Their investigation showed that the effect of pressure on the yield of extracted oil was higher than the other investigated factors such as dynamic extraction time, temperature, and SC-CO$_2$ flow rate. Meanwhile, the maximum yield of oil was obtained at 24 MPa pressure (Figure 3). These results were obtained because increasing pressure may enhance the solubility of plant oil in the SC-CO$_2$ (Jiping Sun and Aiyong, 2002). However, further increasing of the pressure has no considerable role on the yield of plant oil between 24 and 27 MPa (Figure 3).

4.2　Temperature of SC-CO$_2$

One of the most important advantages of SFE technique is the possibility for the extraction of volatile oils from plant matrices via SC-CO$_2$ at moderate temperatures. The SC-CO$_2$ extraction temperature should be set in the lower range 35–50 °C, for the extraction of thermolabile components such as those including essential oils. In other words, in order to avoid degradation of heat-sensitive compounds, the temperature of the fluid should be adjusted in the vicinity of critical point and maintained as low as possible. On the other hand, temperature of the SCF has a considerable effect on the density of the fluid that is important for solubility of the target compounds in SC-CO$_2$ phase. Thus, optimization of this parameter for achieving the highest yield of extraction without degradation of heat-sensitive components of plant extracts is interested and various papers could be found on this subject. Table 4 reviews briefly some of the studies performed on the optimization of SC-CO$_2$ temperature in the extraction and isolation of volatile compounds from herbal samples.

Liu *et al*. (2012) used SC-CO$_2$ extraction for the isolation of volatile oil from pomegranate (*Punica granatum* L.) seeds. In this study, the effect of various parameters including temperature on the changes of yield, chemical composition, and free radical-scavenging activity of volatile oil was investigated. Their results indicated that the temperature of SC-CO$_2$ had a significant effect on the composition of bioactive components and the extracted oil from the seeds. However, as could be seen in Figure 4, the temperature at 30 MPa has no significant effect on the extraction rate and the yield of extracted oil. By increasing the extraction temperature from 35 to 50 °C, the yield improved about 8%, whereas further temperature enhancing caused slower increase in the extraction yield.

Zarena and Udaya Sankar (2011) applied SC-CO$_2$ for the extraction of xanthones from *Mangosteen pericarp* under different pressures (20, 25, and 30 MPa) and various temperatures (40, 50, and 60 °C). The result of experiments was the total xanthone yield extracted on dry weight basis. Figure 5 shows the effect of SC-CO$_2$ temperature on the extraction yield of xanthone. As could be seen in this figure, by increasing extraction temperature from 40 to 60 °C at a constant pressure, the yields of the extracts increased. The finding of their investigation showed that 30 MPa and 60 °C are optimum pressure

and temperature for the extraction, whereas the total yields of xanthone under these optimum conditions of extraction is 7.6% (w/w). These observations showed that increasing temperature and pressure increases the yield of extraction. Because solubility of the most high weight molecular compounds (such as nonvolatile organic compounds) in SC-CO$_2$ is quite low, thus, higher temperatures and pressures are required for substantial loadings (Sauceau *et al*., 2004). Higher temperature decreases the SC-CO$_2$ density, which induces a lower solubility. However, the correlative increase of the vapor pressure enhances the solubility. Meanwhile, temperature increasing also accelerates mass transfer and improves the extraction efficiency (Wang *et al*., 2008). The less yield of extraction at 40 °C may be due to the kinetics of desorption of the components from the pericarp. Increasing the temperature accelerates the desorption process (Rostagno *et al*., 2002).

Ahmed *et al*. (2012) extracted essential oil of Algerian rosemary leaves (*Rosmarinus officinalis*) by the aid of SC-CO$_2$ extraction. They investigated the effect of temperature on the extraction yield of Algerian rosemary essential oil under different temperatures (35, 40, 50, and 60 °C). The results showed that temperature has a complex effect on the extraction yield of plant oil. In pressures of 10 and 14 MPa, the yield of extraction increased with the temperature. However, 18 and 22 MPa have reverse effect. These two opposite effects were observed because increasing the temperature leads to lower SC-CO$_2$ density and thus reduce the solvation capacity of the fluid. On the other hand, the vapor pressure of the solutes enhances and hence their solubility in the SC-CO$_2$ increases.

Turner and McKeon (2002) determined *cis*-vaccenate in the milkweed (*Asclepias sclepias*) seeds via SFE with *in situ* enzymatic methanolysis by considering this fact that temperature has a large impact on enzyme kinetics in SC-CO$_2$, where higher temperatures result in faster reaction rates (Overmeyer *et al*., 1999) and very high temperatures may lead to negative effect on the stability of the enzyme and probably thermal denaturation (Chulalaksananukul *et al*., 1993). They investigated temperature effect on the yield of *cis*-vaccenate from *Asclepias speciosa* seeds because the enzyme reactions in SFE systems at dynamic mode should be fast enough in order to avoid the breakthrough losses of unreacted acylglycerols. Meanwhile, it was found that SCF temperature has no considerable

Table 4 Some of the reported researches on the effect of temperature on the SFE of plants.

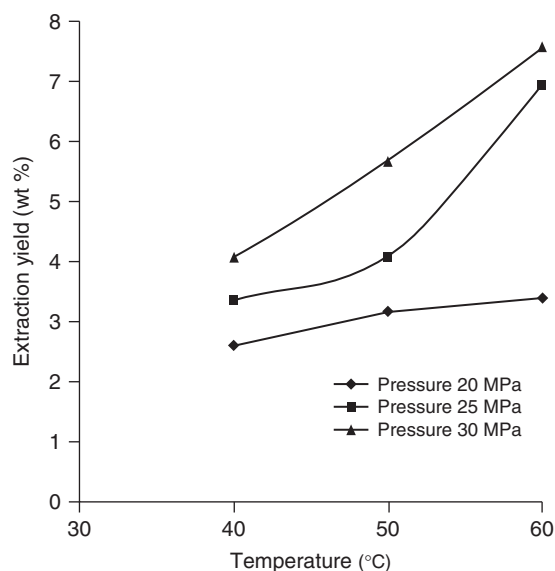| Plant matrix	Range of temperature (°C)	Optimum temperature (°C)	Extraction conditions (pressure, flow rate, etc.)	Observed effect at optimum temperature	Reference
Lippia dulcis Trev.	35–40	40	14 MPa; 5.3×10^{-3} kg min^{-1}; 60 min static extraction time; 260 min dynamic extraction time	Highest yield and maximum percentage of hernandulcin	de Oliveira *et al.* (2012)
Tangor murcote (Blanco)	60–75	75	9.1 MPa; hexane as modifier; conjugated extraction mode	Increasing extraction yield and selective extraction of target analytes	Sargenti and Lanças (1998)
Flax seed (*Linum usitatissimum*)	40, 50, and 60	60	45 MPa; 7.8% EtOH	Maximum extraction of secoisolariciresinol diglucoside	Comin *et al.* (2011)
Trigonella foenum-graecum L.	45, 50, and 55	55	25 MPa; 5 g min^{-1}; EtOH as modifier; 1.5 h extraction time	Maximum extraction of bioactive substance	Tang *et al.* (2007)
Mangosteen pericarp	40–60	60	30 MPa; 300 g g^{-1} solvent ratio to material	Maximum extraction of xanthones	Zarena *et al.* (2012)
Tuna cooking juice (*Thunnus tonggol*)	45, 55, and 65	40	9 MPa	Maximum extraction of fatty acid ethyl esters	Hsieh *et al.* (2005)
Spearmint (*Mentha spicata* L.) leaves	40, 50, and 60	50	20.94 MPa; 7.39 g min^{-1} co-solvent	Highest crude extraction yield of bioactive flavonoid compounds	Bimakr *et al.* (2012)
Peel of *Citrus paradisi*	40.5, 49.5, 58.6, and 63	58.6	9.5 MPa; 15% EtOH	Highest extraction yield of naringin (major flavonoid from the peel of *Citrus paradisi* L.)	Giannuzzo *et al.* (2003)
Acori graminei rhizoma	35, 45, and 55	45	10 MPa; 160 mL min^{-1}; 300 min extraction time	Highest oil yield and selective extraction of β-asarone	Dai *et al.* (2008)
Capsicum annuum L. fruit tissues	40, 50, and 60	40	40 MPa; 2 mL min^{-1}; 116 μm average sample particles diameter	Increasing solubility and extraction yield for capsidiol	Salgın *et al.* (2005)
Geranium (*Pelargonium graveolens*)	40, 50, 60, and 70	40	30 MPa; 4 mL min^{-1}	Maximum extraction yield of oil	Peterson *et al.* (2006)
Okra (*Hibiscus esculentus* L) seeds	40, 50, and 60	60	45 MPa; 6 kg h^{-1} kg	Highest extraction yield and best extraction conditions for sterols	András *et al.* (2005)
Spirulina maxima	20–70	30	18 MPa; 3.33×10^{-5} kg s^{-1}	Maximum extraction of carotene	Canela *et al.* (2002)
Pomegranate (*Punica granatum* L.) seeds	35–50	50	30 MPa	Increasing extraction yield	Liu *et al.* (2012)
Mangosteen pericarp	40, 50, and 60	60	30 MPa	Increasing extraction yield of xanthone	Zarena and Udaya Sankar (2011)
Algerian Rosemary leaves (*Rosmarinus officinalis*)	35, 40, 50, and 60	—	10–22 MPa	Yield of extraction increased in pressures of 10 and 14 MPa, with temperature. Pressures 18 and 22 MPa have reverse effect	Ahmed *et al.* (2012)

(continued overleaf)

Table 4   (*Continued*)

Plant matrix	Range of temperature (°C)	Optimum temperature (°C)	Extraction conditions (pressure, flow rate, etc.)	Observed effect at optimum temperature	Reference
Milkweed (*A. sclepias*) seeds	40, 60, and 80	60	36.6 MPa; 3% (V/V) of MeOH	Novozyme 435 accomplished the highest vaccenate yield at 60 °C; NovoSample 40013 at temperature of 80 °C reaches its highest activity	Turner and McKeon (2002)
Waste hops (*Humulus lupulus* L.)	40–70	50	35 MPa; 50% EtOH	Maximum extraction yield and recovery rate	Guo-qing *et al.* (2005)
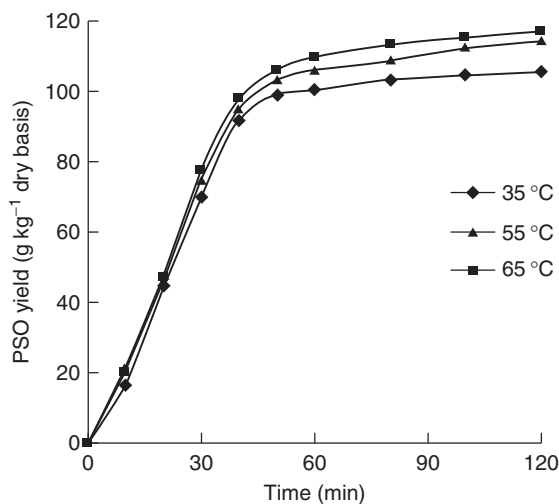
Figure 4   Effect of temperature on the yield of volatile oil extraction via SFE at 30 MPa. (*Source*: Reprinted from Food and Bioproducts Processing, 90, Liu G., Xu X., Gong Y., He L., Gao Y., Effects of supercritical CO_2 extraction parameters on chemical composition and free radical-scavenging activity of pomegranate (Punica granatum L.) seed oil, 333–339, Copyright 2012, with permission from Elsevier.)

Figure 5   Variation of extraction yield by SCF temperature at constant pressure. (*Source*: Reprinted from Separation and Purification Technology, 80, Zarena A.S. and Udaya Sankar K., Xanthones enriched extracts from mangosteen pericarp obtained by supercritical carbon dioxide process, 172–178, Copyright 2011, with permission from Elsevier.)

effect on the kinetics of oil extraction. Therefore, it was assumed that the vaccenate yields obtained in the experiment solely reflected the kinetics of the lipase-catalyzed reaction. SC-CO_2 modified by 3% (v/v) of MeOH was used for extraction at 36.6 MPa under different temperatures of 40, 60, and 80 °C. The obtained results are shown in Figure 6. The figure confirms that enzyme preparation method and temperature have significant effect on the yield of extraction. Although, high stability of Novozyme 435 in pure SC-CO_2 at temperatures higher than 100 °C was demonstrated (Overmeyer *et al*., 1999);

however, Figure 6 shows that Novozyme 435 accomplished the highest vaccenate yield already at 60 °C. This result completely agrees with alcoholysis data obtained for retinyl esters in milk powder (Turner *et al*., 2001). On the other hand, NovoSample 40013 requires a temperature of 80 °C to reach its highest activity under the investigated conditions. However, kinetics correspond to the catalyzed reaction of NovoSample 40013 is still not fast enough to accomplish quantitative vaccenate yields
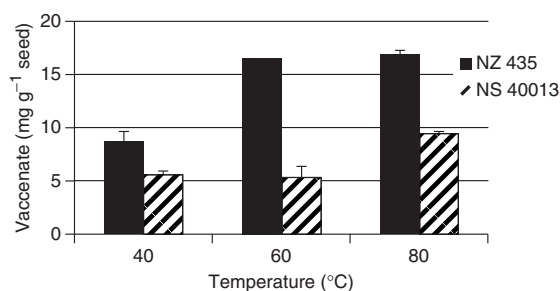
Figure 6 Effect of various temperatures on the extraction yields of *cis*-vaccenate from *A. speciosa* seeds using Novozyme 435 and NovoSample 40013 ($n = 3$). SC-CO$_2$ was modified with 3% (v/v) of MeOH at 36.6 MPa. The error bars correspond to the values of relative standard deviations (RSD; %). NZ: Novozyme 435 and NS: NovoSample 40013. (*Source*: J Am Oil Chem Soc, 79, 2002, 473–478, The use of immobilized Candida antarctica lipase for simultaneous supercritical fluid extraction and in-situ methanolysis of cis-vaccenic acid in milkweed seeds, Turner C. and McKeon T. With kind permission from Springer Science and Business Media.)

at the studied temperatures. The previous study demonstrated that the initial activity of Novozyme 435 is higher than the initial activity of NovoSample 40013 (Husum *et al.*, 2001). Therefore, Novozyme 435 was selected as the preferred immobilized *C. antarctica* lipase B preparation for application in dynamic mode of SFE systems. Meanwhile, performing the reaction at 60 °C is preferred because the lower temperature improves long-term stability of the enzyme. However, 80 °C is the best investigated temperature for NovoSample 40013.

Guo-qing *et al.* (2005) extracted flavonoids components from waste hops (*Humulus lupulus* L.) using SC-CO$_2$. Various parameters such as temperature, pressure, concentrations of EtOH as modifier, and the ratio of solvent to the sample material were investigated in this study. Their finding showed that the extraction yield increases by rising temperature from 40 to 50 °C, whereas the extraction yield shows reverse trend by increasing temperature from 50 to 60 °C. Furthermore, temperature enhancing from 60 to 70 °C again increases the extraction yield. However, the recovery rate is lower than that at 50 °C. They described this surprising trend in the extraction yields by considering overall effect of SCF temperature on the extraction, which is the competition between the increasing in solubility of analyte compounds and the reduction in SC-CO$_2$ density because of the temperature rise (Chiu *et al.*, 2002).

4.3 Effect of Modifier

As the solubility of polar organic compounds in SC-CO$_2$ is limited, the effective extraction of these compounds with pure SC-CO$_2$ is not possible. The additional small amounts of a polar modifier (co-solvent) to the SC-CO$_2$ have been shown to offer enormous increases in the extraction yield of polar organic targets. On the other hand, interaction of a specific modifier with plant sample matrix containing target compounds with various polarities leads to the various extraction yields; thus, selecting the optimum modifier to enhance extraction yield of a group target compounds theoretically is not possible and performing the experiment is essential (Rahimi-Nasrabadi *et al.*, 2012c; Pourmortazavi *et al.*, 2003, 2004b). In order to investigate the effect of various modifiers at a constant pressure and temperature on the yield of extraction efficiency, different volumes of the modifier are added either directly onto the plant matrix in the extraction vessel before the extraction (Ebrahimzadeh *et al.*, 2003) or indirectly via premixing with the CO$_2$ source or simultaneous pumping both of CO$_2$ and co-solvent (Huang *et al.*, 2011a, 2011b). Increase in the extraction yield by the addition of EtOH as co-solvent to the SC-CO$_2$, especially for the extraction of herbal components, has been reported by several authors (Nguyen *et al.*, 2011). The increase could be due to the swollen cellular structure by EtOH, which facilitates the penetration of SC-CO$_2$ into the sample matrix (Casas *et al.*, 2007; Lee *et al.*, 2010). However, some authors (Guclu-Ustundag and Temelli, 2005) reported that this increase is due to the increase in the local density of SCF phase by the addition of EtOH as co-solvent. Nguyen *et al.* (2011) extracted *Moringa oleifera* kernels via SFE. They initially added 10% and 15% EtOH directly to the extractor to combine with SC-CO$_2$ in order to investigate other operating conditions (pressure of 22.5 MPa, temperature of 47.5 °C, CO$_2$ flow rate of 0.5 m^3 h^{-1}, and average particle size of 1.12 mm). The results showed that the addition of 10% EtOH onto the SC-CO$_2$ increases the extraction yield up to 10%; this phenomenon may be due to the role of EtOH in the swelling of the cellular structure of the *M. oleifera* kernels, which facilitated the SC-CO$_2$ penetration to the oil-rich area of the plant cells. In this case, by considering this fact that at CO$_2$ flow rate of 0.5 m^3 h^{-1}, most of EtOH added to the extractor is nearly exhausted during the first 30 min

of the extraction process, and the hypothesis that enhancing the extraction yield because of the rising local density of the solvent is minor. Meanwhile, increasing the EtOH ratio from 10% to 15% shows no significant effect on the oil extraction yield.

Veggi *et al.* (2011) studied the modifier effects on the quality of SFE extracts of some Brazilian plants (*Heteropterys aphrodisiaca*, *Pyrostegia venusta*, *Inga edulis*, *Hymenaea courbaril stilbocarpa*, and *Phaseolus vulgaris* L.). The effectiveness of SFE in the extraction of a rich composition in antioxidant was determined by measuring the overall yields of these plants at a constant pressure and temperature using pure CO_2 and CO_2 as well as EtOH (10% v/v) as modifier. The results of this study showed that the overall yields of extracts richen by antioxidants for all plants increase using the EtOH as co-solvent of CO_2. Among all SFE extracts, jatobá (*H. courbaril stilbocarpa*) possesses the highest antioxidant activity, about 10% and 14% scavenging ability using pure CO_2 and CO_2 modified by EtOH, respectively. Meanwhile, the results showed that the addition of EtOH as modifier in SFE process enhances the extraction efficiency of all raw materials. On the other hand, in some of the samples including jatobá, ingá-cipó (*I. edulis*), and nó-de-cachorro (*H. aphrodisiaca*), using the EtOH as co-solvent caused a great increase in the antioxidant activity of SFE extracts. However, antioxidant recovery for other plants such as bean and cipó-de-são-joão (*P. venusta*) was decreased by the addition of EtOH as co-solvent to the CO_2. In other words, the extracts of jatobá, ingá-cipó, and nó-de-cachorro obtained by SFE using CO_2 modified by EtOH are better antioxidant components sources, whereas SFE with pure CO_2 leads to better antioxidant activity for bean and cipó-de-são-joão.

Andrade *et al.* (2012) described the chemical composition and the antioxidant activity of spent coffee grounds (*Coffea arabica*) and coffee husks extracted by SFE with pure CO_2 and CO_2 + co-solvent. The results of this investigation showed that the addition of the co-solvent to the CO_2 during the extraction with SC-CO_2 enhances the yields of extracts for both coffee samples. The yields intensifying the extracts of spent coffee grounds were more evident. In the identical pressure and temperature (10 MPa and 40 °C), SFE with pure CO_2 yields 0.5 (±0.1)%, whereas in the presence of 8% and 15% EtOH as co-solvent, the yield was increased to 6.3 (±0.5)% and 14 (±2)%, respectively. This increase in the

yield is due to the increase in the solubility of polar compounds in the modified SC-CO_2 in comparison with their solubility in pure SC-CO_2 phase. Meanwhile, the results of this research showed that the number of extracted components by the modified SC-CO_2 increases, which leads to the reduction of the selectivity of extraction process.

Arnaáiz *et al.* (2011) studied SFE and fractionation of lipids from broccoli leaves (*Brassica oleracea* L. var. *Italica*). In this study, some of the effective variables in the extraction process were optimized as temperature of 60 °C, pressure of 30 MPa, and SC-CO_2 flow rate of 3 mL min^{-1}. In order to investigate the effect of modifier on the extraction results, two different fractions were obtained. The first fraction was obtained by the extraction of the samples with pure SC-CO_2. The second fraction was obtained via the extraction of leftover sample material after first step of extraction, whereas in this step, the SC-CO_2 was modified with 15% of MeOH. The fraction obtained with the pure CO_2 was constituted neutral lipids because of the nonpolar nature of CO_2. On the other hand, the residual leaves material extracted by modified CO_2 with the MeOH possesses polar lipids and increasing the ratio of MeOH enhances the total fatty acid content in the extracted lipids (Figure 7). In this case, using 15% MeOH provides the best results. Furthermore, under these optimum conditions, dynamic extraction time of 60 min yields about 80% for the lipids.
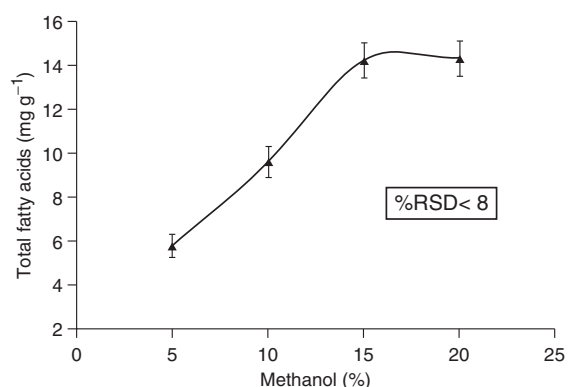
Figure 7 Effect of MeOH ratio as co-solvent on SFE; (extraction conditions: pressure 30 MPa, temperature 60 °C, flow rate of 3 mL min^{-1}, and dynamic extraction time of 30 min). (*Source*: E. Arnaáiz, *et al.* (2011). Reproduced by permission of WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.)

Akay *et al*. (2011) studied the effects of process parameters on the SC-CO_2 extraction of phenols from strawberry (*Arbutus unedo* L.) fruits. Evaluation of the effect of various parameters such as pressure (5–30 MPa), temperature (30–80 °C), and concentration of EtOH as modifier (0–20%) by CO_2 flow rate of 15 g min^{-1} for 60 min was performed using a Box–Behnken statistical design. As shown in Figure 8, co-solvent ratio is the most effective factor, whereas temperature and pressure are not statistically significant as much as co-solvent for determining the total yield of extracted phenol. In addition, the finding of this investigation demonstrated that co-solvent has the most predominant effect on the composition of extracted phenolic compounds.

As discussed, SC-CO_2 as a nonpolar solvent acts efficiently for the extraction of lipophilic analytes, whereas it has a low affinity for polar compounds. Therefore, for effective solubility of polar targets in SC-CO_2 phase, addition of a polar co-solvent is required to increase its solvating power toward polar molecules. In the literature reviewing the extraction of volatile components of herbal samples, especially oxygenated compounds, various investigations for the selection of an efficient co-solvent and optimization of its value could be found. Table 5 presents some of the reported studies on the effect of modifier on the extraction of herbal samples via SC-CO_2.

4.4 Effect of CO_2 Flow Rate

The rate of the SC-CO_2 flowing through the extraction vessel could strongly affect the extraction efficiency. The lower rate of the fluid passing caused deeper penetration of SCF to the plant matrix. The fluid speed could be defined as the linear velocity, which is strongly dependent on both the flow rate and the extraction vessel geometry. In the extraction vessel, the flow rate could be easily changed by varying the inside diameter of restrictor. In many SCF extractions, the value of the SC-CO_2 flow rate has a significant effect on the results (Machmudah *et al*., 2012). Raising SC-CO_2 flow rate leads to the increase in the number of CO_2 molecules in contact with the solute molecules. This more contact increases the intermolecular interaction between CO_2 and the solute, which caused raising dissolution of the solute in the fluid phase (Machmudah *et al*., 2012). Furthermore, enhancing the SC-CO_2 flow rate

may decrease the external mass transfer resistance. However, the higher SC-CO_2 flow rate decreases its residence time, which leads to the lower interaction time between CO_2 and the solute. In addition, the higher flow rates may cause channeling effect, which forced the SC-CO_2 solvent through the sample particles at such a high rate that it passes only around the solid matrix without diffusing to the pores within the plant sample. Meanwhile, the high values of CO_2 flow rate may consolidate the plant sample and restrict the SC-CO_2 movement into and out of the sample, which leads to the reduction of the output of CO_2 passed through the sample (Tonthubthimthong *et al*., 2001). The above discussion confirms that the yield of extraction process is strongly dependent on the value of solute solubility in the SC-CO_2 phase.

Nyam *et al*. (2011) investigated SC-CO_2 extraction of oil from Kalahari melon seeds (*Citrullus lanatus*). The results showed that among all studied extraction parameters, such as pressure, temperature, and SC-CO_2 flow rate, the last parameter strongly affects the extraction of plant oil. As could be seen in Figure 9, the yield of oil extraction was increased by raising the SC-CO_2 flow rate from 10 to 15 mL min^{-1}, because in this step of flow raising, the extraction of the surface attached oil occurs and the increase in the CO_2 flow rate reduces the external mass transfer resistance, which leads to a higher plant oil yield. However, further raising the SC-CO_2 flow rate (20 mL min^{-1}) decreases the oil recovery (Figure 9). This reduction was observed because the extraction mechanism is controlled mainly by diffusion of oil from the plant cell matrix and the increase in the CO_2 flow rate reduces the required SCF diffusion time. Thus, 15 mL min^{-1} is the optimum SC-CO_2 flow rate to achieve the highest percentage of Kalahari oil recovery.

Senyay-Oncel *et al*. (2011) optimized SFE parameters such as pressure, temperature, and flow rate for obtaining the maximum yields of unsaturated fatty acids extracted from *Pistacia terebinthus* berries using a Box–Behnken statistical design. The results showed that the investigated temperature and flow rates are not statistically significant in the SFE yields. However, the interaction between the pressure and the flow rate is significant. As can be seen in Figure 10, at the flow rates between 12 and 20 g min^{-1}, the synergetic effect of temperature and the flow rate provides higher yields. On the other hand, the interaction between the pressure and the
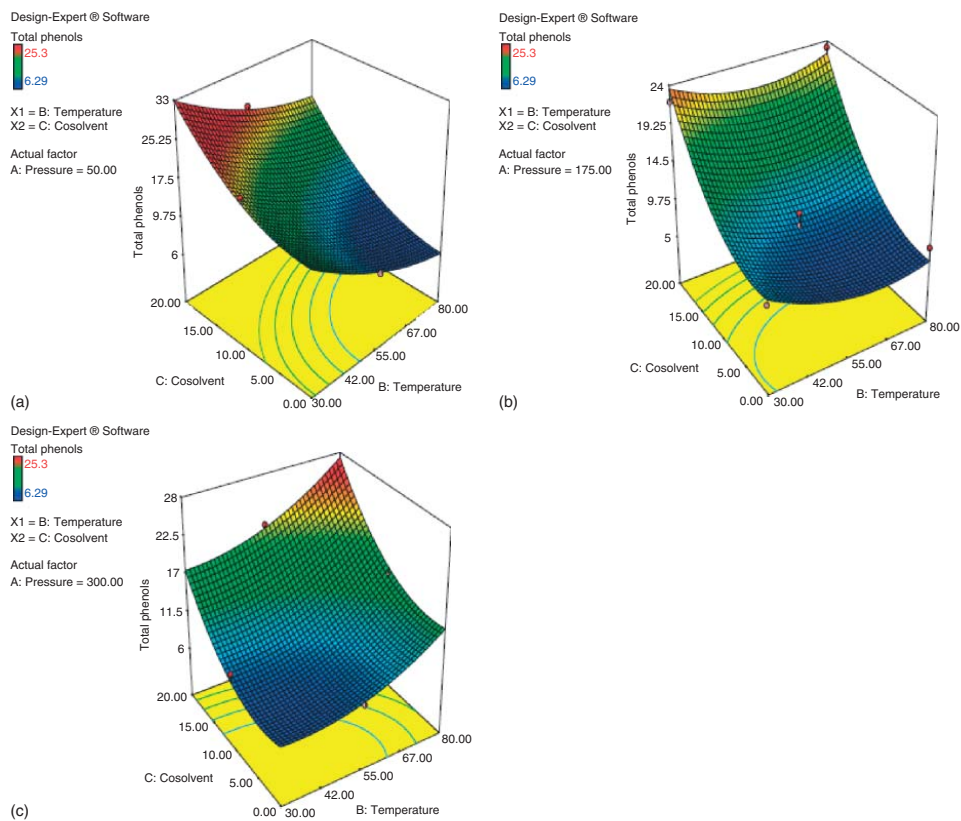
Figure 8   Plots of 3D response surface correspond to the effects of EtOH as co-solvent and temperature at various pressures (A, 4 MPa; B, 14 MPa, and C, 30 MPa). (*Source*: S. Akay, *et al*. (2011). Reproduced by permission of WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.)

Table 5 Some of the investigation on the effect of co-solvent on the plant components extraction via SC-CO_2.

Plant sample	Studied co-solvent	Efficient co-solvent	Observed result by adding co-solvent	Extraction conditions (temperature; pressure)	References
Olive husk (*Olea europa*)	EtOH; 1% and 5% (v/v)	EtOH	Improvement in the total yield of extraction and contents of carotenoids and chlorophyll	40–60 °C; 30–35 MPa	Gracia *et al*. (2011)
Guava (*Psidium guajava* L.) seeds	EtOH	EtOH	Increasing extraction yield	40 °C; 30 MPa	Castro-Vargas *et al*. (2011)
Canola press cake (*Brassica napus L.*)	EtOH; 2.3–10% (v/v); 1 h treatment with sample in extraction vessel	EtOH; 2.3%	Enhancing phospholipid extraction yield	60 °C; 40 MPa; 70 g min^{-1} CO_2 flow rate	Li *et al*. (2010)
Flowers of marigold (*Calendulae flos*), hawthorn (*Crataegi folium* cum flore), and chamomile (*Matricariae flos*)	EtOH; 0–20% (v/v)	EtOH; 20%	Increasing total extraction yield	50 °C; 30 MPa	Hamburger *et al*. (2004)
Jatropha curcas LINN. leaves	MeOH; 0–70% (v/v)	MeOH 70%	Increasing extraction yields about 2.01 times by enhancing modifier concentration from 0% to 70%	50 °C; 40 MPa	Manpong *et al*. (2011a)
Jabuticaba (*Myrciaria cauliflora*) byproducts	EtOH; 0–20% (v/v)	EtOH; 20%	Increasing extraction yield of antioxidant compounds	50 °C; 20 MPa	Cavalcanti *et al*. (2011)
Khoa (*Satureja boliviana* Benth Briq) leaves	EtOH; 0–1% (v/v)	EtOH; 1%	Significant influence on the mass of extract	20 °C; 7 MPa	Hatami *et al*. (2011)
Alnus glutinosa (L.) Gaertn., alder bark	EtOH; 0%, 5%, and 10% (v/v)	EtOH; 10%	Increasing the extraction yield	60 °C; 30 MPa	Felföldi-Gáva *et al*. (2012)
Soybean meal (*Glycine max*)	MeOH; 5.4%, 7.8%, and 10.2% (v/v)	MeOH; 10.2%	Increasing isoflavones yield and total extraction yield	40 °C; 50 MPa	Kumhom *et al*. (2011)
Fruit bodies of *Ganoderma sinense*	EtOH; 0–10%	EtOH; 10%	Higher extraction yield	40 °C; 30 MPa	Chen *et al*. (2011)
Passiflora edulis leaves	EtOH, MeOH, EtAc, CHCl$_3$	MeOH;10%	Increasing extraction yield of flavonoids	60 °C; 7.5 MPa	de Lourdes *et al*. (1997)
Anoectochilus roxburghii	EtOH, MeOH, EtAc	EtOH	Increasing extraction yields of β-sitosterol, stigmasterol and ergosterol	45 °C; 25 MPa	Huang *et al*. (2009)
Haematococcus pluvialis	EtOH; 3.08–12.31 mL g^{-1} (per weight of matrix)	9.23 mL g^{-1}	Increasing astaxanthin yield	50 °C; 30.4 MPa; 6 mL min^{-1} CO_2 flow rate; 20 min extraction time	Pan *et al*. (2012)
Rhizomes of turmeric (*Curcuma longa* L.)	EtOH	EtOH	Increasing extraction yield of curcuminoids and decreasing solvent consumption	70 °C; 30 MPa	Chassagnez-Méndez *et al*. (2000)

Table 5   (*Continued*)

Plant sample	Studied co-solvent	Efficient co-solvent	Observed result by adding co-solvent	Extraction conditions (temperature; pressure)	References
Jatropha curcas LINN. leaves	MeOH; 30–70% (v/v)	MeOH; 30%	Increasing extraction yield of gallic acid, corilagin, and ellagic acid	60 °C; 10 MPa	Manpong *et al.* (2011b)
Moringa oleifera kernels	10% and 15% EtOH	10% EtOH	Increasing extraction yield	47.5 °C; 22.5 MPa; CO_2 flow rate 0.5 $m^3 h^{-1}$; particle size 1.12 mm	Nguyen *et al.* (2011)
Some Brazilian plants (*Heteropterys aphrodisiaca, Pyrostegia venusta, Inga edulis, Hymenaea courbaril stilbocarpa,* and *Phaseolus vulgaris* L.)	10% (v/v) EtOH	10% (v/v) EtOH	Increasing overall yields of extracts richen by antioxidants for all plants	50.0 °C; 35 MPa; 10% (v/v) EtOH	Veggi *et al.* (2011)
Spent coffee grounds (*Coffea arabica*) and coffee husks	8% and 15% (v/v) EtOH	15% EtOH	Increasing extracts yields	60.0 °C; 10 MPa	Andrade *et al.* (2012)
Broccoli leaves (*Brassica oleracea* L. var. *italica*)	5–20% (v/v) MeOH	15% (v/v) MeOH	Increasing extraction yields of polar lipids and total extracts	30 MPa; 60 °C; CO_2 flow rate 3 mL min^{-1}	Arnaáiz *et al.* (2011)
Strawberry (*Arbutus unedo* L.) fruits	0–20% (v/v) EtOH	19.7% (v/v) EtOH	Highest total phenols/g content and radical-scavenging capacity	5–30 MPa; 30–80 °C; CO_2 flow rate 15 g min^{-1}; extraction time 60 min	Akay *et al.* (2011)

flow rate leads to the achievement of maximum oleic acid yield at pressure of 25 MPa and temperature of 55 °C regardless to the flow rate. In this case, it was predicted that the highest yield for oleic acid (51.55%) will be obtained at the pressure of 24 MPa, temperature of 54 °C, and flow rate of 17 g min^{-1}, which is compatible with the obtained experimental yield of 51.19% at pressure of 20 MPa, temperature of 55 °C, and flow rate of 15 g min^{-1}.

Cassel *et al.* (2000) studied dependence of the *Baccharis dracunculifolia*, "Vassoura," oil composition obtained by SC-CO_2 extraction on the operation conditions, that is, CO_2 flow rate. Figure 11 shows that among the tree investigate flow rates (0.5, 1.0, and 2.0 mL min^{-1}), the optimum CO_2 flow rate for the extraction of Vassoura oil is 1.0 mL min^{-1}, whereas the extraction yield at the flow of 2.0 mL min^{-1} was lower because of the partial loss of extracted oil by

CO_2 bubbling in the collector solvent during product collection step.

Gopalan *et al.* (2000) extracted turmeric oil from turmeric (*Curcuma longa*) using SC-CO_2 in a semicontinuous-flow extractor. The extraction yields of turmeric oil as a function of CO_2 flow rates using the sample with average particle diameter of 0.21 mm were studied. As shown in Figure 12, the extraction rate was enhanced with the increase of SC-CO_2 flow rate. On the other hand, the extraction yield at different flow rates enhances by raising the ratio of CO_2 amount/sample weight (Figure 12).

4.5   Effect of Extraction Time

The extraction time could influence the yield of extracted oil via SC-CO_2 process. The effect of extraction time in SFE may be studied in two
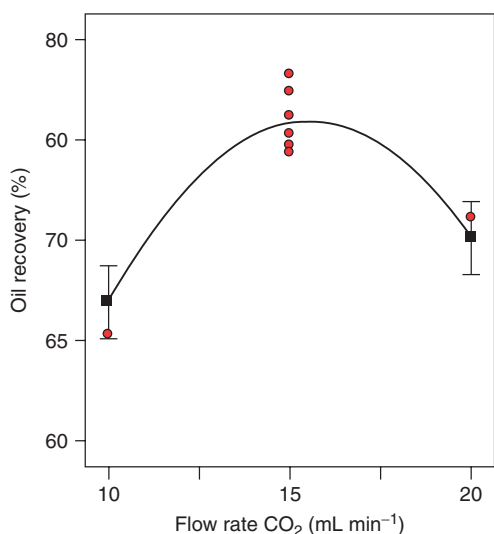
Figure 9 Effect of SC-CO_2 flow rate on Kalahari melon (*Citrullus lanatus*) oil recovery. (*Source*: Food Bioprocess Technol, 4, 2011, 1432–1441, Optimization of supercritical CO_2 extraction of phytosterol-enriched oil from Kalahari melon seeds, Nyam K.L., Tan C.P., Lai O.M., Long K., Che Man YB. With kind permission from Springer Science and Business Media.)

different modes of dynamic extraction time or static extraction time followed by dynamic extraction time. The previous report (Pourmortazavi

et al., 2004a) showed that increasing the extraction time may be also caused variation in the composition of the extracted oil. As an example, at a constant pressure, increasing the dynamic extraction time leads to enhancing the content of heavy compounds with large retention indices in the plant oil.

Guoliang *et al.* (2011) optimized the extraction parameters including SCF pressure, temperature, and time for the oil extraction from *Lycium barbarum* seeds using an orthogonal test design. As shown in Figure 13, the extraction yield was enhanced by increasing the extraction time. In this study, it was found that the yield of extraction increased markedly with the time and reached the maximum at 60 min. However, the extraction yield increased slowly and tended to be saturated at higher extraction times (i.e., 120 min). Thus, in this case, 60 min was employed as optimum extraction time.

Park *et al.* (2012) optimized decaffeination of green tea (*Camellia sinensis*) using SC-CO_2 by the aid of response surface methodology (RSM) for achieving the maximum removal of caffeine from the plant. Figure 14 shows the effect of extraction time at other optimized conditions of decaffeination (pressure of 23 MPa, temperature of 63 °C,
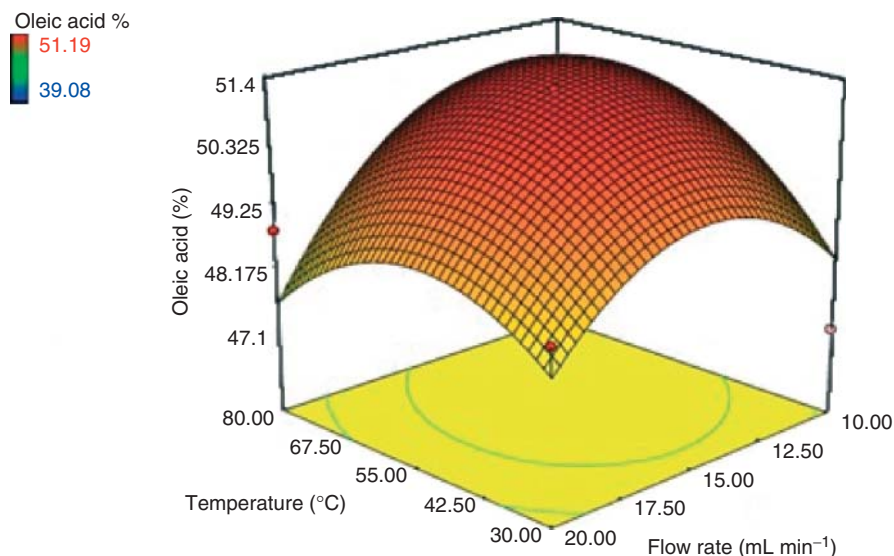
Figure 10 3-D and contour response surface plots for the effects of flow rate and temperature on oleic acid yield at the constant pressure of 20 MPa. (*Source*: J Am Oil Chem Soc, 88, 2011, 1061–1069, Effects of Supercritical Fluid Extraction Parameters on Unsaturated Fatty Acid Yields of Pistacia terebinthus Berries, Senyay-Oncel D., Ertas H. and Yesil-Celiktas O. With kind permission from Springer Science and Business Media.)
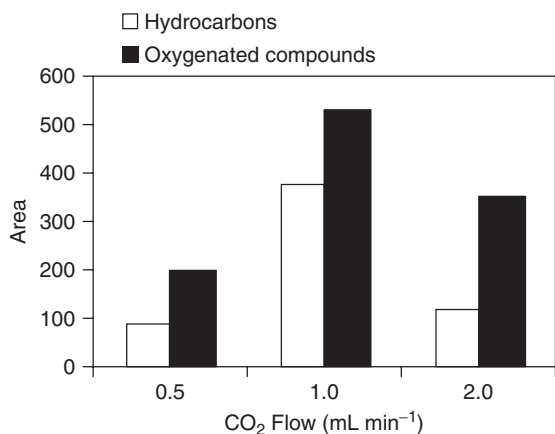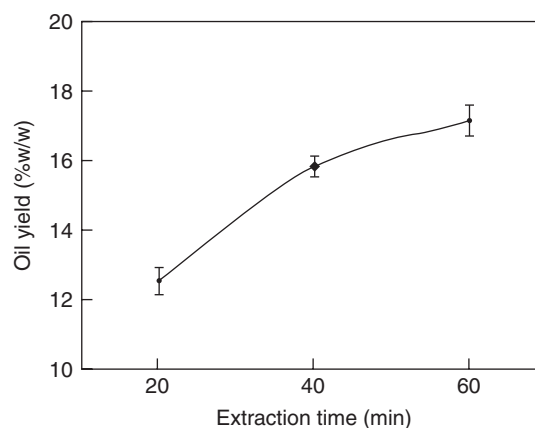
Figure 11 Yield of Vassoura (*Baccharis dracunculifolia)* oil extract at various CO_2 flow rates at the constant temperature of 323.15 K and pressure of 10 MPa. (*Source*: Reprinted with permission from Cassel E, Frizzo CD, Vanderlinde R, Atti-Serafini L, Lorenzo D, Dellacassa E. 2000. Extraction of baccharis oil by supercritical CO_2. Ind Eng Chem Res 39: 4803–4805. Copyright 2000 American Chemical Society.)
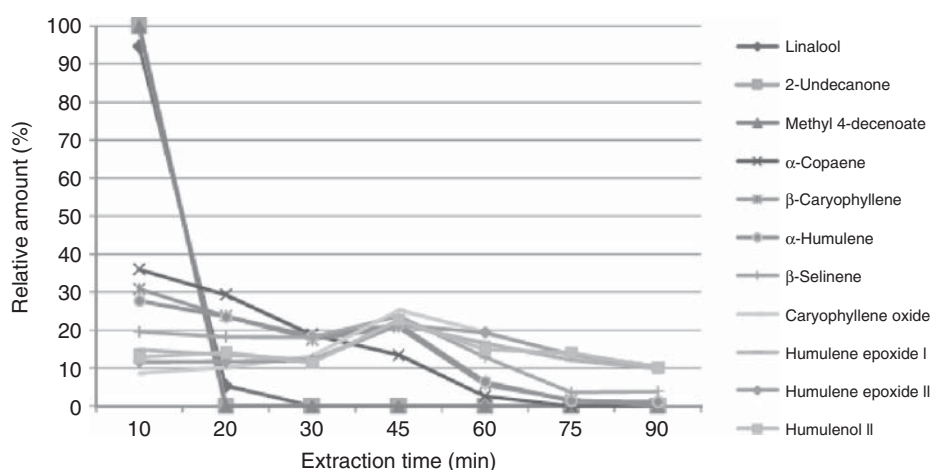

Figure 13 Effect of extraction time on extraction yield of *Lycium barbarum* seed oil. (*Source*: Reprinted from LWT – Food Sci Technol, 44, Guoliang L., Junyou S., Yourui S., Zhiwei S., Lian X., Jie Z., Jinmao Y., Yongjun L., Supercritical CO_2 cell breaking extraction of Lycium barbarum seed oil and determination of its chemical composition by HPLC/APCI/MS and antioxidant activity, 1172–1178, Copyright 2011, with permission from Elsevier.)

Figure 12 Effect of SC-CO_2 flow rate on the extraction yield of turmeric oil (*Curcuma longa*) at temperature of 313 K and pressure of 30 MPa. (*Source*: Reprinted with permission from Gopalan B, Goto M, Kodama A, Hirose T. 2000. Supercritical carbon dioxide extraction of turmeric (Curcuma longa). J Agric Food Chem 48: 2189–2192. Copyright 2000 American Chemical Society.)

and about 3% EtOH as modifier). As shown in the figure, 70% of the initial amount of caffeine was extracted in the first 60 min of the extraction time and then the extraction rate was decreased. The reason for such observation is the large amounts of caffeine retained at the surface or in the exterior locations of green tea, which are extracted in the first 60 min of extraction time by SC-CO_2 at relatively higher rates. The results showed that 96.6%,

40.6%, and 43.1% of the total caffeine, catechin, and chlorophyll were respectively removed from the green tea after 120 min of extraction. It means that caffeine is nearly entirely removed from the green tea, whereas the main quantities of health-benefiting components such as catechins and chlorophylls are remained in the decaffeinated leaves. However, a considerable amount of total chlorophyll is co-extracted during this decaffeination period. This removal of chlorophylls causes negative effects on the taste properties of green tea and hence has an adverse role in the commercializing of the decaffeinated plant (Lee *et al*., 2009).

Van Opstaele *et al*. (2012) reported selective isolation of total hop essential oil from hop pellets (*H. lupulus* L.) by SFE. The finding showed that extraction time has a very important effect in the recovery of total hop oil by SFE. Figure 15 displays the extraction yields for some selected hop oil markers as a function of SFE time. As could be seen in the figure, at the first of extraction period (after 20–30 min), most of the volatile marker components in floral region (such as linalool, 2-undecanone, and methyl 4-decenoate) are totally extracted. On the other hand, isolation of both the sesquiterpene hydrocarbons (e.g., α-humulene) and the oxygenated sesquiterpenoids (e.g., humulene epoxide II) needs the longer extraction time (90 min) in order to their relative
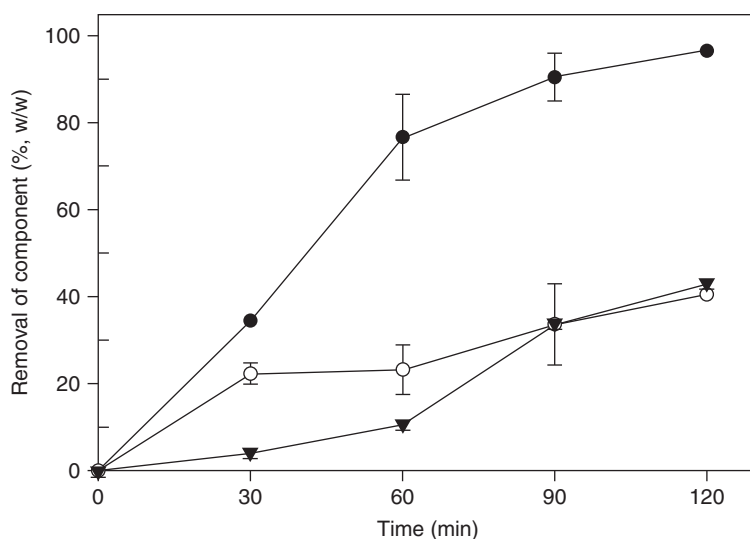
Figure 14 Effect of extraction time on the yields of caffeine (●), total catechins (○), and chlorophylls (▼) from green tea (*Camellia sinensis*) leaves at the optimum conditions for decaffeination; conditions: pressure of 23 MPa, temperature 63 °C, CO_2 flow rate of 8.5 g min^{-1}, and 3% EtOH as modifier). (*Source*: H.S. Park, *et al*. (2012). Reproduced by permission of Elsevier.)


Figure 15 Effect of SCF extraction time on extraction behavior of selected hop (*Humulus lupulus* L.) oil markers (CO_2 density was 0.50 g mL^{-1}). (*Source*: Reprinted from J Supercrit Fluid, 69, Van Opstaele F., Goiris K., De Rouck G., Aerts G., De Cooman L., Production of novel varietal hop aromas by supercritical fluid extraction of hop pellets – Part 1: Preparation of single variety total hop essential oils and polar hop essences, 45–56, Copyright 2012, with permission from Elsevier.)

amount in the plant matrix that becomes approximately 0% and 10%, respectively. Such behavior of hop oil components is due to the fact that the extraction of monoterpenes takes place at the beginning of the extraction process, whereas sesquiterpenes and oxygenated components need the longer extraction times because of their higher molecular weight and polarity.

5 SAMPLE PROPERTIES IN SFE

5.1 Effect of Plant Particle Size

Particle size of plant sample is one of the main parameters affecting the extraction efficiency in SFE using CO_2. Ab Rahman *et al*. (2012) used SC-CO_2 to extract the oil from palm kernel matrix (*Elaes*
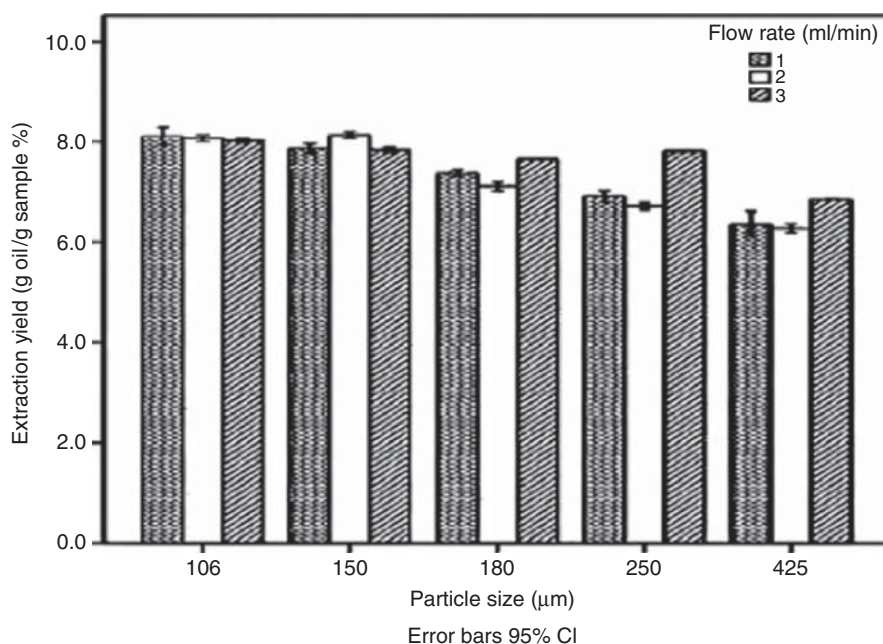
Figure 16 Plant particle size effect on the yield of extraction; operating conditions: temperature of 70 °C, pressure of 34.47 MPa, and SC-CO_2 flow rates of 1, 2, 3 mL min^{-1} for 60 min. (*Source*: Reprinted from J Food Eng, 108, Ab Rahman N.N., Al-Rawi S.S., Ibrahim A.H., Ben Nama M.M., Omar Ab Kadir M., Supercritical carbon dioxide extraction of the residual oil from palm kernel cake, 166–170, Copyright 2012, with permission from Elsevier.)

guineensis). The plant particle size effect on the yield of plant oil was studied using various particle sizes obtained by sieving from the mesh of ≤106, ≤150, ≤180, ≤250, and ≤450 μm. At temperature of 70 °C and pressure of 34.47 MPa with the CO_2 flow rate of 2 mL min^{-1}, as shown in Figure 16, the highest extracted yield was about 8.2%. The palm kernel sample with the particle size of ≤150 μm required the CO_2 flow rate of 2 mL min^{-1} to achieve maximum yield of extracted oil, whereas at smaller particle size of ≤106 μm, slower flow rate is essential to achieve the same yield of oil. The yield of extracted oil from the palm kernel with particle size of ≤425 μm is lower, probably because of restricting the solvent flow by matrix. Removal of the oil from particles with the size larger than 150 μm required higher flow rate, whereas the oil from the plant matrix with the particle size ≤150 μm is extracted effectively with the flow rate of 2 mL min^{-1}. However, plant sample with the particle size ≤106 μm needs slower CO_2 flow rate of 1 mL min^{-1} for effective removal of the oil. It should be noted that the effect of CO_2 flow rate on the extraction yield for the plant sample with smaller particle size is not considerable.

However, increasing the plant particle size leads to clearing the effect of CO_2 flow rate. As could be seen in Figure 16, extraction yield for larger particles enhances by raising the SCF flow rate.

It was explained (Salgın and Salgın, 2006; Salgin, 2007) that the extraction rate decreased as the particle size of plant matrix increased because of reducing the solubility that arisen by increasing the particle size, as well as reducing specific surfaces area of the oilseed material. The analysis of variance showed that the oil extraction yields were significantly dependent on the particle size of plant sample, where by decreasing sample particle size, the amount of collected oil was increased. The reason for oil yield increment is the large surface area contact created by the smaller pieces of the palm kernel cake and CO_2 as extracting solvent. Some reports in the literature explained that the reduction in plant size increases the amount of extracted oil (Boutin and Badens, 2009; del Valle and Uquiche, 2002). For example, Machmudah *et al*. (2007) proposed that the reduction in particle size may cause opening up the plant cells and hence increasing the solvent contact with the surface of solid sample. The reason for such behavior could be

explained as follow: slower SCF flow rates provide enough interaction time between the plant particles and carbon dioxide and hence more solvating of the plant oil in the SCF.

Salgın and Korkmaz (2011) studied the SFE of pumpkin seed (*Cucurbita pepo* L.) oil. The effect of main extraction parameters, that is, plant particle size (250–2360 μm), volumetric flow rate of SC-CO$_2$ (0.06-0.30 L h^{-1}), pressure (20–50 MPa), temperature (40–70 °C), and type and amount of modifier (EtOH and *n*-hexane) on the extraction yield was investigated. Figure 17 shows the yield of extraction versus amount of SCF solvent or SFE time for different particle size of plant sample. It was found that the yield of pumpkin seed oil increased by decreasing particle size, increasing the amount of SCF solvent, and rising the SFE time. The extraction curves in the figure show two linear distinct periods. In the first linear period, the rate of oil recovery was very fast because of the preparation of plant sample with smaller particles and narrower particle size distribution via mechanical grinding of plant materials. This process caused producing smaller particles because

of the disruption in the cell walls of oilseed and hence an increase in the specific surface area and reducing the diffusion path. Furthermore, the intraparticle resistances to the mass transfer may be broken as a result of this disruption, and thus enhance the extraction yield considerably.

The surfaces of particles were covered with more oil after the grinding process. Thus, the amount of releasing oil on the external surfaces of the particles could be increased by the reduction in particle size and hence increasing the surface area of the pumpkin seeds in contact with SC-CO$_2$ solvent. The required extraction time to reach plateau in the yield curve has been increased by decreasing particle size ranges. The extraction yields for a 100-min extraction period were 0.45, 0.41, 0.24, and 0.16 g oil/g of dry seed for the particle size range 250–600 μm, 600–1180 μm, 1180–1700 μm, and 1700–2360 μm, respectively. On the other hand, the yield of oil in the next linear period is very slow. Especially, the SC-CO$_2$ could not almost reach to embedded oil inside the large size of pumpkin seed because of the resistance of large mass transfer. In the longer extraction time, the extraction yield may be reached a certain value to complete the recovery of the plant oil. For each sample, the extraction yield increases with various slopes until a plateau value is reached; above this value, the extraction yield is independent to the above factors. In fact, the plateau region shows the extent efficiency of the investigated factors (amount of SCF solvent and SFE time) to enhance the extraction yield.

Liu *et al*. (2011) reported SC-CO$_2$ extraction of flavonoids from *Maydis stigma*. The effects of various variables including sample particle size, extraction time, and co-solvent (water content in EtOH) on the extraction process were investigated. The effects of particle size of *M. stigma* on the yield of flavonoid were evaluated using six different samples with the mean particle sizes of 0.2, 0.4, 0.6, 0.8, 1.0, and 1.5 mm at the following conditions: pressure of 35 MPa, temperature of 50 °C, CO$_2$ flow rate of 20 L h^{-1}, and 2.0 mL g^{-1} mixture of EtOH/water as co-solvent (with the ratio of 80 : 20 v/v). As could be seen in Figure 18, the extraction time kinetics for the above-mentioned particle size yields the following amounts for flavonoid: 3.81, 3.97, 3.89, 3.78, 3.76, and 3.67 mg g^{-1}. In other words, the yield of flavonoid extracted from *M. stigma* was significantly increased by decreasing the plant particle size from 1.5 to 0.2 mm during an extraction time of 80 min. However, for the extraction times
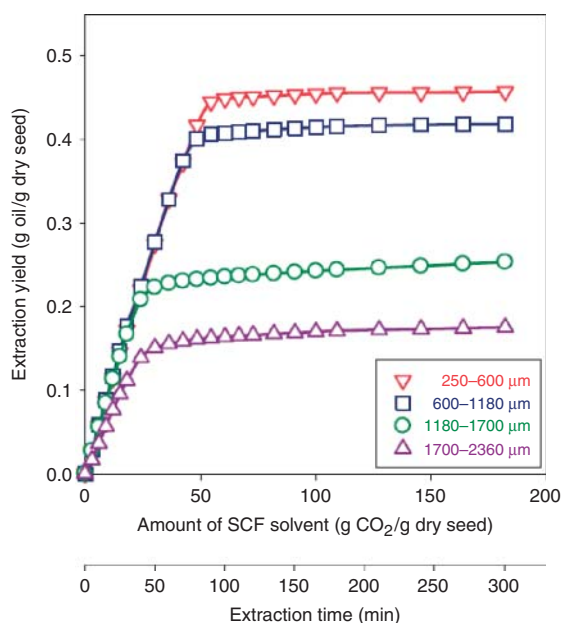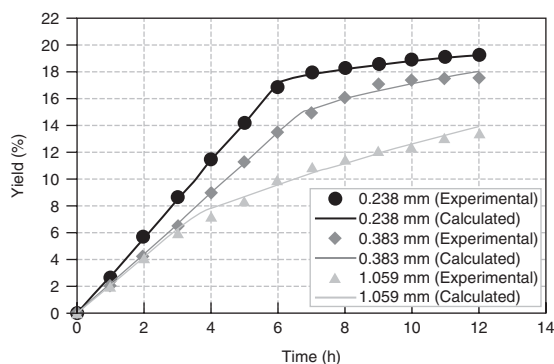

Figure 17 Effect of plant particle size on the extraction yield of pumpkin seed (*Cucurbita pepo* L.) oil obtained via SC-CO$_2$ (30 MPa, 40 °C, 0.12 L h^{-1}). (*Source*: Reprinted from J Supercrit Fluid, 58, Salgn U., Korkmaz H., A green separation process for recovery of healthy oil from pumpkin seed, 239–248, Copyright 2011, with permission from Elsevier.)
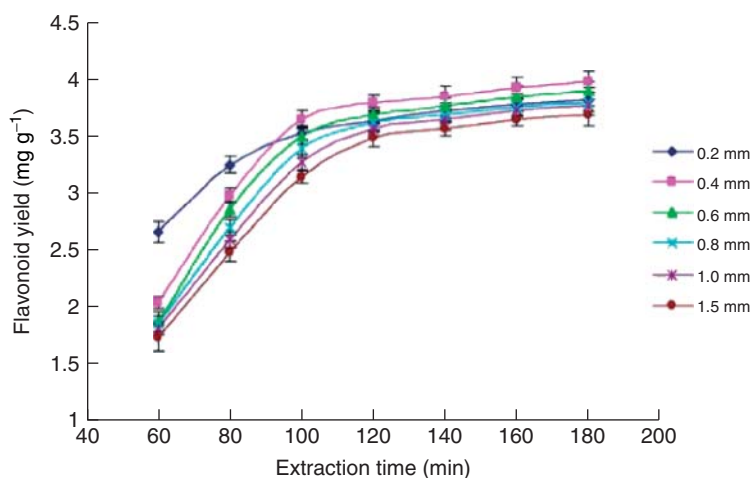
Figure 18 Effect of *Maydis stigma* particle size on flavonoid extraction yields versus time. (*Source*: Reprinted from Food Bioprod Process, 89, Liu J., Lin S., Wang Z., Wang C., Wang E., Zhang Y., Liu J., Supercritical fluid extraction of flavonoids from Maydis stigma and its nitrite-scavenging ability, 333–339, Copyright 2011, with permission from Elsevier.)

higher than 100 min, a slight enhance was observed in the flavonoid yield for all of the studied particle size. These results confirm that the reduction in particle size has a light effect on the total yield of flavonoid extracted from *M. stigma*. These results were in agreement with the SFE of nobiletin and tangeretin from *Citrus depressa* Hayata (Ying-Hung *et al*., 2010). Generally, the reduction in plant particle size leads to more surface area, shortening the diffusion paths in plant solid matrix and improving the SFE yield. However, very small particle sizes make extractions inhomogeneous because of fluid channeling effects in a fixed bed resulting in readsorption of analyte on the matrix surfaces, which disturbs the extraction. In the extraction of flavonoid from *M. stigma*, using the particle size smaller than 0.4 mm causes readsorption of the extracted solutes, which results in lower flavonoid. Hence, 0.4 mm was selected as the optimum particle size for the SFE of the plant.

Jokić *et al*. (2012) investigated a series of operational parameters for the extraction of soybean oil (*Glycine max*) via SFE including sample particle size, pressure, temperature, and CO_2 mass flow rate. At the similar conditions of extraction, the total extraction yield was higher for the plant with smaller particles. The reduction in particle size caused the enhancement of surface area of the sample and hence the amount of oil outside the particles increases. On the other hand, Salgın *et al*. (2006) reported the
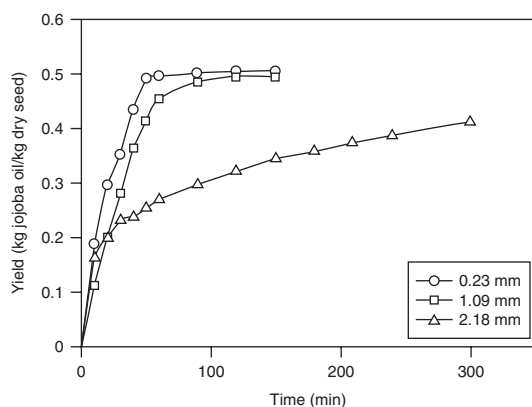

Figure 19 Effect of plant particle size on the extraction yield of soybean oil (*Glycine max*); SFE conditions: $T = 40\,°C$ and $P = 40\,MPa$. (*Source*: Reprinted from J Supercrit Fluid, 38, Reverchon E., De Marco I., Supercritical fluid extraction and fractionation of natural matter, 146–166, Copyright 2006, with permission from Elsevier.)

results of oil extraction from sunflower (*Helianthus annuus*) seeds using SFE at pressure of 40 MPa and temperature of 40 °C with a similar particle size distribution (from 0.23 to 1.09 mm). Their results showed no significant differences in the extraction yield for the plant with different particle sizes of 1.09 and 2.18 mm. However, Reverchon and De Marco (2006) reported that the extraction yield decreases by enhancing the average particle size of plant approximately from 0.25 to 2.0 mm (see Figure 19).

As was discussed, very small particles may cause some problems with channeling inside the extraction bed, which caused the low efficiency and decreased yield. Meanwhile, the sample preparation costs increase because of further milling of plant. On the other hand, as seen in Figure 19, using the larger particles leads to incomplete extraction of oil because of the requirement of a very long time for solvent diffusion inside the plant particles. Thus, the extraction yield rose more slowly by enhancing SCF flow rate for the large particles in comparison with small particles, which indicates that at the higher flow rates, the CO_2 is not saturated with the plant oil (Kriamiti *et al.*, 2002; Özkal *et al.*, 2005; Han *et al.*, 2009; Mezzomo *et al.*, 2009; Döker *et al.*, 2010).

Salgın *et al.* (2004) applied SC-CO_2 in the extraction of jojoba oil from *Simmondsia chinensis* seeds. The effect of plant particle size on the extraction yield was studied and the results are shown in Figure 20. Totally, the yields of extraction enhanced as the extraction time increased; during long extraction periods, the extraction yield could be similar for the plant with various particle sizes. However, during the first few minutes of extraction, particle size effect could be observed very clearly; that is, small particles yield higher oil. As shown in Figure 20, by reducing the particle sizes from about 2.18 to 1.09 mm, a considerable improvement was observed in the extraction yield. On the other hand, more reduction of

particle size to about 1/8 of initial value (0.23 mm) shows no considerable increase in the extraction yield. Such trend for larger particles is due to the effect of intraparticle diffusion, which causes a substantial decrease in the yield of extraction. Figure 20 also indicates that an extraction time of about 100 min is enough to reach the maximum possible yield with the smaller plant particles (0.23 and 1.09 mm) under the above-mentioned extraction conditions.

5.2 Effect of Plant Moisture

In the SC-CO_2 extraction of herbal samples, the presence of water could affect the yield and composition of the extracts. Water in SFE process of plant materials may be present because of the residual moisture of fresh plant or directly be added to the sample before or during the extraction as a co-solvent. The use of water as a natural modifier in a high-pressure extraction processes showed very interested and advantageous results in the most cases. Among various common co-solvents, water and EtOH are attractive modifiers for food-grade extractions. However, the following points should be considered in application of moisture effect: (1) increase of moisture may be lead to the formation of ice blockages because of the Joule–Thompson cooling during the fluid expanding; (2) the ionization of liable compounds by water may cause reduction of their solubility and hence reduction in the rates of extraction (although this phenomenon may be desirable while the compounds are not interested in the final products); (3) the possibility of components hydrolysis; and (4) probable effect on the shelf life of the products (Leeke *et al.*, 2002). However, commonly, the influence of moisture on the oil mass transfer and oil solubility is considered to be a moderate factor that enhances the oil yield in the range between 3% and 18% (Goodrum and Kilgo, 1987; Eggers, 1996). It has been reported that natural moisture content of tobacco affects nicotine SCF extraction and increases the yield to 25% in comparison with dried sample. In addition, moisturizing of the coffee beans accelerates the caffeine removal and increases the SCF extraction yield of vanilla from vanilla beans (Balachandran *et al.*, 2006). Variation of the bed moisture levels may have different effects on mass transfer of solute in the solid phase of plant matrix (Leeke *et al.*, 2002). In SC-CO_2 extraction of *Origanum vulgare* L., the discontinuous addition

Figure 20 Effect of particle size on the SFE yield of jojoba oil from *Simmondsia chinensis* seeds versus extraction time; SFE conditions: 60 MPa pressure, 90 °C temperature, and 2 mL min^{-1} CO_2 flow rate. (*Source*: J Am Oil Chem Soc, 81, 2004, 293–296, Supercritical fluid extraction of jojoba oil, Salgn U., Çalml A., Uysal B.Z. With kind permission from Springer Science and Business Media.)

of water causes an increase in the extraction yield of essential oils, however, reduces the amount of co-extracted waxy materials (Leeke *et al*., 2002). It has been reported that the enhancement of extraction yield is due to the changes of the solvent transport properties and reduction of intraparticle resistance for the solute within the plant particles (Balachandran *et al*., 2006). It was found that increasing the moisture level in oilseeds and paprika samples has a negative effect on the extraction efficiency and required extraction time for reaching the desired yield of extracts (Eggers, 1996; Nagy, 2010), whereas it enhances the loss of volatile components (He *et al*., 2003). Yoda *et al*. (2003) reported influence of 5–10% moisture in the *Stevia rebaudiana* matrix on extraction yield at supercritical and subcritical conditions.

Ivanovic *et al*. (2011) extracted volatile components of *Helichrysum italicum* flowers by SC-CO$_2$ at pressure of 10 MPa and temperature of 40 °C using flowers containing various amounts of moisture (10.5% and 28.4%). The results showed that the higher moisture content of the plant material increases the extraction rate at the initial step of extraction and enhances the extraction yield. The moisture content in the *H. italicum* flowers was about three times larger than the dry flowers. They determined the moisture contents of air-dried and wetted plant samples by Karl Fischer volumetric titration. In order to obtain plant samples with increased moisture content, they sprayed one part of plant particles with distilled water and left in the plastic bag overnight. Obviously, moisturizing of *H. italicum* flowers by water presoaking raises the extraction rate about 40%. This faster extraction leads to diminishing 25% consumption of SC-CO$_2$ for obtaining a similar extraction yield at the identical pressure and temperature. On the basis of the previously published reports (Wang and Weller, 2006; Pourmortazavi and Hajimirsadeghi, 2007), the increase in the extraction yield due to the water content of plant materials might be due to the following factors: (1) dissolving the water into the SC-CO$_2$ phase and enhancing the solubility of solutes through a modifier effect; and (2) the water-swelling of plant matrix could affect the internal or solid mass transfer resistance and transferring of solute components to the surface of plant particle. On the basis of the previous report (Ivanovic *et al*., 2011), an increase in the moisture content of matrix of plant sample (approximately three times rather than dried flowers) results in five times higher solubility of analytes and slight increase of mass transfer resistance in the solid phase.

Devittori *et al*. (2000) extracted crude oil from *Panicum miliaceum* L. with SC-CO$_2$. They investigated the effects of plant matrix moisture on the oil extraction yield. It has been reported that moisture of the plant raw material is one of the parameters that affects SC-CO$_2$ oil extraction (Eggers, 1996). However, drying and reducing the moisture content of ground pellets of millet bran (*Panicum miliaceum*) from 8% to 1–2% could not modify the oil mass transfer during SC-CO$_2$ extraction. This result confirms the previously reported data, which shows that moisture content between 3% and 12% has negligible effect on the mass transfer of oil during SC-CO$_2$ extraction (Snyder *et al*., 1984). Unexpectedly, the oil mass transfer was slightly more efficient in wet pellets (with the moisture content around 8%) in comparison with dried pellets (with 1–2% moisture content). It should be noted that cereals drying, especially millet bran, before oil extraction may cause a bitter flavor in the cake (Dendy, 1995).

Leeke *et al*. (2002) extracted essential oil from the herb *O. vulgare* L. ssp. *virens* (Hoffm. et Link) *letswarrt* using SC-CO$_2$ at 10 MPa and 40 °C in the presence various amounts of water (0–80% w/w). It has been reported (Eggers, 1996) that water could be used as a co-solvent in industrial-scale extraction via SC-CO$_2$. The most important advantage of water over other organic modifiers (i.e., acetone and ethyl acetate) is remaining the extraction process totally clean, which makes the extraction process green. The organic co-solvent would be extracted besides the products, and thus a further separation step will be required for the removal of the solvent residues. Nonpolar essential oil components (i.e., γ-terpinene, *p*-cymene, and α-pinene) possess very low water solubility; thus, using the SC-CO$_2$ modified with water shows minimal increase in the extraction yield. In other words, increasing the polarity of SC-CO$_2$ has no considerable effect on their extraction yield. On the other hand, functional groups of the oxygenated components in the essential oil cause interactions with the water as co-solvent, which lead to the increase of their extraction yield. These interactions are including the formation of charge-transfer complex, dipole (induced dipole), and hydrogen bonding.

Mobin Siddique *et al*. (2011) extracted candlenut (*Aleurites moluccanus*) oil by SC-CO$_2$ and optimized the extraction parameters via RSM. In order to investigate the effect of plant matrix moisture

on the extraction process, various ground candlenut samples including untreated and dried samples in either a heat oven (at 45 °C for 12 h with the sample moisture content of 2.91%) or a vacuum oven (for 5 days with the sample moisture content of 1.98%) were prepared. The untreated plant sample (with the moisture content of 4.87%) was used as a control sample. The results showed that the obtained yield of oil from the samples decreased in the following order: heat-oven-dried plant (77.27%), vacuum-oven-dried sample (74.32%), and the untreated sample (70.12%). The plant dried with the oven heat was containing the highest value of linoleic acid, followed by the untreated plant and vacuum-oven-dried sample. The yield of extraction was lower for the untreated plant and heat-oven-dried sample, which may be due to the aggregation of particles in these samples. The aggregation of plant particles prevents passing of SC-CO$_2$ through the sample matrix. In other words, aggregation of the plant particles caused an increase in the effective size of the sample particles, and hence increased the required path length for the diffusion of SCF to reach the oil. Furthermore, aggregation of the particles and formation of bigger particles caused the absorption of oil in the created wide spaces, which make the oil unable to reach the surface. As a result, the plant particles become sticky in nature and the extraction yield decreases.

6   COMPARISON OF SFE WITH OTHER CONVENTIONAL EXTRACTION TECHNIQUES

SCF could be applied for the extraction and isolation of various components from plant matrices, and the composition of the SC-CO$_2$ extracts may be varied depend on the operational conditions. Usually, to investigate the efficiency of the method for the extraction of herbal volatile components, the SFE results are compared with other conventional techniques, that is, distillation, solvent extraction (SE), Soxhlet, and ultrasound-assisted extraction. Distillation techniques, that is, steam and hydrodistillation (HD) have traditionally been applied for the removal of the essential oil from plant materials. These techniques have some shortcomings, namely losses of volatile compounds, low extraction efficiency, and long extraction time. In addition, elevated temperatures and water could be caused the

degradation or chemical modifications of essential oil components (El-Ghorab *et al*., 2004; Fischer *et al*., 1988, Rubio-Rodríguez *et al*., 2012). As an alternative, SFE is a particularly preferred method for the extraction of natural plant materials. The importance of SFE has been increased in the food and plant extraction during recent years (Herrero *et al*., 2006; Ávila *et al*., 2011; Kim *et al*., 2008). Among various available liquids and gaseous compound, the most popular SCF is CO$_2$, which is an excellent medium for extraction of nonpolar species from plant matrices such as alkenes and terpenes. It is also somewhat suitable for the extraction of moderately polar species, including oxygenated compounds, aldehydes, esters, and alcohols, but is less useful for polar components (Zougagh *et al*., 2004; Castro-Vargas *et al*., 2011; Couto *et al*., 2009). The extraction of polar molecules requires addition of a modifier, most commonly MeOH or EtOH (Punín Crespo and Lage Yusty*et al*., 2005; Ziémons *et al*., 2005; Chikushi *et al*., 2009). Pourmortazavi *et al*. (2004a, 2004b) showed that different extract compositions could be obtained by different extraction methods applied to herbal samples. Their results showed significance difference between the composition of SC-CO$_2$ extracts and the hydrodistilled essential oils obtained from *Juniperus communis* L. leaves. The hydrodistilled oil possesses more volatile components in comparison with the SC-CO$_2$ extracts.

Costa *et al*. (2012) extracted the volatile components from the aerial parts of *Thymus lotocephalus* via HD and SFE techniques. The HD yields about 0.3% (w/w) oil for the aerial parts of *T. lotocephalus*, whereas SFE at pressures of 12 and 18 MPa yields 6.0% and 7.8% (w/w), respectively. The results of this investigation showed that a higher number of compounds might be extracted by HD method; however, the highest extraction yields are obtained by SFE technique. In addition, separation and identification of the extracts indicated different qualitative and quantitative chemical compositions in the HD oil and SFE extracts of the plant. *T. lotocephalus* hydrodistilled oil and its SFE extracts showed suitable antioxidant activity, identified by a hydrogen atom transfer mechanism. Regarding anti-ChE activity, both hydrodistilled oil and SFE extracts showed a considerable activity, mainly against the BChE enzyme. It should be noted that the AChE and BChE inhibitory activities exhibited by hydrodistilled oil were mainly due to the presence of 1,8-cineole and caryophyllene

Table 6 Values of extraction yield (expressed as percentage of fresh plant weight) obtained for six different Tunisian olive leaf (*Olea europaea*) varieties using different extraction processes (SFE, PLE, MAE, and conventional SE).

Extraction process (conditions: solvent, temperature, time, and pressure)	Oueslati (1) yield (%)	Chetoui (2) yield (%)	Chemlali (3) yield (%)	El Hor (4) yield (%)	Jarboui (5) yield (%)	Chemchali (6) yield (%)
SFE (CO_2 + EtOH, 40 °C, 1 h, 15 MPa)	9.5	8.9	5.8	8.2	9.7	5.8
PLE (EtOH, 150 °C, 20 min, 10 MPa)	19.9	19.5	14.8	19.6	16.7	22.4
PLE (water, 150 °C, 20 min, 10 MPa)	10.4	7.5	8.4	8.9	11.2	11.0
MAE (MeOH : water 80 : 20, 80 °C, 6 min)	16.7	10.6	11.2	11.6	5.2	12.1
Conventional (MeOH : water 80 : 20, room temperature, 24 h)	9.4	8.1	8.2	8.2	9.1	16.8

Source: Reprinted from Food Chem Toxicol, 50, Taamalli A., Arráez-Román D., Barrajón-Catalán E., Ruiz-Torres V., Pérez-Sánchez A., Herrero M., Ibaóez E., Micol V., Zarrouk M., Segura-Carretero A., Fernández-Gutiérrez A., Use of advanced techniques for the extraction of phenolic compounds from Tunisian olive leaves: Phenolic composition and cytotoxicity against human breast cancer cells, 1817–1825, Copyright 2012, with permission from Elsevier.

oxide, respectively. Generally, the essential oil obtained by HD method has the most activity.

Taamalli *et al*. (2012) compared different extraction techniques such as microwave-assisted extraction (MAE), SFE, pressurized liquid extraction (PLE), and solid–liquid extraction to investigate their efficiency in the extraction of phenolic components from the leaves of six kinds of Tunisian olive (*Olea europaea*). Among the investigated techniques, PLE yields higher extraction efficiency, whereas phenolic components in the extracts were principally influenced by the solvent type used in the various extraction techniques. The extracted samples by MAE possess a larger number of phenolic compounds, mostly with a polar character. Table 6 presents the yield of extraction for different methods obtained from six kinds of Tunisian olive leaves. As shown in the table, PLE using EtOH as extraction solvent presented the highest yield for all of the studied varieties of plant, whereas among these varieties, Chemchali showed the highest yield. Meanwhile, MAE produces high extraction yields, but considerably less than PLE using EtOH. As an exception, the extraction efficiency for Jarboui variety in MAE method was less in terms of total extracted yield. Finally, PLE method performed with water as solvent, solid–liquid extraction using MeOH : H_2O as solvent and SFE using CO_2 as main solvent and EtOH as co-solvent presented comparable yields.

Huang *et al*. (2011a, 2011b) compared the composition of the volatile compounds obtained from two varieties of *Perilla frutescens* (Baisu and Zisu) via SFE, HS-SPME, and HD. Perillaldehyde and perilla ketone were the main components obtained from Zisu and Baisu, respectively. SFE produced high yield oil and its extracts contained some nonaroma components with the high molecular weights. The optimized conditions of SFE for the extraction of oil from both varieties (Zisu and Baisu) were 30 MPa and 45 °C, which yields about 1.4%. In the SFE extracts of Zisu and Baisu, 35 and 32 compounds were identified, respectively. Meanwhile, some particular compounds with high molecular weights, that is, vitamin E, were extracted via SFE. In addition, HS-SPME as a simple and sensitive technique was used for the analysis of volatile compounds of fresh perilla. HS-SPME was performed at the optimum extraction conditions of 45 °C for 20 min. In the fresh leaves of Zisu and Baisu samples, 34 and 31 components were determined, respectively. On the other hand, the oil yields obtained by HD technique based on the dry plant weight for Zisu and Baisu were 0.51% and 0.38%, respectively. Meanwhile, 38 and 31 compounds were identified by GC/MS of hydrodistilled oil of Zisu and Baisu, respectively.

Nisha *et al*. (2012) studied SFE and organic SE (using hexane) of lipids from freeze-dried *Mortierella alpina* biomass. The composition of extracted lipids via organic solvent method was similar to SC-CO_2 extracts, whereas the pretreatment of the plant cells with HCl and enzyme-treated biomass caused a considerable increase in lipid recovery. The value of recovery increase is proportional to the percentage of ruptured plant cells. In spite of similar composition for the lipids extracted by two methods, SFE and organic SE, a significant difference was observed in the yields of recovered lipids. In
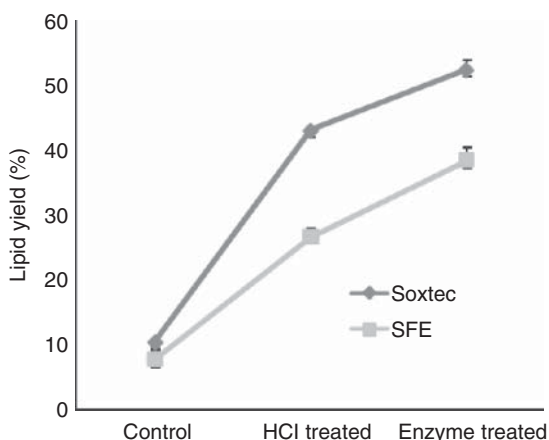
Figure 21 Total yields of extracted lipids by SFE and SE method in the control and treated biomass of *M. alpina*. (*Source*: Reprinted from Food Chem, 133, Nisha A., Udaya Sankar K., Venkateswaran G., Supercritical CO_2 extraction of Mortierella alpina single cell oil: Comparison with organic solvent extraction, 220–226, Copyright 2012, with permission from Elsevier.)

terms of total lipid yield of *M. alpina*, SFE method showed a lower efficiency rather than organic SE (Figure 21). The low yields of SC-CO_2 extractions can be attributed to the various factors affecting the breaking dry cell walls for the extraction of their complex contents including neutral and insoluble lipids. Meanwhile, the low solubility of oils rich in C20 fatty acids in SC-CO_2 may be another reason

for lower extraction yield (Catchpole *et al.*, 2009). On the other hand, the lipid materials together with the triacylglycerols have a suitable solubility in hexane. This high solubility is responsible for the higher extraction yield when hexane was used as the organic solvent in SE. In both SFE and SE methods, the yield of extracted lipid is dependent on the pretreatment of the sample with the biomass.

Parisotto *et al*. (2012) studied the antitumor activity of *Cordia verbenacea* extracts obtained by SFE (with CO_2 at 30 MPa and 50 °C) and conventional organic SE using EtOH. The antitumor *in vitro* assays for both SC-CO_2 and EtOH extracts were carried out and the results indicated that SC-CO_2 extract causes a superior decrease in tumor cells viability and their proliferation, whereas the most probable type of death for these cells is apoptosis. The reason for inducing the cell death by the SFE extracts may be the morphology of the cells stained by acridine orange and ethidium bromide. The results of investigation for the inhibition ability of tumor growth and antitumor activity showed that only SFE extracts are efficient to reduce the expression of cyclooxygenase in the human cancer cell lines MCF-7 (breast). Meanwhile, based on the ratio of unviable cells/viable cells, both SE and SFE extracts of *C. Verbenacea* presented a high potential for the inhibition of tumor
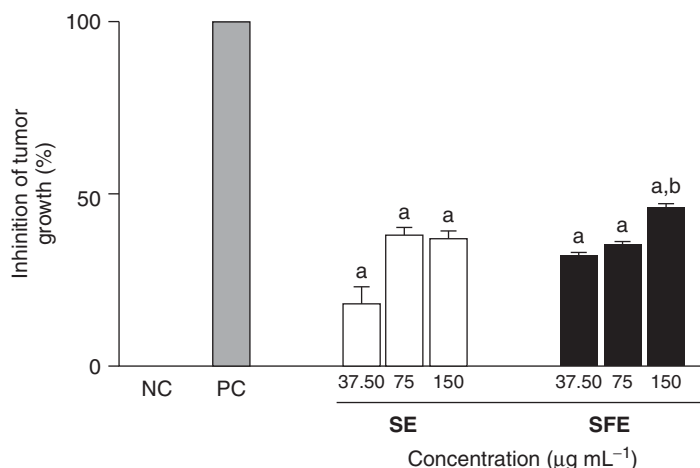
Figure 22 Inhibition of tumor growth in mice inoculated (0.2 mL of 5×10^6 cells/mice) by treatment with SE and SFE extracts of *Cordia verbenacea* (37.5, 75, 150 mg kg^{-1} of mice) during 9 days; (a) marks: the significant differences in relation to control group (NC) and (b) marks: the significant differences between two extracts tested. (*Source*: Reprinted from J Supercriti Fluid, 61, Parisotto E.B., Michielin E.M.Z., Biscaro F., Ferreira S.R.S., Wilhelm Filho D., Pedrosa R.C., The antitumor activity of extracts from Cordia verbenacea D.C. obtained by supercritical fluid extraction, 101–107, Copyright 2012, with permission from Elsevier.)

Table 7 Comparison of SFE method for the extraction and isolation of herbal volatile components with other conventional extraction techniques.

Sample	Analyte	Applied techniques	Optimum conditions of SFE (pressure, temperature, modifier, etc.)	Comparison of results	References
Quince (*Cydonia oblonga* Miller) seeds	Fatty acids	SFE and UAE	35.3 MPa; 35 °C; MeOH 150 µL; static and dynamic extraction times 10 and 60 min	Similar oil composition, but higher extraction yield was achieved with SFE rather than USE	Daneshvand *et al.* (2012)
Olive oil mill wastes (*Olea europa*)	Phenolic compounds	SE and SFE	35 MPa, 40 °C, MeOH, extraction time 60 min	Higher extraction yield was achieved by SE using polar solvent	Lafka *et al.* (2011)
Cordia verbenacea DC	Antioxidant compounds	SFE, SE, HD, and maceration	30 MPa; 50 °C; 8% EtOH; 4 h extraction time	EtAc fraction obtained by maceration and extract isolated by SE using 25% aqueous mixture of EtOH have the highest scavenger activity against DPPH radical; the highest total phenol content value was obtained by SE using EtAc	Michielin *et al.* (2011)
Cordia verbenacea D.C.	Volatile components	SFE and SE	30 MPa and 50 °C	SCF extract causes superior reduction in tumor cells	Parisotto *et al.* (2012)
Wormwood (*Artemisia absinthium* L.)	Volatile components	HD, SE, and SFE	18 MPa; 40 °C and EtOH	SCF extracts exhibited stronger antifeedant effects than the traditional ones (up to eight times more active)	Martín *et al.* (2011)
Drosera intermedia	Plumbagin	SE, UAE, and SFE	20 MPa; 40 °C; 3.62 g min^{-1} SC-CO$_2$ flow rate; extraction time 2 h	Higher yield of plumbagin was achieved by UAE; the highest extraction yield was obtained with SFE	Grevenstuk *et al.* (2012)
Thyme (*Thymus vulgaris*), rosemary (*Rosmarinus officinalis*), and sage (*Salvia officinalis*)	Extracts and essential oil	SFE and HD	35 MPa; 100 °C; 0.3 kg h^{-1} SC-CO$_2$ flow rate	Higher yields in the SFE process	Ivanovic *et al.* (2011)
Spearmint (*Mentha spicata* L.) leaves	Flavonoid compounds	Soxhlet and SFE	20 MPa; 60 °C; extraction time 60 min	Soxhlet extraction yields a higher crude extract; SFE extract contains more main flavonoid compounds	Bimakr *et al.* (2012)
Kenaf seeds (*Hibiscus cannabinus*)	Oil	Sonication, Soxhlet, and SFE	60 MPa; 40 °C	Soxhlet gave the highest oil yield; oil from sonication was the most cytotoxic toward ovarian cancer cell; oil from SFE gives the lowest IC$_{50}$ against colon cancer	Saiful Yazan *et al.* (2011)
Microalgae (*Chlorella vulgaris*)	Lipids	SE and SFE	30.6 MPa; 60 °C; extraction time 6 h	SFE provides higher selectivity than SE for triglyceride extraction though the total extracted lipids	Chen *et al.* (2011)

Table 7 (*Continued*)

Sample	Analyte	Applied techniques	Optimum conditions of SFE (pressure, temperature, modifier, etc.)	Comparison of results	References
Mitragyna speciosa leaves	Alkaloids	UAE, MAE, and SFE	30 MPa; 65 °C; 28.8% EtOH, 12 kg h^{-1} SC-CO$_2$ flow rate; extraction time 45 min	MAE in a closed vessel at 110 °C (60 W, MeOH/water 1 : 1) gives the highest alkaloid fraction amount; UAE with an immersion horn at 25 °C (21.4 kHz, 50 W, MeOH) shows best yield for mitragynine	Orio *et al.* (2012)
Curcuma longa	Curcumin	Soxhlet, MAE, UAE, and SFE	30 MPa; 50 °C; 10% EtOH, SC-CO$_2$ flow rate 5 mL min^{-1}; static time 60 min and dynamic time 300 min	Maximum extraction rate constant was observed by MAE acetone extract of water-soaked curcuma rhizomes; MAE technique is more efficient for curcumin extraction	Wakte *et al.* (2011)
Kenaf (*Hibiscus cannabinus*) seed oil	Lipids (fatty acids, tocopherols, and sterols)	SFE and SE with hexane	60 MPa; 40 °C	No differences between fatty acid compositions of various oil extracts; higher yield of tocopherols was achieved by SFE at high pressure; extraction of kenaf seed oil using SFE at high temperature (80 °C) gives higher amounts of sterols	Adam Mariod *et al.* (2011)
Acorn fruit of *Quercus rotundifolia* L.	Oils	SFE and SE with hexane	18 MPa; 40 °C; superficial velocity of 2.5×10^{-4} ms^{-1}; plant particle diameter of 2.7×10^{-4} m	Higher yield of cholesterol was achieved by SFE; no phospholipids in SC-CO$_2$ extract; identical total amount of tocopherols in acorn oils by extraction via *n*-hexane or SFE	Lopes *et al.* (2005)
Nigella sativa L. seeds	Volatile compounds	SFE and HD	28 MPa/50 °C for total extract; 12 MPa/40 °C for major volatile part	Better recovery of phenolic compounds by SFE	Tiruppur *et al.* (2010)
Flaxseed (*Linum usitatissimum*)	Oil	SFE and SE with petroleum ether	55 MPa; 70 °C; 1 L min^{-1} SC-CO$_2$ flow rate; extraction time 3 h	Higher extraction yield for petroleum ether extraction; higher content of α-linolenic acid in SFE extract	Bozan *et al.* (2002)
Krill (*Euphausia superba*)	Oil containing astaxanthin	SFE and SE with hexane	25 MPa; 45 °C	Higher percentage of polyunsaturated fatty acids in SFE extract; lower acidity and peroxide value of krill oil and more stability than SE oil	Ali-Nehari *et al.* (2012)
Rice bran	Lipids	SFE and Soxhlet with hexane	600 bar; 100 °C	Higher extraction yield with Soxhlet extraction	Kuk *et al.* (1998)

(*continued overleaf*)

Table 7   (*Continued*)

Sample	Analyte	Applied techniques	Optimum conditions of SFE (pressure, temperature, modifier, etc.)	Comparison of results	References
Juniperus communis L. leaves	Essential oils	SFE and HD	20.3 MPa; 45 °C; 30 min dynamic extraction time	Higher concentration for more volatile components in hydrodistilled oil and lower content in SC-CO$_2$ extracts	Pourmortazavi *et al.* (2004a, 2004b)
Thymus lotocephalus	Volatile components	SFE and HD	18 MPa; 40 °C	Higher number of compounds extracted by HD; highest extraction yields obtained by SFE	Costa *et al.* (2012)
Tunisian olive (*Olea europaea*)	Phenolic components	MAE, SFE, PLE, solid–liquid extraction	15 MPa; 40 °C; 6.6% EtOH as modifier	PLE using EtOH showed highest yield; PLE using water, solid–liquid extraction using MeOH : H$_2$O and SFE by CO$_2$ and EtOH co-solvent presented comparable yields	Taamalli *et al.* (2012)
Perilla frutescens (Baisu and Zisu)	Volatile compounds	SFE, HS-SPME, HD	30 MPa; 45 °C	SFE produced higher oil yield and its extracts contained some nonaroma components with high molecular weights	Huang *et al.* (2011a, 2011b)
Mortierella alpina biomass	Lipids	SFE; SE	25 MPa; 50 °C	Composition of extracted lipids via SE was similar to SC-CO$_2$ extracts; total lipid yield of SFE was lower than SE	Nisha *et al.* (2012)
Cordia verbenacea	Extracts	SFE; SE	30 MPa; 50 °C	SC-CO$_2$ extract causes a superior decrease in tumor cells viability and their proliferation	Parisotto *et al.* (2012)
Satureja hortensis L.		SFE; Soxhlet extraction	10% EtOH as co-solvent	Soxhlet extracts with EtOH showed highest yield and antioxidant activity	Plánder *et al.* (2012)

growth in mice in comparison with the negative control (NC) as shown in Figure 22.

Plánder *et al.* (2012) analyzed different extracts of *S. hortensis* L. obtained by various solvents (i.e., SC-CO$_2$ and organic solvent). Various organic solvents including EtOH at different concentrations, acetone, isopropanol, ethyl-acetate, and pentane were used for Soxhlet extraction of summer savory (*S. hortensis* L.). SFE was also carried out with pure CO$_2$ and CO$_2$ modified with EtOH. Among the SFE extracts, the product obtained by CO$_2$ modified with 10% EtOH as co-solvent showed the highest antioxidant activity. In comparison with SC-CO$_2$ extracts, the Soxhlet extracts in the presence of EtOH as solvent showed the highest values of the yield

and antioxidant activity. The results of this work confirm that 50% and 70% aqueous EtOH is the most efficient solvent for the isolation of the antioxidant components from *S. hortensis* L. Meanwhile, among the various organic solvents studied, 50% EtOH presents the highest yield, whereas the extraction yield was decreased with the decreasing the polarity of solvents. On the other hand, the antioxidant activity could be attributed to the presence of other phenolic compounds such as flavonoids in the extracts. The total flavonoid content in the Soxhlet extracts of *S. hortensis* L. decreases as the following order: ethyl-acetate > acetone > EtOH–water (70 : 30 v/v) > isopropanol > pure EtOH > pentane > EtOH–water (96 : 4 v/v) > EtOH–water (50 : 50 v/v),

whereas this order for SC-CO_2 extracts is as follows: CO_2 modified with 15% EtOH > CO_2 + 10% EtOH > CO_2 + 5% EtOH > pure CO_2. Table 7 shows some of the recent reports on the comparison of SFE method efficiency for the extraction of volatile compounds from plant materials with other conventional techniques.

7 CONCLUSION

Application of SCF technology for the isolation and extraction of volatile component from herbal samples has been developed during recent years. The possibility in the tuning of the operational parameters involved in the SFE process, that is, extraction conditions and plant sample properties permit the optimization of the yield and the composition of the extracted products. The SFE method appears to be a green and cost-effective technique at laboratory scale; however, some supplementary information is required for the accurate economic evaluation of its application in large scales. The advantages of SC-CO_2 extraction over other conventional techniques especially HD and Soxhlet extraction are low operating temperature and, hence, prevention of thermal degradation of most of the labile compounds; shorter extraction period; higher selectivity in the extraction of interested compounds; and no organic solvent residue in the extracts, which make the method more interested for plant oil isolation and extraction.

ABBREVIATIONS

AChE	acetylCholinEsterase
BChE	butyrylCholinEsterase
P_c	critical pressure
T_c	critical temperature
DPPH	(2,2-diphenyl-1-picrylhydrazyl)
EtOH	ethanol
EtAc	ethyl acetate
GC-MS	gas chromatography-mass spectroscopy
HD	hydrodistillation
HS-SMPE	head space-solid phase microextraction
MeOH	methanol
MAE	microwave-assisted extraction
NC	negative control
PLE	pressurized liquid extraction
SC-CO_2	supercritical carbon dioxide
RSM	response surface methodology
SCFs	supercritical fluids
SD	standard deviation
SE	solvent extraction
SFE	supercritical fluid extraction
UAE	ultrasound-assisted extraction

REFERENCES

Ab Rahman, N. N., Al-Rawi, S. S., Ibrahim, A. H., *et al.* (2012) *J. Food Eng.*, **108**, 166–170.

Adam Mariod, A. , Matthaus, B., and Ismail, M. (2011) *J Am Oil Chem Soc*, **88**, 931–935.

Aghel, N., Yamini, Y., Hadjiakhoondi, A., *et al.* (2004) *Talanta*, **62**, 407–411.

Ahmed, Z., Abdeslam-Hassan, M., Ouassila, L., *et al.* (2012) *Energy Procedia*, **18**, 1038–1046.

Akay, S., Alpak, I. and Yesil-Celiktas, O. (2011) *J. Sep. Sci.*, **34**, 1925–1931.

Akgun, I. H., Erkucuk, A., Pilavtepe, M., *et al.* (2011) *J. Supercrit. Fluid*, **57**, 31–37.

Ali-Nehari, A., Kim, S-B., Lee, Y-B., *et al.* (2012) *Korean J. Chem Eng*, **29**, 329–336.

Andrade, K. S., Gonçalvez, R. T., Maraschin, M., *et al.* (2012) *Talanta*, **88**, 544–552.

András, C. D., Simándi, B., Örsi, F., *et al.* (2005) *J. Sci. Food Agric.*, **85**, 1415–1419.

Arnaáiz, E., Bernal, J., Teresa Martín, M., *et al.* (2011) *Eur. J. Lipid Sci. Technol.*, **113**, 479–486.

dos Santos, A., Garcia, V., Ferreira Cabral, V., *et al.* (2012) *J. Supercrit. Fluid*, **69**, 75–81.

Ávila, M., Zougagh, M., Escarpa, A., *et al.* (2011) *J. Supercrit. Fluid*, **55**, 977–982.

Balachandran, S., Kentish, S. E. and Mawson, R. (2006) *Sep. Purif. Technol.*, **48**, 94–105.

Bayat, Y., Pourmortazavi, S. M., Iravani, H., *et al.* (2012) *J. Supercrit. Fluid*, **72**, 248–254.

Bimakr, M., Abdul Rahman, R., Ganjloo, A., *et al.* (2012) *Food Bioprocess Technol.*, **5**, 912–920.

Boutin, O. and Badens, E. (2009) *J. Food Eng.*, **92**, 396–402.

Bozan, B. and Temelli, F. (2002) *J Am Oil Chem Soc*, **79**, 231–235.

Brunner, G. (2005) *J. Food Eng.*, **67**, 21–33.

Canela, A. P. R. F., Rosa, P. T. V., Marques, M. O. M., *et al.* (2002) *Ind. Eng. Chem. Res.*, **41**, 3012–3018.

Casas, L., Mantell, C., Rodriguez, M., *et al.* (2007) *J. Supercrit. Fluid*, **41**, 43–49.

Cassel, E., Frizzo, C. D., Vanderlinde, R., *et al.* (2000) *Ind. Eng. Chem. Res.*, **39**, 4803–4805.

Castro-Vargas, H. I., Rodríguez-Varela, L. I. and Parada-Alfonso, F. (2011) *J. Supercrit. Fluid*, **56**, 238–242.

Catchpole, O. J., Tallon, S. J., Eltringham, W. E., *et al.* (2009) *J. Supercrit. Fluid*, **47**, 591–597.

Cavalcanti, R. N., Veggi, P. C. and Meireles, M. A. A. (2011) *Procedia Food Sci.*, **1**, 1672–1678.

Chang, C. J. and Randolph, A. D. (1989) *AIChE J.*, **35**, 1876–1881.

Chassagnez-Méndez, A. L., Machado, N. T., Araujo, M. E., *et al.* (2000) *Ind. Eng. Chem. Res.*, **39**, 4729–4733.

Chen, Q., Li, P., He, J., Zhang, Z., Liu, J. (2008) *J. Sep. Sci.*, **31**, 3218–3224.

Chen, Q., Li, P., Yang, H., *et al.* (2009) *J. Sep. Sci.*, **32**, 3152–3156.

Chen, T., Zhao, X., Wu, J., *et al.* (2011) *J. Taiwan Inst. Chem. E*, **42**, 428–434.

Chikushi, H., Hirota, K., Yoshida, N., *et al.* (2009) *Talanta*, **80**, 738–743.

Chiu, K. L., Cheng, Y. C., Chen, J. H., *et al.* (2002) *J. Supercrit. Fluid*, **24**, 77–87.

Chulalaksananukul, W., Condoret, J. S. and Combes, D. (1993) *Enzyme Microb. Technol.*, **15**, 691–698.

Cocero, M. J., González, S., Pérez, S., *et al.* (2000) *J. Supercrit. Fluid*, **19**, 39–44.

Comin, L. M., Temelli, F. and Aranda, S. M. (2011) *J. Am. Oil Chem. Soc.*, **88**, 707–715.

Cooper, A. I. (2000) *J. Mater. Chem.*, **10**, 207–234.

Cossuta, D., Simándi, B., Hohmann, J., *et al.* (2007) *J. Sci. Food Agric.*, **87**, 2472–2481.

Costa, P., Goncalves, S., Grosso, C., *et al.* (2012) *Ind. Crop Prod.*, **36**, 246–256.

Couto, R. M., Fernandes, J., Gomes da Silva, M. D. R., *et al.* (2009) *J. Supercrit. Fluid*, **51**, 159–166.

Dai, J., Ha, C. and Shen, M. (2008) *J. Sep. Sci.*, **31**, 714–720.

Dandekar, D. V. and Gaikar, V. G. (2002) *Sep. Sci. Technol.*, **37**, 2669–2690.

Daneshvand, B., Ara, K. M. and Raofie, F. (2012) *J Chromatogr A*, **1252**, 1–7.

Darr, J. A. and Poliakoff, M. (1999) *Chem. Rev.*, **99**, 495–541.

Daukšas, E., Venskutonis, P. R., Povilaityte, V., *et al.* (2001) *Nahrung*, **45**, 338–341.

de Lourdes, L., Moraes, M., Vilegas, J. H. Y., *et al.* (1997) *Phytochem. Anal.*, **8**, 257–260.

Dendy, D. A. V. (1995) Sorghum and the Millets: Production and Importance, in *Sorghum and Millets. Chemistry and Technology*, ed. D. A. V. Dendy, American Association of Cereal Chemists, Inc., St. Paul, pp. 11–26.

del Valle, J. M. and Uquiche, E. L. (2002) *J. Am. Oil Chem. Soc.*, **79**, 1261–1266.

del Valle, J. M. and Urrego, F. A. (2012) *J. Supercrit. Fluid*, **66**, 157–175.

de Oliveira, P. F., Francisco Machado, R. A., Bolzan, A., *et al.* (2012) *J. Supercrit. Fluid*, **63**, 161–168.

Devittori, C., Gumy, D., Kusy, A., *et al.* (2000) *J. Am. Oil Chem. Soc.*, **77**, 573–579.

Döker, O., Salgin, U., Yildiz, N., *et al.* (2010) *J. Food Eng.*, **97**, 360–366.

Ebrahimzadeh, H., Yamini, Y., Sefidkon, F., *et al.* (2003) *Food Chem.*, **83**, 357–361.

Eggers, R. (1996) Supercritical fluid extraction (SFE) of oilseeds/lipids in natural products, in *Supercritical Fluid Technology in Oil and Lipid Chemistry*, eds. J. W. King and G. R. List, AOCS Press, Champaign, IL, pp. 35–64.

El-Ghorab, A. H., Mansour, A. F. and El-massry, K. F. (2004) *Flavour Fragr. J.*, **19**, 54–61.

Erkey, C. (2000) *J. Supercrit. Fluid*, **17**, 259–287.

Fischer, N., Nitz, S. and Drawert, F. (1988) *J. Agric. Food Chem.*, **36**, 996–1003.

Förg, A., Leupold, G., Parlar, H., *et al.* (2002) *Eng. Life Sci.*, **2**, 79–82.

Felföldi-Gáva, A., Szarka, S., Simándi, B., *et al.* (2012) *J. Supercrit. Fluid*, **61**, 55–61.

Gaspar, F. (2002) *Ind. Eng. Chem. Res*, **41**, 2497–2503.

Gholivand, M. B., Rahimi-Nasrabadi, M., Batooli, H., *et al.* (2010) *Food Chem. Toxicol.*, **48**, 24–28.

Gholivand, M. B., Rahimi-Nasrabadi, M., Batooli, H., *et al.* (2012) *Nat. Prod. Res.*, **26**, 883–891.

Gholivand, M. B., Rahimi-Nasrabadi, M., Mehraban, E., *et al.* (2011) *Nat. Prod. Res.*, **25**, 1585–1595.

Giannuzzo, A. N., Boggetti, H., Nazareno, M. A., *et al.* (2003) *Phytochem. Anal.*, **14**, 221–223.

Glisic, S. B., Ristic, M. and Skala, D. U. (2011) *Ultrason. Sonochem.*, **18**, 318–326.

Goodrum, J. W. and Kilgo, M. B. (1987) *Trans. ASAE*, **30**, 1865–1868.

Gopalan, B., Goto, M., Kodama, A., *et al.* (2000) *J. Agric. Food Chem.*, **48**, 2189–2192.

Gouveia, A. F., Duarte, C., Beirão da Costa, M. L., *et al.* (2006) *Eur. J. Lipid Sci. Technol.*, **108**, 421–428.

Gracia, I., Rodríguez, J. F., de Lucas, A., *et al.* (2011) *J. Supercrit. Fluid*, **59**, 72–77.

Grevenstuk, T., Goncalves, S., Nogueira, J. M. F., *et al.* (2012) *Ind. Crop Prod.*, **35**, 257–260.

Guclu-Ustundag, O. and Temelli, F. (2005) *J. Supercrit. Fluid*, **36**, 1–15.

Guoliang, L., Junyou, S., Yourui, S., *et al.* (2011) *LWT – Food Sci. Technol.*, **44**, 1172–1178.

Guo-qing, H., Hao-ping, X., Qi-he, C., *et al.* (2005) *J. Zhejiang Univ. Sci.*, **6B**(10), 999–1004.

Hamburger, M., Baumann, D. and Adler, S. (2004) *Phytochem. Anal.*, **15**, 46–54.

Han, X., Cheng, L., Zhang, R., *et al.* (2009) *J. Food Eng.*, **92**, 370–376.

Hatami, T., Glisic, S. B. and Orlovic, A. M. (2012) *J. Supercrit. Fluid*, **62**, 102–108.

Hatami, T., Rahimi, M., Veggi, P. C., *et al.* (2011) *J. Supercrit. Fluid*, **55**, 929–936.

Herrero, M., Cifuentes, A. and Ibañez, E. (2006) *Food Chem.*, **98**, 136–148.

He, H. P., Corke, H. and Cai, J. G. (2003) *J. Agric. Food Chem.*, **51**, 7921–7925.

Huang, Z., Xu, L. and Li, J. (2011a) *Chem. Eng. J.*, **166**, 461–467.

Hsieh, C., Chang, C. J. and Ko, W. (2005) *Fisheries Sci.*, **71**, 441–447.

Huang, L., Zhong, T., Chen, T., *et al.* (2007) *Rapid Commun. Mass Spectrom.*, **21**, 3024–3032.

Huang, B., Lei, Y., Tang, Y., *et al.* (2011b) *Food Chem.*, **125**, 268–275.

Huang, Z., X-h, S. and W-j, J. (2012) *J. Chromatogr. A*, **1250**, 2–26.

Husum, T. L., Christensen, T. W., Christensen, M. W. and Andersen, L. E. (2001) Application for new immobilized *Candida antarctica* lipase B. Information Sheet, Oil and Fats, Novozymes A/S, Bagsvaerd, Denmark.

Ivanovic, J., Ristic, M. and Skala, D. (2011) *J. Supercrit. Fluid*, **57**, 129–136.

Jiping Sun, Y. J. and Aiyong, Q. (2002) *J. Cereals Oils*, **5**, 2–4(in Chinese with English abstract).

Jokić, S., Nagy, B., Zeković, Z., *et al.* (2012) *Food Bioprod. Proc.*, **90**, 693–699.

Ju, H., Huang, K., Chen, J., *et al.* (2010) *J. Am. Oil Chem. Soc.*, **87**, 1063–1070.

Kagliwal, L. D., Patil, S. C., Pol, A. S., *et al.* (2011) *Sep. Purif. Technol.*, **80**, 533–540.

Kaškonienė, V., Kaškonas, P., Maruška, A., *et al.* (2011) *Acta Physiol. Plant*, **33**, 2377–2385.

Kerrola, K., Galambosi, B. and Kallio, H. (1994) *J. Agric. Food Chem.*, **42**, 1979–1988.

Khajeh, M. (2011) *J. Supercrit. Fluid*, **55**, 944–948.

Kim, W., Kim, J., Kim, J., *et al.* (2008) *J. Food Eng.*, **89**, 303–309.

Kriamiti, H. K., Rascol, E., Marty, A., *et al.* (2002) *Chem. Eng. Proc.*, **41**, 711–718.

Kuk, M. S., Dowd, M. K., (1998) *J Am Oil Chem Soc*, **75**, 623–628.

Kumhom, T., Elkamel, A., Douglas, P. L., *et al.* (2011) *Chem. Eng. J.*, **172**, 1023–1032.

Lafka, T-I., Lazou, A. E., Sinanoglou, V. J., *et al.* (2011) *Food Chem*, **125**, 92–98.

Lee, Y., Charles, A. L., Kung, H., *et al.* (2010) *Ind. Crop Prod.*, **31**, 59–64.

Lee, S. M., Lee, H. S., Kim, K. H., *et al.* (2009) *J. Food Sci.*, **74**, S135–S141.

Leeke, G., Gaspar, F. and Santos, R. (2002) *Ind. Eng. Chem. Res.*, **41**, 2033–2039.

Lehotay, S. J. (1997) *J. Chromatogr. A*, **785**, 289–312.

Li, L., Lee, W., Jong Lee, W., *et al.* (2010a) *Food Sci. Biotechnol.*, **19**, 405–410.

Li, H., Wu, J., Rempel, C. B., *et al.* (2010b) *J. Am. Oil Chem. Soc.*, **87**, 1081–1089.

Lina, S., Fei, R., Xudong, Z., *et al.* (2010) *J. Am. Oil Chem. Soc.*, **87**, 1221–1226.

Liu, J., Lin, S., Wang, Z., *et al.* (2011) *Food Bioprod. Proc.*, **89**, 333–339.

Liu, G., Xu, X., Gong, Y., *et al.* (2012) *Food Bioprod. Proc.*, **90**, 573–578.

Lopes, I. M. G., Bernardo-Gil, M. G. (2005) *Eur J Lipid Sci Technol*, **107**, 12–19.

Lucien, F. P. and Foster, N. R. (2000) *J. Supercrit. Fluid*, **17**, 111–134.

Luque de Castro, M. D. and Garcia-Ayuso, L. E. (1998) *Anal. Chim. Acta*, **369**, 1–10.

Machmudah, S., Kawahito, Y., Sasaki, M., *et al.* (2007) *J. Supercrit. Fluid*, **41**, 421–428.

Machmudah, S., Zakaria, W. S., *et al.* (2012) *J. Food Eng.*, **108**, 290–296.

Manpong, P., Douglas, S., Douglas, P. L., *et al.* (2011a) *J. Food Process Eng.*, **34**, 1661–1681.

Manpong, P., Douglas, S., Douglas, P. L., *et al.* (2011b) *J. Food Process Eng.*, **34**, 1408–1418.

Mariod, A. A., Abdelwahab, S. I., Gedi, M. A., *et al.* (2010) *J. Am. Oil Chem. Soc.*, **87**, 849–856.

Martín, L., González-Coloma, A., Díaz, C. E., *et al.* (2011) *J. Supercrit. Fluid*, **57**, 120–128.

Mattea, F., Martín, Á. and Cocero, M. J. (2009) *J. Food Eng.*, **93**, 255–265.

Mezzomo, N., Martínez, J. and Ferreira, S. R. S. (2009) *J. Supercrit. Fluid*, **51**, 10–16.

Michielin, E. M. Z., de Lemos Wiese, L. P., Ferreira, E. A., *et al.* (2011) *J. Supercrit Fluid*, **56**, 89–96.

Mobin Siddique, B., Ahmad, A., Alkarkhi, A. F. M., *et al.* (2011) *J. Food Sci.*, **76**, C535–C542.

Mukhopadhyay, M. (2000) *Natural Extracts using Supercritical Carbon dioxide*, CRC Press LLC, N.W. Corporate Blvd., Boca Raton, Florida 33431.

Nagy, B. (2010) Modelling of supercritical fluid extraction. Ph.D. Thesis, Budapest University of Technology and Economics Faculty of Chemical and Bioengineering György Oláh Doctoral School, http://www.omikk.bme.hu/collections/phd/Vegyeszmernoki es Biomernoki Kar/2010/.

Nguyen, H. N., Gaspillo, P. D., Maridable, J. B., *et al.* (2011) *Chem. Eng. Process*, **50**, 1207–1213.

Nisha, A., Udaya Sankar, K. and Venkateswaran, G. (2012) *Food Chem.*, **133**, 220–226.

Nyam, K. L., Tan, C. P., Lai, O. M., *et al.* (2011) *Food Bioprocess Technol.*, **4**, 1432–1441.

Oliveira, E. L. G., Silvestre, A. J. D. and Silva, C. M. (2011) *Chem. Eng. Res. Des.*, **89**, 1104–1117.

Orio, L., Alexandru, L., Cravotto, G., *et al.* (2012) *Ultrason Sonochem*, **19**, 591–595.

Overmeyer, A., Schrader-Lippelt, S., Kasche, V., *et al.* (1999) *Biotechnol. Lett.*, **21**, 65–69.

Özkal, S. G., Yener, M. E. and Bayındırlı, L. (2005) *J. Supercrit. Fluid*, **35**, 119–127.

Palavra, A. M. F., Coelho, J. P., Barroso, J. G., *et al.* (2011) *J. Supercrit. Fluid*, **60**, 21–27.

Park, H. S., Im, N. G. and Kim, K. H. (2012) *LWT – Food Sci. Technol.*, **45**, 73–78.

Pan, J., Wang, H., Chen, C., *et al.* (2012) *Eng. Life Sci.*, **12**, 1–10.

Parisotto, E. B., Michielin, E. M. Z., Biscaro, F., *et al.* (2012) *J. Supercrit. Fluid*, **61**, 101–107.

Pereira, C. G. and Meireles, M. A. A. (2010) *Food Bioprocess Technol.*, **3**, 340–372.

Peterson, A., Machmudah, S., Roy, B. C., *et al.* (2006) *J. Chem. Technol. Biotechnol.*, **81**, 167–172.

Plánder, S., Gontaru, L., Blazics, B., *et al.* (2012) *Eur. J. Lipid Sci. Technol.*, **114**, 772–779.

Pourmortazavi, S. M., Baghaee, P. and Mirhosseini, M. (2004a) *Flavour Fragr. J.*, **5**, 417–420.

Pourmortazavi, S. M. and Ghadiri, M. (2005) *J. Food Comp. Anal.*, **18**, 439–446.

Pourmortazavi, S. M. and Hajimirsadeghi, S. S. (2005) *Ind. Eng. Chem. Res.*, **44**, 6523–6533.

Pourmortazavi, S. M. and Hajimirsadeghi, S. S. (2007) *J. Chromatogr. A*, **1163**, 2–24.

Pourmortazavi, S. M., Hajimirsadeghi, S. S., Kohsari, I., *et al.* (2004b) *J. Chem. Eng. Data*, **49**, 1530–1534.

Pourmortazavi, S. M., Rahimi-Nasrabadi, M. and Hajimirsadeghi, S. S. (2014) *Curr. Anal. Chem.*, **10**, 3–28.

Pourmortazavi, S. M., Sefidkon, F. and Hosseini, S. G. (2003) *J. Agric. Food Chem.*, **51**, 5414–5418.

Punín Crespo, M. O. and Lage Yusty, M. A. (2005) *Chemosphere*, **59**, 1407–1413.

Rahimi-Nasrabadi, M., Ahmadi, F. and Batooli, H. (2012a) *Nat. Prod. Res.*, **26**, 669–674.

Rahimi-Nasrabadi, M., Ahmadi, F. and Batooli, H. (2012b) *Nat. Prod. Res.*, **26**, 637–642.

Rahimi-Nasrabadi, M., Gholivand, M. B. and Batooli, H. (2009) *Dig. J. Nanomater. Bios.*, **4**, 819–822.

Rahimi-Nasrabadi, M., Gholivand, M. B., Vatanara, A., *et al.* (2012c) *J. Herb. Spices Med. Plants*, **18**, 318–330.

Rahimi-Nasrabadi, M., Nazarian, S., Farahani, H., *et al.* (2013a) *Int. J. Food Prop.*, **16**, 369–381.

Rahimi-Nasrabadi, M., Pourmortazavi, S. M., Nazarian, S., *et al.* (2013b) *Int. J. Food Prop.*, DOI: 10.1080/10942912. 2011.558227.

Reverchon, E. and De Marco, I. (2006) *J. Supercrit. Fluid*, **38**, 146–166.

Romero, N., Fernández, A. and Robert, P. (2012) *Eur J. Lipid Sci. Technol.*, **114**, 951–957.

Rostagno, M. A., Araújo, J. M. A. and Sandi, D. (2002) *Food Chem.*, **78**, 111–117.

Rubio-Rodríguez, N., de Diego, S. M., Beltrán, S., *et al.* (2012) *J. Food Eng.*, **109**, 238–248.

Sahena , F., Zaidul, I. S. M., Jinap, S., *et al.* (2009) *J. Food Eng.*, **95**, 240–253.

Saiful Yazan, L., Foo, J. B., Ghafar, S. A. A., *et al.* (2011) *Food Bioprod Process*, **18**, 328–332.

Saldaña, M. D. A., Sun, L., Guigard, S. E., *et al.* (2006) *J. Supercrit. Fluid*, **37**, 342–349.

Salgin, U. (2007) *J. Supercrit. Fluid*, **39**, 330–337.

Salgin, U., Çalımlı, A. and Uysal, B. Z. (2004) *J. Am. Oil Chem. Soc.*, **81**, 293–296.

Salgin, U., Döker, O. and Çalımlı, A. (2006) *J. Supercrit. Fluid*, **38**, 326–331.

Salgin, U. and Korkmaz, H. (2011) *J. Supercrit. Fluid*, **58**, 239–248.

Salgin, S. and Salgin, U. (2006) *Eur. J. Lipid Sci. Technol.*, **108**, 577–582.

Salgin, U., Üstün, A. S., Mehmetoğlu, Ü., *et al.* (2005) *J. Chem. Technol. Biotechnol.*, **80**, 124–132.

Sargenti, S. R. and Lanças, F. M. (1998) *J. Microcolumn Sep.*, **10**, 213–223.

Sauceau, M., Letourneau, J. J., Freiss, B., *et al.* (2004) *J. Supercrit. Fluid*, **31**, 133–140.

Senyay-Oncel, D., Ertas, H. and Yesil-Celiktas, O. (2011) *J. Am. Oil Chem. Soc.*, **88**, 1061–1069.

Snyder, J. M., Friedrich, J. P. and Christianson, D. D. (1984) *J. Am. Oil Chem. Soc.*, **61**, 1851–1856.

Sovova, H. (2012) *J. Supercrit. Fluids*, **66**, 73–79.

Taamalli, A., Arráez-Román, D., Barrajón-Catalán, E., *et al.* (2012) *Food Chem. Toxicol.*, **50**, 1817–1825.

Tang, G. W., Yang, C. J. and Xie, L. D. (2007) *J. Pest. Sci.*, **80**, 151–157.

Tiruppur Venkatachallam, S. K., Pattekhan, H., Divakar, S., *et al.* (2010) *J. Food Sci. Technol.*, **47**, 598–605.

Tonthubthimthong, P., Chuaprasert, S., Douglas, P., *et al.* (2001) *J. Food Eng.*, **47**, 289–293.

Topal, U., Sasaki, M., Goto, M., *et al.* (2006) *J. Agric. Food Chem.*, **54**, 5604–5610.

Torres, C. F., Fornari, T., Torrelo, G., *et al.* (2009) *Eur. J. Lipid Sci. Technol.*, **111**, 459–463.

Turner, C., King, J. W. and Mathiasson, L. (2001) *J. Agric. Food Chem.*, **49**, 553–558.

Turner, C. and McKeon, T. (2002) *J. Am. Oil Chem. Soc.*, **79**, 473–478.

Van Opstaele, F., Goiris, K., De Rouck, G., *et al.* (2012) *J. Supercrit. Fluid*, **69**, 45–56.

Veggi, P. C., Cavalcanti, R. N. and Meireles, M. A. A. (2011) *Procedia Food Sci.*, **1**, 1717–1724.

Visentín, A., Cismondi, M. and Maestri, D. (2011) *Innov. Food Sci. Emerg.*, **12**, 142–145.

Wakte, P. S., Sachin, B. S., Patil, A. A., *et al.* (2011) *Sep. Purif. Technol.*, **79**, 50–55.

Wang, L. and Weller, C. L. (2006) *Trends Food Sci. Technol.*, **17**, 300–312.

Wang, L. Z., Yang, B., Dua, X., *et al.* (2008) *Food Chem.*, **108**, 737–741.

Wenli, Y., Bo, S. and Yaping, Z. (2005) *J. Sci. Food Agric.*, **85**, 489–492.

Xia, L., You, J., Li, G., *et al.* (2011) *J. Am. Oil Chem. Soc.*, **88**, 23–32.

Yamini, Y., Sefidkon, F. and Pourmortazavi, S. M. (2002) *Flavour Fragr. J.*, **17**, 345–348.

Ye, C. and Lai, Y. (2012) *Chem. Eng. Technol.*, **35**, 646–652.

Ying-Hung, L., Albert, L. C., Hsien-Feng, K., *et al.* (2010) *Ind. Crops Prod.*, **31**, 59–64.

Yoda, S. K., Marques, M. O. M., Petenate, A. J., *et al.* (2003) *J. Food Eng.*, **57**, 125–134.

Zarena, A. S., Manohar, B. and Udaya Sankar, K. (2012) *Food Bioproc. Technol.*, **5**, 1181–1188.

Zarena, A. S. and Udaya Sankar, K. (2011) *Sep. Purif. Technol.*, **80**, 172–178.

Zhi-ling, C., Jian-ping, C., Hui-lin, C., *et al.* (2011) *Procedia Environ. Sci.*, **8**, 426–432.

Ziémons, E., Goffin, E., Lejeune, R., *et al* (2005) *J. Supercrit. Fluid.*, **33**, 53–59.

Zougagh, M., Rios, A. and Valcarcel, M. (2004) *Anal. Chim. Acta*, **524**, 279–285.

Zougagh, M., Valcárcel, M., and Ríos, A. (2004) Supercritical fluid extraction: a critical review of its analytical usefulness. *TrAC Trend Anal. Chem.* **23**: 399–405.

Zuknik, M. H., Nik Norulaini, N. A. and Mohd Omar, A. K. (2012) *J. Food Eng.*, **112**, 253–262.

4

New Trends in Extraction of Natural Products: Microwave-Assisted Extraction and Pressurized Liquid Extraction

**Philippe Christen and Beatrice Kaufmann**

School of Pharmaceutical Sciences, University of Geneva, University of Lausanne, Geneva, Switzerland

1 INTRODUCTION

Natural products, especially those originating from plants, are still an important source of chemotherapeutic agents. Because plants may contain thousands of primary and secondary metabolites with very different physicochemical properties, the extraction process is very often tedious and time consuming. The first goal of a plant extraction process is to separate the desired compounds from the solid matrix. Active constituents from plant materials are often still extracted using conventional solid–liquid extraction (SLE) techniques. Such techniques involve an adequate choice of solvents and use of a shaking and/or heating mechanism in order to increase the mass transfer and solubilization of the active compounds. The traditional extraction methods for determining a wide variety of natural compounds remain maceration, ultrasonication, percolation, turbo-extraction (high speed mixing), heating under reflux, steam distillation, and Soxhlet extraction. The last of these is frequently used for plant extraction. These techniques are, however, often considered to be a bottleneck for most analytical procedures, being one of the least evolved parts of the whole process, as illustrated in Figure 1 (Rouessac and Rouessac, 2004). They are both labor intensive and time-consuming. In addition to a prolonged extraction time (typically 16–24 h for Soxhlet), they suffer from several limitations including high solvent consumption, low selectivity, and an often unsatisfactory reproducibility. Furthermore, large volumes of polluting organic solvents have to be used, most of which are hazardous and/or expensive, which is not compatible with clean and "green" chemistry. In addition, some of these methods (steam distillation and Soxhlet) require an extended heating period, which can create artifacts or even degrade heat-sensitive compounds.

With the advent of fast analytical methods (fast- and ultra-fast GC, UHPLC, and capillary electrophoresis) that allow extracts to be analyzed in a few minutes, or even seconds (Bieri *et al.*, 2006), there is a need for fast and efficient extraction of samples in a way that is both cost-effective and environmentally friendly. Modern extraction technologies should involve low solvent consumption, be as exhaustive as possible, be rapid and easy-to-use, and offer the possibility of automation. They should result in a significant reduction of solvent consumption and thus decrease the waste generated during the process. Furthermore, they allow the use of solvents that are widely acceptable for food and
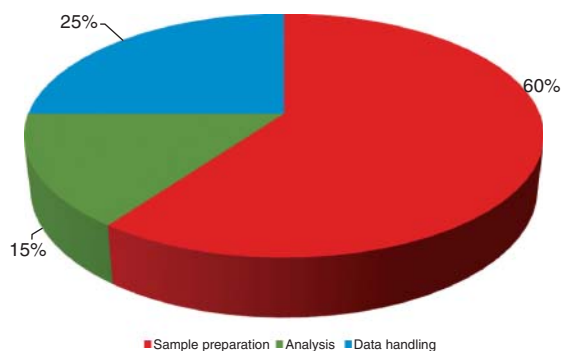
Figure 1 Graphical representation of the global analytical process (Rouessac and Rouessac, 2004).

pharmaceutical processing. Among these technologies are those based on compressed fluids, such as supercritical fluid extraction, microwave-assisted extraction (MAE), pressurized liquid extraction (PLE), and pressurized hot water extraction (PHWE).

During the last decade, some exhaustive review articles have been published concerning the use of modern techniques for natural product isolation (Huie, 2002; Romanik *et al.*, 2007; Sticher, 2008). This chapter focuses on two of the alternative approaches to traditional extraction methods, namely MAE and PLE. PHWE can be considered as a special case of PLE in which the solvent is water and is briefly considered in this chapter.

The influence of important parameters (e.g., temperature, pressure, nature of the solvent, solvent flow rate, when relevant, and extraction time) on the extraction efficiency and characteristics of the extracted products is discussed, based on an overview of the recent literature related to plant materials.

2   MICROWAVE-ASSISTED EXTRACTION

Although many initially reported applications of MAE dealt with recovery of pollutants from environmental matrices, the interest in this process for the extraction of active constituents from plant materials increased dramatically over the last decade, as shown by the large number of papers published on this topic. A few review articles have also been released concerning the application of MAE to natural product isolation (Ananth *et al.*, 2009; Chan *et al.*, 2011; Mandal *et al.*, 2007; Nuechter *et al.*, 2005).

2.1   Principle of the Method and Heating Mechanism

MAE is a process that uses microwave energy to heat a solid sample in order to desorb analytes from the matrix (Chen *et al.*, 2008a).

Microwaves are electromagnetic radiation with a frequency ranging from 300 MHz to 300 GHz (Figure 2). In order to avoid interference with official communications, domestic, industrial, and scientific microwave devices operate at 2.45 GHz.

Directly related to their electromagnetic nature, microwaves possess two oscillating perpendicular magnetic and electric fields. The latter is responsible for the transformation of electromagnetic energy into calorific energy via two simultaneous mechanisms occurring in the sample, namely, dipole rotation and ionic conduction. Dipole rotation occurs as a result of the alignment within the electric field of those molecules possessing a dipole moment (either permanent or induced by the electric field), in both the solvent and the solid sample. This fast oscillation generates heat through collisions and friction between surrounding molecules. When working at the usual frequency of 2.45 GHz, this phenomenon occurs 4.9×10^9 times per second and the resultant heating is very fast. Indeed, the larger the dielectric constant of the solvent, the more optimal is the heating. If the medium is more viscous, then the molecular rotation is affected and this phenomenon is diminished. Consequently, unlike classical conductive heating methods, microwaves heat the whole sample simultaneously, thus leading to very fast and efficient extractions. Another advantage of MAE is the disruption of weak hydrogen bounds promoted by the dipole rotation of the molecules. Furthermore, the migration of the dissolved ions increases solvent
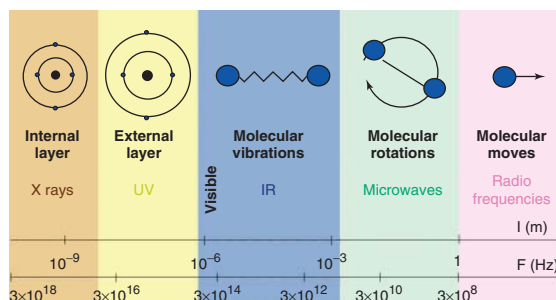
Figure 2   The electromagnetic spectrum.

penetration into the matrix and thus facilitates the solvation of the analytes.

Changing the electric field also induces ionic currents in the solution. As the medium resists the electrophoretic migration of dissolved ions, friction occurs and heat is liberated by a Joule effect. This phenomenon depends on the size and charge of the ions present in the solution. Roughly, the higher the dielectric constant of a compound, the greater the resistance to the microwave energy that is observed in the sample (Belanger *et al.*, 2008). Solvents having a permanent dipole moment (e.g., methanol or water), as well as ionic liquids, strongly absorb microwave energy, whereas apolar solvents such as hexane are transparent to microwave irradiation. The ability of the solvent to dissipate the heat, also called dielectric loss (ε''), is in fact of utmost importance (Jain *et al.*, 2009). Indeed, if the difference between the dielectric constant (ε') and the dielectric loss is large, microwaves will be very strongly absorbed. Water, for instance, has a ε' of 80 and a ε'' of 12. This difference implies a greater ability to absorb microwave energy than to dissipate the heat in the system. This leads to the "superheating" phenomenon, which occurs only in matrices containing water. Methanol does not present this "superheating" behavior, because it has a ε' of 23.9 and a ε'' of 15.2 and therefore, a better ability to dissipate the heat generated by absorbing the microwave energy.

This liquid phase extraction approach is based on the unique ability of microwaves to selectively heat part of the sample, which in most cases comprises substances that are present directly in the matrix itself (Pare and Belanger, 2010).

Figure 3 shows the three basic situations that can be occur in MAE: (a) absorbing sample in an absorbing medium (bulk heating), (b) absorbing sample in a transparent solvent, and (c) a rehydrated sample dipped in either a transparent (C1) or an absorbing (C2) solvent.

One particular application is the extraction of essential oils, for which a cold apolar solvent can be used (e.g., hexane; Ferhat *et al.*, 2007). Indeed, microwaves interact selectively with the polar molecules present in glands, trichomes, or vascular tissues (selective heating and extraction). Localized heating leads to the expansion and rupture of cell walls and is followed by the liberation of essential oils into the solvent (Figure 3b).

This situation presents obvious advantages in the case of thermosensitive substances.

A similar case is obtained by moisturizing a dry sample with the appropriate volume of polar solvent before performing the extraction (Figure 3c).

2.2 Instrumentation

A microwave-assisted extractor is basically made up of a magnetron (i.e., microwave generator), a wave guide, and an extraction chamber (Pare and Belanger, 2010).

Despite the increasing use of domestic ovens (modified or not) in laboratories, two technologies

(a) Absorbing sample in absorbing solvent (bulk heating mechanism)

(b) Absorbing sample in transparent solvent

(c) Moisturizing of the sample and extraction using transparent (C1) or absorbing (C2) solvent
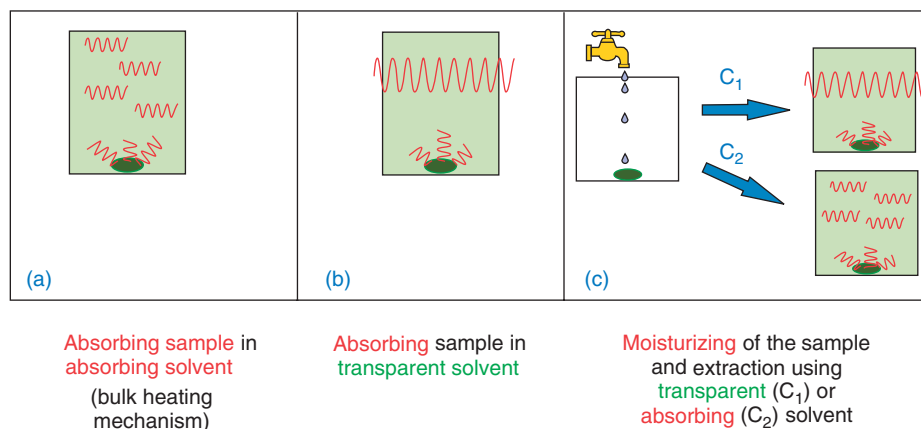
Figure 3 Three basic extraction schemes of solid matrices by MAE. (a) Absorbing sample in an absorbing solvent. (b) Absorbing sample in a transparent solvent. (c) Rehydrated sample dipped in either a transparent (C1) or an absorbing solvent (C2).

are commercially available: closed vessel systems working under controlled pressure and temperature, and open extraction systems operating at atmospheric pressure. Owing to hazards related to highly flammable organic solvents and the potential lack of homogeneity of the microwave field in domestic ovens, it is strongly recommended to use only equipments approved for MAE.

2.2.1   Closed Vessel Systems

In closed MAE, several samples can be extracted at once using the same set of experimental conditions. Figure 4 depicts the MARS 6 system (CEM Corp.) that allows up to 40 simultaneous extractions. The solid samples are placed into the cells with adequate solvent and subjected to microwave irradiation, causing instantaneous efficient heating in the presence of microwave-absorbing substances (Camel, 2002). The cells are made of nonmicrowave-absorbing material that is able to withstand high temperatures (up to 300 °C) and high pressures (as high as 13 bar). Furthermore, the part of the cell in contact with the sample must be totally inert. High quality glass, quartz, or polymer transparent to microwaves, such as PTFE, are generally used (Majors, 2006). The closed vessels are placed on a turntable in order to overcome possible nonhomogeneity of the electric field in the cavity. These devices are generally equipped with a high level of security features (Majors, 2006), such as high capacity exhaust fans, solvent sensors, or pressure-burst safety membranes placed on each vessel.

Pressure is generated in the closed vessels by heating. In some cases, the inner temperature can grow up to circa 100 °C above the atmospheric boiling point of the solvent used. Temperature can also be set at a fixed value by adjusting the microwave power applied. The selected power has to be set at a sufficient level to avoid solute degradation due to excessive temperature and overpressure problems.

Both speed and efficiency of extraction are enhanced in this procedure. It is notable that extraction vessels must be cooled down before opening, for reasons of pressure and temperature. This is particularly important in the case of volatile compounds that partition between the headspace and the liquid phase when the samples are heated, thus leading to potential losses and quantification errors. After irradiation,

Figure 4   Typical closed vessel MAE system (Mars 6) from CEM Corp. (*Source*: Reproduced by permission of CEM Corporation.)

the solid residue needs to be removed by filtration or centrifugation, and the liquid phase is collected.

2.2.2   Open Cells

The open cell system operates at atmospheric pressure. Thus, the maximum temperature that can be achieved is determined by the boiling point of the solvent or mixture of solvents used. The samples are immersed into the solvent and placed in quartz microwave transparent containers, which are topped by a vapor condenser. The microwaves are often focused on the sample, thus allowing homogeneous and very efficient heating. The sample to be extracted may be directly dipped into the solvent or can be placed into a Soxhlet-type cellulose cartridge to avoid filtration steps. This approach can be particularly interesting for recovery of thermosensitive compounds.

Compared to closed vessel extractions, open cells offer increased safety in sample handling and, furthermore, allow larger samples to be extracted. However, the commercial supply of such systems (e.g., Star 2 from CEM Corp.; Figure 5) is limited and this technique presents a lower degree of automation.

In both closed and open systems, the sample can be concentrated by evaporation of the solvent. The dry

Figure 5 Typical open vessel MAE system (Star 2) from CEM Corp. (*Source*: Reproduced by permission of CEM Corporation.)

residue is dissolved into an appropriate vehicle before chromatographic analysis.

2.2.3 Microwave-Assisted Hydrodistillation (MAHD)

MAHD is a particular procedure in which the water contained in the sample itself (either native or added to the plant material) heats under microwave irradiation, thus causing rupture of some of the specific structures of the matrix itself (Deng *et al.*, 2007). The essential oils or volatiles are released and condensed by a vapor condenser and separated from the water present in the collecting vial as a two-phase system, similar to conventional steam distillation (Iriti *et al.*, 2006; Lo Presti *et al.*, 2005; Riela *et al.*, 2011; Tigrine-Kordjani *et al.*, 2011; Yahaya *et al.*, 2010). It has been demonstrated that essential oil obtained by this technique contains a higher concentration of oxygenated compounds in comparison to the same oil obtained by traditional hydrodistillation (Iriti *et al.*, 2006).

2.2.4 Microwave-Assisted Extraction-Headspace-Solid-Phase Microextraction (MAE-Headspace-SPME)

Solid-phase microextraction (SPME) is a relatively recent sampling technique, introduced by Arthur and Pawliszyn in the early 1990s (Arthur and Pawliszyn, 1990). This technique uses small diameter fibers coated with a stationary phase on which the analytes extracted from either liquid or solid samples are adsorbed. This procedure is particularly suitable for volatiles, and several techniques can be applied to enhance the partition of analytes between the solid or liquid phase and the vapor phase, including microwave heating. The microwave-assisted extracted volatiles are trapped directly on an SPME fiber placed above the microwave heated plant sample (Fan *et al.*, 2009; Yang *et al.*, 2010b).

Combination of both microwave irradiation and SPME online adsorption allows the selectivity of extraction to be increased and was successfully applied to the analysis of volatiles from *Magnolia grandiflora* L. (Fan *et al.*, 2009) or *Polygala furcata* Royle (Yang *et al.*, 2010b). All these techniques take advantage of the fact that water molecules are selectively heated by the microwaves and that the energy thus released can be transferred to surrounding molecules that are relatively transparent to microwaves (Belanger *et al.*, 2008). Instead of water (Yang *et al.*, 2010b), it is also possible to add inert microwave-absorbing material (e.g., carbonyl iron powder (CIP), graphite powder, or activated carbon powder (Wang *et al.*, 2006a, 2006b) to the sample. These substances act only as microwave heating catalysts (Wang *et al.*, 2006c). Using this method, MAE is even possible in transparent organic solvents (Majors, 2006).

2.3 Influencing Parameters in MAE

The factors affecting MAE are strongly dependent on the nature of both the solvent and the solid matrix. Basically, three factors are important for the extraction of a compound from a plant matrix: the solubility of the analyte in the selected solvent, the mass

transfer kinetic of the analyte from the matrix to the solvent, and the strength of the matrix/analyte interactions (Jain *et al.*, 2009).

2.3.1 Nature of the Solvent

There are two basic approaches to using microwave heating for sample preparation: bulk and selective heating (Pare and Belanger, 2010). While the former does not particularly take advantage of the specific properties of microwaves, mimicking classical thermal resistive extraction, the latter allows the selective heating of part of the system (the matrix or the surrounding solvent), thus exploiting the unique selectivity of extraction provided by the application of microwaves. This selectivity of heating is directly related to the interaction of microwaves with polar compounds or solvents.

Solvents generally used for this purpose cover a wide range of polarities, from heptane to water. In most cases, the selected solvent has a high dielectric constant and strongly absorbs microwave energy (bulk heating). However, the extracting selectivity and the ability of the medium to interact with

microwaves can be modulated using mixtures of solvents, such as alcohols and water, which are often used for natural product extractions. The fastest extractions have been reported to take place in a mixture of solvents of different polarities, typically methanol and water or hexane and ethanol (Chen and Spiro, 1995; Spiro and Chen, 1995). Figure 6 shows a GC–MS chromatogram of steroidal saponins extracted by focused open MAE (Soxwave 3.6, Prolabo, Fontenay-sous-Bois, France) from the seeds of *Trigonella foenum-graecum* L. using a mixture of 2-propanol (*i*PrOH) and water. The solid sample (200 mg) was placed inside a quartz tube topped with a vapor condenser. Ten milliliters of extracting solvent (42% *i*PrOH in water) was added and the extraction was performed over 30 min at 40 W.

Some particular applications of MAE using ionic liquids (e.g., 1-butyl-3-methylimidazolium bromide) for the extraction of natural compounds are reported for phenols (Du *et al.*, 2009), anthraquinones (Lu *et al.*, 2011), flavonoids (Zeng *et al.*, 2010), alkaloids (Du *et al.*, 2010; Lu *et al.*, 2008; Ma *et al.*, 2010), lignans (Yuan *et al.*, 2011), and phthalides (Chi *et al.*, 2011).
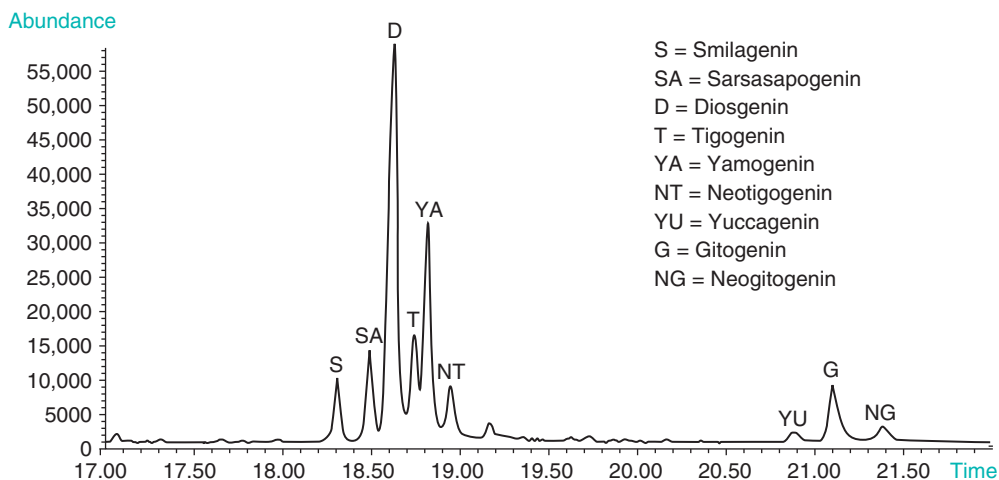
Figure 6 GC–MS chromatogram of *Trigonella foenum-graecum* seeds extract obtained by MAE (SIM on fragment 139). Experimental conditions: all extractions were performed using a focused microwave system operating at atmospheric pressure (Soxwave 3.6, Prolabo, Fontenay-sous-Bois, France). The solid sample (200-mg seeds) was placed into a quartz tube topped with a vapor condenser. Ten milliliters of solvent (42% isopropanol in water) were added. The extraction was performed during 30 min at a power of 40 W. GC–MS was performed with a HP 5890 series II gas chromatograph coupled with a HP 5972 (Agilent Technologies, Palo Alto, CA, USA) mass spectrometer. The mass selective detector was operated in the electron impact ionization mode with an ionization potential of 70 eV. Helium was used as carried gas at a constant flow rate of $1 \, mL \, min^{-1}$. Injection was performed in the splitless mode (time: 1.5 min) and injection volume was 1 μL. The gas chromatograph was fitted with a $30 \, m \times 0.25 \, mm$ ID fused silica capillary column coated with a phenyl-methyl silicone phase HP-5MS with a film thickness of 0.25 μm.

Attempts have also been made to use nonionic surfactants for micellar extraction of triterpenic substances (Sun *et al.*, 2007a). To improve interactions between the matrix and the microwaves, CIP was mixed to the plant material before extraction with a transparent solvent (anhydrous diethylether) in order to extract volatile compounds from ginger (Yu *et al.*, 2007): this produced an excellent extraction selectivity. A similar approach was applied using hexane for recovery of monoterpenes (Chemat *et al.*, 2005), ethylacetate for triterpenes (Asghari *et al.*, 2011), and petroleum ether for alkaloids (Raman and Gaikar, 2002). In all these cases, water contained in the matrix heats locally and promotes release of secondary metabolites into the cold surrounding medium.

The ability of substances to interact at various degrees with microwaves enables fine modulation of the extraction selectivity by varying the nature of the solvent used. Figure 7 shows the significant increase in the selectivity of extraction of artemisinin, an antimalarial sesquiterpene extracted from *Artemisia annua* L. The extraction rate was in fact identical for all solvents tested (Figure 7a), whereas the cleanliness of the extracts increased with decreasing dielectric constant of the solvents used (Figure 7b). Hexane produced the best results in terms of both extraction yield and selectivity. Residual water still present in the plant material causes localized superheating, resulting in explosion of plant structure and in release of the active substance(s) into the cold surrounding medium.

A particular technique was applied for the extraction of polyphenols using a mixture of water and diethylether (Lu *et al.*, 2007): the upper organic phase remains cold, whereas the aqueous solvent and the plant matrix are heated. This procedure allows MAE and purification of the extracts to be simultaneously carried out by liquid–liquid extraction. Similarly, lipophilic compounds were extracted using a mixture of water and petroleum ether (Shen *et al.*, 2008).
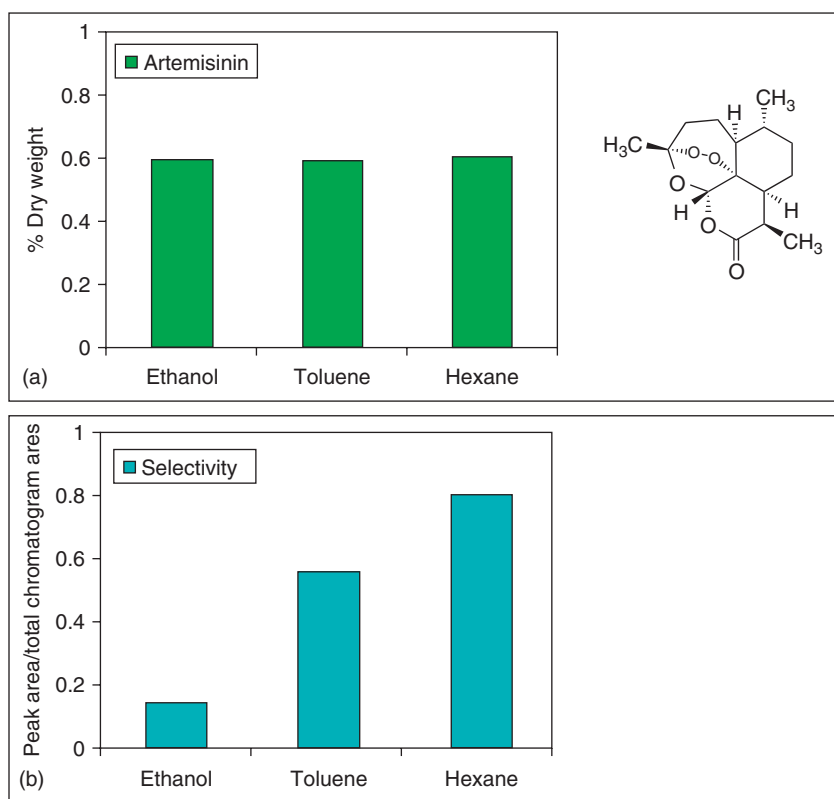
Figure 7 MAE of artemisinin from the leaves of *Artemisia annua* with hexane, toluene, and ethanol for 40 s. (a) The extraction rate was similar for the three solvents but selectivity increased significantly with decreasing dielectric constant (b).

2.3.2 Solvent Volume

The volume of organic solvent used for MAE can be dramatically reduced to only a few milliliters (typically <40 mL for plant metabolites) (Romanik *et al.*, 2007) in comparison with conventional SLE techniques. Generally, the amount of solvents may be about 10 times less than those used in classical SLEs. In the particular case of microwave-assisted hydrodistillation, no organic solvent is used. Unlike conventional extraction methods, for which the extraction efficiency increases simultaneously with increasing solvent volume, in MAE, a higher liquid-to-matrix ratio may give lower recoveries (Xiao *et al.*, 2008). This could partially be explained by the poor microwave heating of the matrix if the large volume of solvent absorbs most of the microwave energy (Chan *et al.*, 2011). Inversely, a low liquid-to-matrix ratio may result in solvent with an insufficient solubilization capacity.

Once a suitable solvent or mixture of solvents has been selected, its quantity has to be determined as this affects the extraction efficiency and therefore the extraction yield (Chan *et al.*, 2011). Du *et al.* (2010b) shows that the solvent/matrix ratio has a significant influence on the extraction yield of flavonoids from *Pueraria lobata* (Willd.) Ohwi and *Pueraria thomsonii* Berth.

2.3.3 Influence of the Matrix Characteristics

In many cases, the matrix moisture improves extraction recovery. The effect of this parameter also depends on the extraction solvent used in the method. The water added (or naturally occurring in the sample) will always have an effect on the microwave-absorbing ability and hence facilitate the heating process. Thus, precise control of the sample's moisture leads to more reproducible results. In addition to the moisture content, a reduced particle size enables a larger surface area in contact with the solvent and thus a better extraction yield (Kaufmann *et al.*, 2001a). However, too small particle size may cause difficulty in separating the extract from the plant residue without an additional clean-up step (Chan *et al.*, 2011). Furthermore, one must take into account the plant organ (seed, leaf, root, bark, etc.) considered when optimizing MAE conditions. For leaves, a short MAE exposure with approx. 20–50% water is generally suitable, whereas for roots and rhizomes, more extreme conditions, that is, longer extraction time, high water proportion, and high power setting are required. Matrix porosity directly influences the diffusion of analytes to the surface of the solid sample. Higher temperature and swelling of the matrix may therefore promote faster extraction kinetics by increasing the diffusion rate (Jain *et al.*, 2009). The targeted heating of water or polar molecules present in the matrix leads to local superheating, causing significant expansion with subsequent rupture of cell walls and liberation of the analytes into the surrounding medium. Thus, optimized parameters depend mainly on the matrix rather than on the compounds to be extracted. Kaufmann *et al.* (2007) report the influence of the plant matrix in MAE of diosgenin extracted from seeds, leaves, and roots from fenugreek. Chemat *et al.* (2005) observe that microwave irradiation of caraway seeds seems to affect the internal structure of the cells. The higher terpene extraction yield was directly related to ruptures observed in the glandular tissues after explosion promoted by intracellular superheating.

2.3.4 Extraction Time

To avoid thermal degradation of compounds, the extraction time should be minimized. In the case of natural products from plants, the irradiation periods vary from a few seconds (Jain *et al.*, 2009) to 30 min (Deng *et al.*, 2007; Huie, 2002; Sticher, 2008) with most of the applications reporting an average extraction time of 10 min. The extraction of artemisinin, a heat-sensitive sesquiterpene, with three solvents of different polarities (ethanol, toluene, and hexane) for 40 s showed that the extraction yield was identical in all three cases (Figure 7a). This confirms that a short irradiation time, even in the presence of a solvent with high dielectric constant, avoids thermal degradation of compounds. Overexposure may lead to degradation of fragile compounds, as well as oxidation risks. If a longer extraction time is needed, performing repeated shorter irradiation cycles using fresh solvent may avoid thermal degradation.

2.3.5 Microwave Power and Temperature

The effect of microwave energy is strongly dependent on the dielectric constant of both the solvent and

the plant matrix. Microwave power and temperature are two parameters that are closely related. In open systems, applying a higher microwave power allows a faster temperature rise. Therefore, the extraction kinetic can be accelerated (Chemat *et al.*, 2005). This is related to a drop in viscosity and surface tension (Chan *et al.*, 2011). At ambient pressure, the maximal temperature that can be obtained is the boiling point of the solvent or mixture of solvents used. As long as the applied power is sufficient to reach this temperature and to desorb the metabolites, no additional improvement is observed when applying higher power settings.

The power must be selected carefully to avoid excessive temperature, particularly when working with a closed vessel system. Even if a temperature above the boiling point of the solvent can improve extraction efficiency because of increased desorption from the matrix, this could lead to compound degradation and overpressure in the vessel. Some commercially available laboratory equipment enables application of finely tuned microwave power, whereas domestic ovens work at various preset power settings. For closed systems, typical power settings are between 400 and 1000 W, whereas open focused systems operate at 25–250 W.

2.4 Application of MAE to Natural Products Recovery

MAE of natural products covers a broad range of compounds, from nonpolar to very polar substances. Some interesting reviews have been published on this topic during recent years (Chan *et al.*, 2011; Huie, 2002; Jain *et al.*, 2009; Mandal *et al.*, 2007; Sticher, 2008).

In the case of natural product recovery, the most widely used solvent is ethanol, at a proportion ranging from 40% to 100%. Water is also often used, and some particular applications take advantage of the localized superheating phenomenon occurring in the core of the matrix in presence of water, whether already present or added by a presoaking step or sample pretreatment (Yang *et al.*, 2010b). In this case, water pretreatment could be responsible for the extraction of desired compounds. The moisture level of the matrix may in fact directly influence the optimal operating conditions needed (Gao *et al.*, 2004).

Depending on the active compounds to be recovered, extraction time ranges between 30–40 s and 30 min for most of the volatile oils, and even 60 min (Yahaya *et al.*, 2010) or more than 70 min for polysaccharides extracted with pure water (Wang *et al.*, 2010b).

Some applications require several cycles using fresh solvent. This is the case for curcumin extracted from *Curcuma longa* L. (Mandal *et al.*, 2008) or the flavonoid scutellarin from *Erigeron breviscapus* Hand.-Mazz. (Gao *et al.*, 2007). For the extraction of coumarins from *Angelica dahurica* Fish ex. Hoffm. (Xie *et al.*, 2010b), a first extraction was performed with water for 1 min, followed by extraction with 30% ethanol for 1 min and finally for 10 min using ethanol 90%. In a similar approach, lignans were extracted from *Magnolia biondii* Pamp. with water during two cycles of 5 min each (Zhao *et al.*, 2007).

Optimization of extraction parameters can be performed applying chemometric models, response surface methodology, or orthogonal optimization (Hu *et al.*, 2008; Japon-Lujan *et al.*, 2006; Jyothi *et al.*, 2010; Kaufmann *et al.*, 2007; Kim *et al.*, 2007; Kong *et al.*, 2010; Lee *et al.*, 2006; Liao *et al.*, 2008; Mandal and Mandal, 2008; Mirzajani *et al.*, 2010; Pai *et al.*, 2011; Sun *et al.*, 2007b; Teng *et al.*, 2009; Xie *et al.*, 2010a; Yan *et al.*, 2009, 2010a; Yoshida *et al.*, 2010; Zhong *et al.*, 2010).

Table 1 shows a selection of MAE applications dedicated to natural products that have been published during the last decade. The extraction conditions are summarized with respect to the main influencing parameters involved.

3 PRESSURIZED LIQUID EXTRACTION

PLE, also known as pressurized solvent extraction (PSE) or pressurized fluid extraction (PFE), was introduced in the mid-1990s under the trade name: accelerated solvent extraction – ASE™, developed by Dionex. PLE was primarily used for extraction of environmental samples such as soils and sediments. It was reported for the first time by Richter *et al.* (1996) for the extraction of polycyclic aromatic hydrocarbons, polychlorinated biphenyls, and petroleum hydrocarbons. PLE is now a well-established extraction method in many laboratories. This is confirmed by the increasing number of papers appearing on the topic (Haglund and Spinnel, 2011). In 1996, the US Environment Protection Agency adopted PLE for

Table 1   Selected applications of MAE to the extraction of natural products.

Class of compounds	Plants	MAE conditions						References
		System	Solvent	Power (W)	Time (min.)	Temp. (°C)	Pressure (bar)	
Phenolic compounds								
Bergenin	*Ardisia crenata* Sims. *Rodgersia sambucifolia* Hemsl.	c	60% MeOH	n.i.	15	60	n.i.	Deng *et al*. (2010)
Tanshinones	*Salvia miltiorrhiza* Bge	o	Water	800	12–16	70	Atm.	Fang *et al*. (2009)
Phenolic acid	*Lonicera japonica* Thunb.	▲	70% EtOH	ca. 400	4	n.i.	Atm.	Hu *et al*. (2009)
Cinnamic acid + harpagoside	*Scrophularia ningpoensis* Hemsl.	▲	75% EtOH	ca. 400	4	n.i.	Atm.	Huang *et al*. (2011)
Oleuropein	*Olea europaea* L.	o	80% EtOH	200	8	n.i.	Atm	Japon-Lujan *et al*. (2006)
Curcumin	*Curcuma longa* L.	o	Acetone	Various	1 (repeat)	n.i.	Atm.	Mandal *et al*. (2008)
Tyrosol	*Rhodiola crenulata* L.	▲ (o)	50% MeOH	400	5	n.i.	Atm.	Mao *et al*. (2007)
Phenolic acid	*Larrea tridentata* (DC.) Coville	c	MeOH	800	1	70 ± 2	n.i.	Martins *et al*. (2010)
Phenolic compounds	Various aromatic plants	▲	60% MeOH, 60% acetone	750	4	n.i.	n.i.	Proestos and Komaitis (2008)
Phenolic compounds (acids and aldehydes)	Various plants	c	2 mol L^{-1} HCl + 0–30% MeOH	500	20	70	n.i.	Sterbova *et al*. (2004)
Phenolic acids	*Sambucus nigra* L. *Polygonum aviculare* L.	o/c	80% MeOH	240–360	30	n.i.	n.i.	Waksmundzka-Hajnos *et al*. (2007)
Phenolic acids	*Ligustrum lucidum* Ait.	c	80% EtOH	500	30	70	n.i.	Xia *et al*. (2011a)
Chlorogenic acid	*Lonicera japonica* Thunb.	▲	50% EtOH	700	5	60	Atm.	Zhang *et al*. (2008)
Gallic acid	*Camellia oleifera* C. Abel	o	Water	n.i.	35	76	Atm.	Zhang *et al*. (2011)
Phenolic compounds	*Phaseolus vulgaris* L.	c	50% EtOH	n.i.	15	50–150	n.i.	Sutivisedsak *et al*. (2010)
Phenolic compounds	*Geranium sibiricum* L.	c	Water pH 5.2	500	9	33	n.i.	Yang *et al*. (2010a)
Polyphenols	*Uncaria sinensis* (Oliv.) Havil	c	Water	600	20	100	n.i.	Tan *et al*. (2011)
Anthraquinones	*Morinda citrifolia* L.	c	80% EtOH	720	30	60	n.i.	Hemwimon *et al*. (2007)
Terpenes								
Monoterpenes (*C10*)								
Carvone and limonene	*Carum carvi* L.	o	Hexane	120	60	n.i.	Atm.	Chemat *et al*. (2005)
Camphor and borneol	*Chrysanthemum indicum* L.	▲	Water	700	4	n.i.	n.i.	Deng *et al*. (2006a)
Paeonol	*Cynanchum paniculatum* (Bge) Kitag. *Paeonia suffruticosa* Andrews	▲	Water	700	4	n.i.	n.i.	Deng *et al*. (2006b)

Table 1 (*Continued*)

Class of compounds	Plants	MAE conditions						References
		System	Solvent	Power (W)	Time (min.)	Temp. (°C)	Pressure (bar)	
Sesquiterpenes (C15)								
Curcumol, curdione, and germacrone	*Curcuma wenyujin* Y.H. Chen & C. Ling, *C. phaeocaulis* Valeton, *C. kwangsiensis* S.G. Lee & C.L. Liang	▲	Water	700	4	n.i.	n.i.	Deng *et al.* (2006c)
Artemisinin	*Artemisia annua* L.	o	n.i.	n.i.	12	n.i.	Atm.	Hao *et al.* (2002)
Diterpenes (C20)								
Andrographolide and dehydrographolide	*Andrographis paniculata* Nees.	c	60% MeOH	80	6	n.i.	n.i.	Chen *et al.* (2007c)
Tanshinones	*Salvia miltiorrhiza* Bge	▲	95% EtOH	n.i.	2	80	n.i.	Pan *et al.* (2002)
Stevioside Rebaudioside	*Stevia rebaudiana* Bertoni	c	Water	n.i.	5–45	60–120	n.i.	Teo *et al.* (2009)
Triterpenes (C30)								
Ginsenosides	*Panax quinquefolium* L.	o	Aq. MeOH	90	5–60	n.i.	Atm.	Dai and Orsat (2010)
Ginsenosides	*Panax quinquefolium* L.	c	70% EtOH	n.i.	10	125	65	Qu *et al.* (2009), Shi *et al.* (2007)
Ginsenosides	*Panax ginseng* C.A. Meyer	c	70% EtOH	n.i.	15	n.i.	72.5	Shi *et al.* (2007)
Ginsenosides	*Panax ginseng* C.A. Meyer	▲	70% EtOH	50	1–10	n.i.	Atm.	Kim *et al.* (2007)
Withanolides	*Withania somnifera* Dunal	o	MeOH	140	2	50	Atm.	Jyothi *et al.* (2010)
Withanolides	*Withania somnifera* Dunal	c	75% MeOH	n.i.	150 sec	68	n.i.	Mirzajani *et al.* (2010)
Withanolides	*Iochroma gesnerioides* (Kunth) Miers	o	MeOH	120	40 s	n.i.	Atm.	Kaufmann *et al.* (2002), Kaufmann *et al.* (2007)
Oleanolic acid	*Gymnema sylvestre* R.Br.	o	EtOH	500	8	n.i.	n.i.	Mandal and Mandal (2010)
Betulinic acid	*Ancistrocladus heyneanus* Wall. ex J. Graham	▲	95% MeOH or 90% EtOH	180	1–5	n.i.	n.i.	Pai *et al.* (2011)
Triterpenic acids and triterpenic dialcohols	*Olea europaea* L.	o	80% EtOH	180	5	n.i.	Atm.	Sanchez-Avila *et al.* (2009)
Asiaticoside	*Centella asiatica* L.	c	90% MeOH	n.i.	20	70	n.i.	Shen *et al.* (2009)
Astragalosides	*Astragalus membranaceus* var. *mongholicus* (Bge)	▲	MeOH	300	4	n.i.	n.i.	Song *et al.* (2007)
Astragalosides	*Astragalus membranaceus* var. *mongholicus* (Bge)	c	80% EtOH	700	15	70	n.i.	Yan *et al.* (2009)
Flavonoids								
Quercitrin	*Platycladus orientalis* (L.) Franco	o	80% MeOH	80	5	n.i.	Atm.	Chen *et al.* (2007b)

Table 1   (*Continued*)

Class of compounds	Plants	MAE conditions						References
		System	Solvent	Power (W)	Time (min.)	Temp. (°C)	Pressure (bar)	
Scutellarin	*Erigeron breviscapus* Hand.-Mazz.	▲	Aq. EtOH	700	Cycles power ON/OFF	60–100	Atm.	Gao *et al*. (2007)
Quercetin and its glycosides	*Psidium guajava* L.	c	EtOH	300	5	100	n.i.	Huang and Zhang (2004)
Various flavonoids and isoflavonoids	*Labisia pumila* Benth	c	MeOH	750	2	60	n.i.	Karimi and Jaafar (2011)
Flavonoids	*Acanthopanax senticosus* Harms	c and ▲(o)	90% EtOH 50% EtOH	700 n.i.	10 6	n.i. n.i.	Atm. 72.5	Liu (2009)
Isoflavones	*Astragalus membranaceus* var. *mongholicus* (Bge)	▲ (c)	MeOH	300	4	n.i.	n.i.	Song *et al*. (2007)
Flavanone	*Cajanus indicus* Spreng.	c	80% EtOH	300	2	65	n.i.	Kong *et al*. (2010)
Anthocyanins	*Rubus idaeus* L.	o	1.5 M HCl : 95% EtOH (15 : 85)	366	12.1	n.i.	n.i.	Sun *et al*. (2007b)
Narirutin, naringin, and neohesperidin	*Citrus aurantium* L. or *Citrus sinensis* Osbeck	o	90% EtOH,	800	1.5	80	n.i.	Wang *et al*. (2010a)
Flavonoids	*Astragalus mongolicus* Bge	c	90% EtOH	1000	25	110	n.i.	Xiao *et al*. (2008)
Total flavonoids	*Gossampinus malabarica* (DC.) Merr.	o	60% EtOH	240	100 s	n.i.	Atm.	Liang *et al*. (2010)
Total flavonoids	*Chaenomeles sinensis* Lindl.	o	46.24% EtOH	108.84	4.91	n.i.	Atm.	Teng *et al*. (2009)
Coumarins								
Decursin	*Angelica gigas* Nakai	o	67% EtOH	100	6	n.i.	Atm.	Lee *et al*. (2006)
Tetrahydropalmatine imperatorin and isoimperatorin	*Corydalis yanhusuo* *Angelica dahurica* Fish ex. Hoffm.	o	70% EtOH	500		n.i.	Atm.	Liao *et al*. (2008)
Furanocoumarins	*Archangelica officinalis* Hoffm.	o and c	80% MeOH	600	31	n.i.	Atm and n.i.	Waksmundzka-Hajnos *et al*. (2004b)
Furanocoumarins	*Pastinaca sativa* L.	o and c	80% MeOH	600	31	n.i.	Atm and n.i.	Waksmundzka-Hajnos *et al*. (2004a)
Isofraxidin	*Sarcandra glabra* (Thunb.)	o	60% EtOH	n.i.	20	65	Atm.	Xiao *et al*. (2009)
Coumarins	*Angelica dahurica* Fish. ex Hoffm.	o	1 step: water then 30% EtOH then 90% EtOH	800	1 + 1 + 10	70	Atm.	Xie *et al*. (2010b)
Saponins								
Saikosaponins	*Bupleurum chinense* D.C.	▲	47–50% EtOH	360–400	5.8–6.0	73–74	Atm.	Hu *et al*. (2008)
Diosgenin	*Trigonella foenum-graecum* L.	o	40% *i*PrOH	40	5–30	n.i.	Atm.	Kaufmann *et al*. (2007)
Saponins	*Cicer arietinum* L.	c	MeOH, EtOH or 70% EtOH, butanol or 50% butanol	300	10 or 20	60	n.i.	Kerem *et al*. (2005)

Table 1 (*Continued*)

Class of compounds	Plants	MAE conditions						References
		System	Solvent	Power (W)	Time (min.)	Temp. (°C)	Pressure (bar)	
Saponins	*Acanthopanax senticosus*	▲	80% EtOH	510	30	n.i.	Atm.	Wang *et al.* (2003)
Volatiles and essential oils								
	Lavandula angustifolia L.	o	no solvent	500	20	n.i.	Atm.	Iriti *et al.* (2006)
	Rosmarinus officinalis L.	o	No solvent	900	18	n.i.	Atm.	Lo Presti *et al.* (2005)
	Ferulago campestris (Bess.) = *F. galbanifera* (Mill.) Kock. =*F. ferulago* L.	o	No solvent	200	30	100	Atm.	Riela *et al.* (2011)
	Zygophyllum album L.	o	No solvent	n.i.	30	100	Atm.	Tigrine-Kordjani *et al.* (2011)
	Cuminum cyminum L. and *Zanthoxylum bungeanum* Maxim.	▲	No solvent	85	30	100	Atm.	Wang *et al.* (2006c)
	Illicium verum Hook. f, and *Zingiber officinale* Rosc.	▲	No solvent	85	30	100	Atm.	Wang *et al.* (2006b)
	Mentha haplocalyx Briq. *Citrus reticulata* Blanco	▲	No solvent	85	30	100	Atm.	Wang *et al.* (2006a)
	Aquilaria malaccensis Lam.	o	No solvent	750	30–60	n.i.	Atm.	Yahaya *et al.* (2010)
	Magnolia grandiflora L.	▲	No solvent	700	4	n.i	Atm.	Fan *et al.* (2009)
	Toona sinensis (A. Juss.) M. Roem.	▲	Water	400	4	n.i.	Atm.	Mu *et al.* (2007)
	Polygala furcata Royle	▲	Sample rehydrated (water)	400	4	n.i.	Atm.	Yang *et al.* (2010b)
Alkaloids								
Homoharringtonine	*Cephalotaxus koreana* Nakai	▲	MeOH	150	20	30	Atm.	Kim *et al.* (2010)
Oxymatrine	*Sophora flavescens* Aiton	c	60% EtOH	500	10	50	n.i.	Xia *et al.* (2011b)
Protopine and allocryptopine	*Macleaya cordata* (Willd) R.Br.	▲	45.2% EtOH	350	6	54.71	Atm.	Zhong *et al.* (2010)
Berberine	*Mahonia bealei* (Fort.)	o	MeOH, 90% EtOH, EtOH	650	6 (54% moisture content) or 4 (10% moisture content)	n.i.	Atm.	Gao *et al.* (2004)
	Chrysanthemum morifolium (Ramat.)							
Lignans								
Lignans	*Magnolia biondii* Pamp.	o	Water	n.i.	2 × 5	85	Atm.	Zhao *et al.* (2007)
Lignans	*Schisandra chinensis* (Turcz.) Baill.	o	75% EtOH	200	45 s	76	Atm.	Zhu *et al.* (2007)

Table 1 (*Continued*)

Class of compounds	Plants	MAE conditions						References
		System	Solvent	Power (W)	Time (min.)	Temp. (°C)	Pressure (bar)	
Sugars								
Simple sugars	*Zea mays* L.	c	Water	1000	3–30	140–220	n.i.	Yoshida *et al*. (2010)
Polysaccharides	*Cyclocarya paliurus* (Batal.) Iljinskaja	c	Water	800	20.8	100.9	n.i.	Xie *et al*. (2010a)
Polysaccharides	*Potentilla anserina* L.	c	Water	369	76.8	63.3	n.i.	Wang *et al*. (2010b)
Polysaccharides	*Lycoris aurea* Herb.	c	Water	n.i.	10–20	80–100	n.i.	Ru *et al*. (2009)
Mucilage	*Trichosanthes dioica* Roxb.	▲	Water	320	20	n.i.	Atm	Shah *et al*. (2010)
Miscellaneous								
Carotenoid	*Bixa orellana* L.	o	EtOAc + H$_2$O	210	18	n.i.	Atm.	Vasu *et al*. (2010)
Glucosinolates	*Eruca sativa* Mill.	c	Methanol	250	10	80	n.i.	Omirou *et al*. (2009)
Naphthohydroquinone	*Rubia cordifolia* L.	n.i.	70% EtOH	460	4	n.i.	n.i.	Ma *et al*. (2009)

c = closed system; o = open system; ▲ = domestic oven (modified or not); n.i. = not indicated.

determining persistent organic pollutants in environmental samples (Anonymous, 1995). A list of specific compounds highlighted in this EPA method has recently been published (Dean and Ma, 2010).

Numerous applications have been developed for a great variety of compounds from different matrices, including environmental (Luque-Garcia and Luque de Castro, 2004; Nieto *et al*., 2008, 2010; Ramos *et al*., 2002; Schantz, 2006) and biological (Andreu *et al*., 2007; Bjoerklund *et al*., 2006; Carabias-Martinez *et al*., 2005; Mendiola *et al*., 2007, 2008; Ridgway *et al*., 2007; Rostagno *et al*., 2010; Runnqvist *et al*., 2010) matrices. Surprisingly, only a few reviews dedicated to the extraction of medicinal plants have been published in the last decade (Huie, 2002; Kaufmann and Christen, 2002; Mustafa and Turner, 2011; Seidel, 2012; Sticher, 2008).

3.1 Principle of the Method

The technique combines elevated temperature and pressure with liquid solvents under an inert nitrogen atmosphere devoid of light that is able to improve the speed and efficiency of extraction by significant amounts. The elevated pressure used means that the solvent is retained in a liquid state at a temperature above its boiling point. This improves the contact between the solvent and the solid matrices and thus facilitates the desorption of the analytes from the plant material and increases mass transfer kinetics.

The advantages of PLE compared to traditional methods are less solvent consumption, less time required for extraction, greater process automation to achieve fast and efficient extractions, and the ability to control the temperature of extraction, which is important for thermally labile compounds. It is reported that PLE allows the reduction of solvent consumption by up to 90% (Zgorka, 2009). Furthermore, no additional filtration step is required because the matrix and the undissolved components are retained inside the extraction cell. This is particularly important in the case of online coupling of PLE with analytical methods and automation. Conversely, a major drawback is the low selectivity of PLE. Undesirable coextracted compounds therefore have to be eliminated before analysis, which is time-consuming and costly. Therefore, systems allowing simultaneous extraction and clean-up are particularly welcome in high throughput analysis. To achieve this, various sorbents or chemicals (silica gel, Florisil, ion-exchange resin, and carbon) may be mixed (Haglund and Spinnel, 2011) and packed in the extraction cell before PLE.

Another possibility is to use different solvents of increasing strength sequentially (Papagiannopoulos and Mellenthin, 2002). The extract complexity may also be reduced using a preextraction with a weak solvent to remove interfering compounds (Haglund and Spinnel, 2011).

Very recently, a new modular approach to PLE has been developed (Haglund and Spinnel, 2011). In this system, ASE cells are coupled and may

be loaded with either solid samples or adsorbents allowing the selective retention of analytes or matrix to achieve simultaneous extraction and clean-up or fractionation. Such an approach reduces the total sample preparation time and improves the selectivity of the PLE process.

After drying, grinding, homogenization, and sieving, the solid samples to be extracted are placed in a stainless steel extraction cell. Depending on the liquid content of the sample, solid dispersant or inert diluter (quartz sand and diatomaceous earth) may be added if necessary to avoid sample aggregation and clogging of the extraction cell. The cell is then filled automatically with one or more suitable solvents. The temperature is increased, thus increasing pressure in the closed cell. The selected temperature and pressure are maintained for the desired period of time. The solvent containing the analytes is then flushed into the collection vial. One or more extraction cycles can be repeated if necessary.

PLE also enables the extractions to be performed under an inert atmosphere protected from light. In addition, PLE apparatus offers the possibility of extracting multiple samples. The automation made possible, along with the reduced extraction time and solvent exposure, has contributed significantly to the development of this method over the last decade.

3.2 Methodology

PLE instruments are made of a pump able to deliver solvents in stainless steel extraction cells of various volumes (typically 1–250 mL), a flow restrictor to maintain the desired pressure (up to 200 bar), an oven to heat the extraction cell (up to 200 °C), and one or more collection vials. For samples that require extraction in basic or acidic conditions, the manufacturers provide special extraction cells.

There are two ways to conduct PLE: static and dynamic modes.

3.2.1 Static Extraction

In the static extraction, the cell is filled with dried powdered solid matrix, loaded into the oven and the solvent is added and pressurized at a user-defined setting. The vessel is heated and the solvent is held in the cell for a specified period of time. If there is overpressure, excess solvent is automatically expulsed into the collecting vial. This operation characterizes the first static extraction cycle. After the desired time, the solvent is flushed into the collecting vial using pressurized nitrogen flux. If multiple extraction cycles are required, the cell is refilled with fresh solvent and pressurized again. These cycles are repeated for as long as necessary. Once the last cycle is completed, the cell is purged of solvent with a stream of nitrogen (Smith and Strickland, 2007). The number of cycles and the duration of the static extraction periods are important for the extraction efficiency and depend on the analytes extracted and on the matrix; they can vary greatly (Runnqvist *et al.*, 2010). Performing single extraction steps and analyzing each extract allows the number of cycles required to obtain exhaustive extraction of the compounds to be determined. The optimal conditions are frequently a compromise between thorough extraction and acceptable run time. Mostly, the applied static periods range from 5 to 10 min and two to three cycles are required (Runnqvist *et al.*, 2010). Static extraction is preferred for compounds strongly bonded to the matrix. To date, the majority of PLE applications (about 75%) have been performed in the static extraction mode (Haglund and Spinnel, 2011).

3.2.2 Dynamic Extraction

In the dynamic extraction process, the first steps are similar to those for static extraction. After heating and pressurization, the static valves are released and a few milliliters of solvent are pumped continuously through the extraction cell to elute the analytes. Typically, the pump will be set at a constant flow rate of 0.5–2.0 mL min^{-1}. The solvent coming out is collected in the collection vial. Dynamic extraction is preferably used for easily extractable compounds.

In both modes, at the end of the process, a nitrogen purging step flushing through the extraction cell allows recovery of the extracted analytes in the collection vials. It is notable that for recovery of highly volatile compounds, a cooling step should be included. ASE instruments are presently the only systems commercially available that allow static and dynamic extractions in the same run.

3.3 PLE Instrumentation

The first PLE instrument was marketed by Dionex. A schematic diagram of the Thermo Scientific™
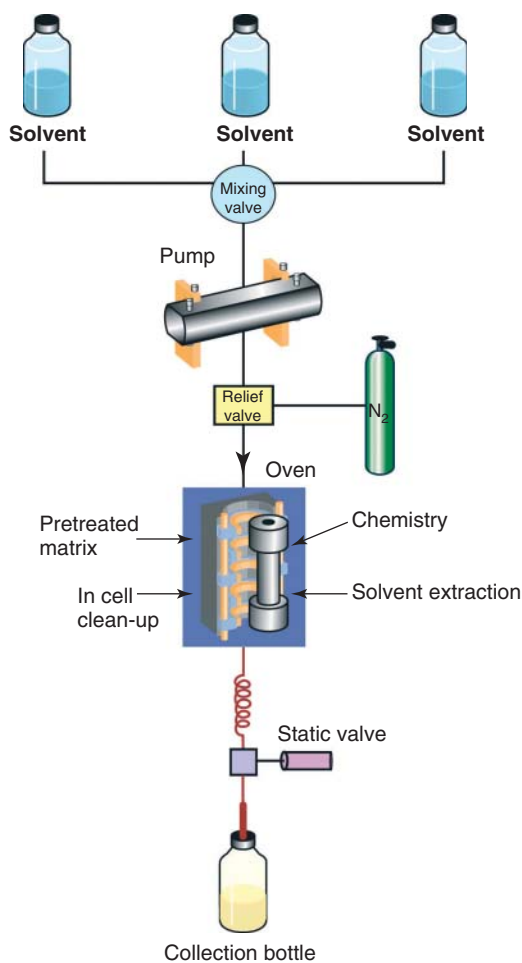
Figure 8 Schematic diagram of Thermo Scientific™ Dionex™ ASE™ equipment. (*Source*: Reproduced with permission of Thermo Fisher Scientific Inc.)

Dionex™ ASE™ equipment is shown in Figure 8. A Thermo Scientific™ Dionex™ ASE™ system 350 apparatus is shown in Figure 9. This extractor is an automated sequential solvent extraction system with a carousel that holds up to 24 samples. Other companies manufactured systems that are commercially available, such as SpeedExtractor E-916 from BUCHI Labortechnik AG (Flawil, Switzerland) (Figure 10). Up to six samples can be extracted simultaneously. The extracts are collected in vials and can be immediately evaporated on the parallel evaporator Multivapor™ P-6 without sample transfer. Another possibility is the parallel evaporation and concentration to residual volume using the Syncore® Analyst

Figure 9 Thermo Scientific™ Dionex™ ASE system 350 apparatus from Dionex. Up to 24 samples (1–100 g each) may be sequentially extracted in 1, 5, 10, 22, 34, 66, and 100 mL cell sizes. (*Source*: Reproduced with permission of Thermo Fisher Scientific Inc.)

with different vessel options (150 and 250 mL). Fluid Management Systems Inc. (Watertown, MA, USA) is another company manufacturing PLE equipment.

3.4 Influencing Factors in PLE

Several parameters, including those related to the matrix, to the solvent, and to the extraction mode, may affect the efficiency of PLE. Among them, the nature of the solvent, the temperature, and the pressure are the most significant for consideration in the PLE process. They are briefly discussed in the following sections.

3.4.1 Nature of the Solvent

The selection of a suitable solvent is frequently a critical step in the optimization of a PLE process and should be undertaken with care in order to obtain

Figure 10 E-916 SpeedExtractor device from BUCHI Labortechnik AG. Up to six samples can be simultaneously extracted in 10, 20, and 40 mL cells. The collection vessels (60, 220, and 240 mL) can be transferred directly to the parallel evaporator Multivapor™ P-6. (*Source*: Reproduced by permission of BUCHI Labortechnik AG.)

good extraction efficiency and selectivity. A wide variety of organic solvents may be used in PLE for the extraction of phytoconstituents from herbs. The most important criteria are that the solvent or a combination of solvents must be able to solubilize the analytes of interest, improve the extraction efficiency, and minimize the coextraction of undesirable components from the matrix. Economic and environmental aspects should also be considered when selecting the solvents. In this respect, ethanol, methanol, water, or alcohol/water mixtures are the most frequently used solvents for the PLE applications (Table 2). Mixtures of solvents generally provide more efficient extraction than single solvents, especially when analytes of a wide range of polarities are desired. When choosing an extraction solvent, it could be worthwhile to take into account its compatibility with future analytical steps (HPLC and GC separations). PHWE is a particular case of PLE in which water is used as an extraction solvent. A more detailed description of the basic principles of PHWE is described in the following sections.

3.4.2 Temperature

Temperature is another important factor contributing to increased extraction efficiency in PLE. Owing to the elevated pressure, PLE can be performed well above the normal boiling point of the employed solvent. High temperature increases the capacity of solvents to solubilize analytes and improves their diffusivity, resulting in increased extraction speed. It

reduces solvent viscosity and surface tension, thus enhancing its penetration into the solid matrices. At high temperature, the strong matrix–analyte interactions (Van der Waals forces, dipole interactions, and hydrogen bonding), as well as dipole attractions of the solute molecules and active sites on the matrix, are disrupted, leading to improved analyte desorption from the matrix (Richter *et al.*, 1996). Solvents have an enhanced solvation power that speeds up the release of analytes from the matrix, increasing mass transfer rate, and thus accelerating the extraction kinetics. Maximum temperature is limited by the thermal degradation of analytes. The main drawback of using high temperatures is that the quantity of coextracted compounds is also increased and thus may also affect selectivity of PLE.

3.4.3 Pressure

High pressure is mainly used to keep solvents in the liquid state at temperatures above their boiling points during the whole extraction process and does not therefore appear to be a critical parameter. It is generally not optimized except when the nature of the matrix and its physical characteristics should be considered. For instance, the efficacy of PLE for analytes that have to be isolated from wet and highly adsorptive samples can be significantly improved by increasing pressure (Ramos *et al.*, 2002). High pressure will also facilitate the penetration of the solvent into a porous matrix, which makes the extraction much faster and more efficient than conventional atmospheric pressure methods, thereby enhancing extraction recovery.

Optimization of the extraction parameters is frequently realized using the one-variable-at-a-time method in which one parameter is modified while the others are kept constant. However, these parameters may also be optimized by means of experimental designs to evaluate the robustness of the extraction method. Using this strategy, all investigated parameters are varied in each experiment. In this way, it is possible to detect the influencing factors in a minimum number of trials. For example, a central composite design was used to optimize temperature, pressure, and extraction time and allowed to evaluate the robustness of the method by drawing response surfaces (Brachet *et al.*, 2001).

Table 2   Selected applications of PLE to the extraction of natural products.

Class of compounds	Plants	PLE conditions				References
		Pressure (bar)	Temp. (°C)	Solvent	Time (min)	
Hydrocarbons						
C_{21}–C_{36} n-alkanes	*Lolium rigidum* Gaudin *Trifolium subterraneum* L.	103	100	n-Hexane	5	Smith and Strickland (2007)
Carbohydrates						
Monosaccharides and polysaccharides	*Cordyceps sinensis* (Berk.) Sacc.; *C. militaris* (L.) Link.	103	100	EtOH/H_2O (7 : 3)	10	Guan *et al.* (2010)
Oligosaccharides	*Lupinus albus* L.; *L. angustifolius* L.	103	60	EtOH/H_2O (48 : 52)	5 × 5	Bansleben *et al.* 2008)
Lipids						
Oxylipins	*Arabidopsis thaliana* (L.) Heynh.	200	RT	iPrOH	30	Thiocone *et al.* (2008)
Phospholipids	*Glycine max* (L.) Merr.	103	120–150	$CHCl_3$/CH_3OH (2 : 1)	5	Zhou *et al.* (2010)
Fatty acids	*Zea mays* L.	120	110	$CHCl_3$/CH_3OH (1:1)	20	Poerschmann *et al.* (2009)
	Ziziphus jujube Mill. var. spinosa (Bge) Hu ex H.F. Chou	83	140	CH_3OH/EtOAc (95 : 5)	15	Zhao *et al.* (2006a)
	Cordyceps sinensis (Berk.) Sacc.; *C. liangshanensis* M. Zhang, D. Liu and R. Hu; *C. gunni* Berk.; *C. militaris* (L.) Link	103	160	Petroleum ether	10	Yang *et al.* (2009)
Polyynes; polyene	*Oplopanax horridus* (Smith) Miq. *O. elatus* (Nakai) Nakai	103	100	CH_3OH	5	Huang *et al.* (2010)
α-Tocopherol	*Corylus avellana* L.	103	60	n-Hexane + 0.01% BHT	15	Sivakumar *et al.* (2005)
Phenolic compounds						
Phenolic acids	*Cimicifuga racemosa* (L.) Nutt.	69	90	CH_3OH/H_2O (6 : 4)	5	Mukhopadhyay *et al.* (2005)
	Humulus lupulus L.	n.i.	60	Pentane; acetone/H_2O (4 : 1)	20	Papagiannopoulos and Mellenthin (2002)
Gingerol, shogaol, and diarylheptanoids	*Zingiber officinale* Roscoe	103	100	EtOH/H_2O (7 : 3)	20	Hu *et al.* (2011)
Biphenols	*Magnolia biloba* (Rehd. et Wils.) Cheng	103	70–100	n-Hexane	5–15	Cheah *et al.* (2010)
Phenolic antioxidants	*Rosmarinus officinalis* L.	103	50–200	EtOH; H_2O	20	Herrero *et al.* (2010)
	Origanum vulgare L. ; *Artemisia dracunculus* L.; *Thymus serpyllum* L.	103	50–200	EtOH; H_2O or mixtures	20	Miron *et al.* (2011)

Table 2 (*Continued*)

Class of compounds	Plants	PLE conditions				References
		Pressure (bar)	Temp. (°C)	Solvent	Time (min)	
Anthraquinones + naphthalenes	*Rheum officinale* Baill.	103	140	CH_3OH	5	Gong *et al*. (2005)
	Rumex nepalensis Spreng.	100	60	CH_3OH	10	Gautam *et al*. (2011)
Chromones	*Saposhnikovia divaricata* (Turcz.) Schischk.	103	140	$EtOH/H_2O$ (1 : 1)	8	Li *et al*. (2010)
Xanthones	*Garcinia mangostana* L.	n.i.	100	EtOH	5	Destandau *et al*. (2009)
Coumarins	*Heracleum leskowii* L.	n.i.	80–110	CH_2Cl_2; petroleum ether; CH_3OH	10	Skalicka-Wozniak and Glowniak (2012)
Flavonoids	*Epimedium* sp.	103	120	$EtOH/H_2O$ (7 : 3)	10	Chen *et al*. (2008b)
	Glycyrrhiza uralensis Fisch.; *G. inflata* Bat.; *G. glabra* L.	103	100	$EtOH/H_2O$ (7 : 3)	5	Chen *et al*. (2009)
	Costus speciosus (Koen.) Sm.	20–30	n.i.	0.2% SDS in H_2O; 0.1% Triton X-100 in H_2O	30	Chang *et al*. (2011)
	Scutellaria baicalensis Georgi	10–20	120	CH_3OH	20	Ong *et al*. (2004)
	Petroselinum crispum (Mill.) A.W. Hill.	69–103	40–160	$EtOH/H_2O$ (1 : 1) or acetone/H_2O (1 : 1)	5–15	Luthria (2008)
	Trifolium incarnatum L.; *T. pannonicum* L.; *T. pratense* L.; *T. rubens* L.; *T. arvense* L.; *T. medium* L.	103	75–125	CH_3OH/H_2O (3 : 1) or $EtOH/H_2O$ (3 : 1)	5	Zgorka (2009, 2011)
	Humulus lupulus L.	n.i.	60	Pentane; acetone/H_2O (4 : 1)	20	Papagiannopoulos and Mellenthin (2002)
Rutin	*Sambucus nigra* L.	100	100	CH_3OH/H_2O (8 : 2)	10	Dawidowicz and Wianowska (2005), Dawidowicz and Wianowska (2009)
Anthocyanins + phenolic compounds	*Myrciaria cauliflora* (Mart.) O. Berg.	117	80	H_2O at pH 2.5	5	(Santos and Meireles (2011)
Rotenone	*Derris elliptica* Benth.; *D malaccensis* Prain.	138	50	$CHCl_3$	30	Sae-Yun *et al*. (2006)
Terpenoids and steroids						
Terpenes	*Picea engelmannii* Parry ex. Engelm.	60	100	*n*-Hexane	10	Mardarowicz *et al*. (2004)
Sesquiterpenes	*Curcuma* sp.	103	120	CH_3OH	5	Yang *et al*. (2005)
	Curcuma longa L.	69	140	CH_3OH	5	Qin *et al*. (2007)
Artemisinin	*Artemisia annua* L.	100	40	$EtOH/H_2O$ (3 : 1)	n.i.	Quennoz *et al*. (2010)
Diterpenoids	*Salvia miltiorrhiza* Bunge	103	100	EtOH	10	Li *et al*. (2008)
Triterpenes + ergosterol	*Ganoderma lucidum* (Leyss. ex. Fr.) Karst.; *G. sinensis*	103	100	CH_3OH	5	Zhao *et al*. (2006b)

Table 2 (*Continued*)

Class of compounds	Plants	PLE conditions				References
		Pressure (bar)	Temp. (°C)	Solvent	Time (min)	
Charantin	*Momordica charantia* L.	103	120	EtOH/H$_2$O (1:1)	40	Pitipanapong *et al.* (2007)
Saponins	*Panax ginseng* C.A. Meyer; *P. quinquefolius* L.; *P. notoginseng* (Burk.) F.H. Chen	69	150	CH$_3$OH	15	Wan *et al.* (2007); Wan *et al.* (2006a); Wan *et al.* (2006b)
	P. ginseng; *P. quinquefolius*	25–30	140	CH$_3$OH	20	Lee *et al.* (2002)
Gypenosides	*Gynostemma pentaphyllum* Makino	15	100	EtOH/H$_2$O (8:2)	180	Chen *et al.* (2007a)
Glycyrrhizin	*Glycyrrhiza glabra* L.	10–30	100	CH$_3$OH	20	Ong (2002)
Withanolides	*Iochroma gesnerioides* (Kunth) Miers	250	60, 80, 100	CH$_3$OH/H$_2$O (1:1)	10	Kaufmann *et al.* (2001b)
Alkaloids						
Amaryllidaceae	*Narcissus* sp.; *Crinum moorei* Hook. *f.*; *Scadoxus puniceus* (L.) Friis & Nordal	100	140	CH$_3$OH	2×10	Mroczek (2009)
Indole der.	*Isatis tinctoria* L.	120	70	CH$_2$Cl$_2$	2×5	Mohn *et al.* (2007)
Protoberberines	*Corydalis decumbens* (Thunb.) Pers.	103	100	EtOH/H$_2$O (9:1)	16	Shen *et al.* (2011)
Tropanes	*Datura* sp.	n.i.	90	CH$_3$OH (Hy.)	15	Mroczek *et al.* (2006)
	Erythroxylum coca Lam. var. *coca*	200	80	CH$_3$OH/tartaric acid (99:1) (Scop.) CH$_3$OH	10	Brachet *et al.* (2001)
Miscellaneous						
Benzoxazinone derivatives	*Triticum* sp.	103	150	CH$_3$OH/CH$_3$COOH (99:1)	3×5	Bonnington *et al.* (2003)
Caffeine	*Camellia* sp.; *Coffea* sp.	60	100	H$_2$O	10	Dawidowicz and Wianowska (2005)
Glucosinolates	*Isatis tinctoria* L.; *Isatis. indigotica* Fort.	120	50	CH$_3$OH/H$_2$O (7:3)	3x5	Mohn *et al.* (2008)
Nucleosides	*Cordyceps sinensis* (Berk.) Sacc.; *Cordyceps. militaris* (L.) Link.	103	160	CH$_3$OH	5	Fan *et al.* (2006); Yang and Li (2008)
Phthalides	*Angelica sinensis* (Oliv.) Diels; *A. acutiloba* (Sieb. et Zucc.); *Angelica. gigas* Nakai	83	100	CH$_3$OH	10	Lao *et al.* (2004)
Phytosterols	*Piper gaudichaudianum* Kunth	103	85	Petroleum ether	10	Peres *et al.* (2006)

3.5 Extraction of Plant Material

PLE is now extensively applied in natural product analysis. These compounds present a large variety of physicochemical properties, which complicates method development. Despite these difficulties, PLE frequently enables higher extraction yields to be obtained than extraction protocols of classical pharmacopoeia monographs (Basalo *et al.*, 2006). Ideally, the choice of extraction parameters should allow exhaustive extraction of the compounds of interest. However, this exhaustive extraction frequently

leads to coextraction of undesirable compounds and requires a clean-up procedure before analysis. Owing to the complexity of plant matrices, extraction of natural products is often a compromise among extraction efficiency, selectivity, and quantity of coextracted components. Recently, a modular approach has been developed by Haglund and Spinnel (2011). It consists of coupling extraction cells loaded both with sample for extraction and adsorbent, allowing extraction and clean-up to be carried out simultaneously. This interesting approach could be useful for complex matrices such as those found in plant materials.

Several parameters should be considered before starting PLE of natural products. While knowledge of physicochemical properties of pharmaceuticals to be extracted from relevant matrices is easily obtainable, this is not the case for the extraction of natural products from plant materials. Frequently, these values are not known, and thus many aspects of the extraction process remain unpredictable. As a rule, the solvent used for analysis in classical extraction procedures involving heat treatment can also be used for PLE. The limitation is that the conditions should not cause degradation of compounds.

It is advantageous to powder the matrix to improve extraction efficiency. In most cases, sample materials are ground to create more surface area in contact with the solvent. Sometimes, the addition of a dispersion agent avoids sample aggregation and extraction cell clogging. Inert material such as sand or diatomaceous earth may be used for this purpose.

Solvent choice is of prime importance in PLE to obtain good extraction efficiency. The major criteria for a good solvent are its ability to solubilize the analytes of interest and to minimize the coextraction of undesirable compounds (Runnqvist *et al.*, 2010). Even if a single organic solvent such as *n*-hexane (Mardarowicz *et al.*, 2004), chloroform (Sae-Yun *et al.*, 2006), methanol (Mroczek, 2009), or ethanol (Li *et al.*, 2008) can be used, a combination of organic solvents (e.g., chloroform/methanol (Poerschmann *et al.*, 2009), methanol/ethyl acetate (Zhao *et al.*, 2006a)) or a combination of an organic solvent with water (Li *et al.*, 2010; Mukhopadhyay *et al.*, 2005) may be of interest because of the large variety of polarities. There are some applications where natural products are extracted under pH control (Brachet *et al.*, 2001; Santos and Meireles, 2011).

A central composite design has been used to study and optimize extraction of cocaine and benzoylecgonine from coca leaves (Brachet *et al.*, 2001). Several parameters, including the nature of the extraction solvent, the pressure, the temperature, and the extraction time, were investigated. Extraction with toluene was less efficient than with methanol or water or a mixture of them (Figure 11). Optimal
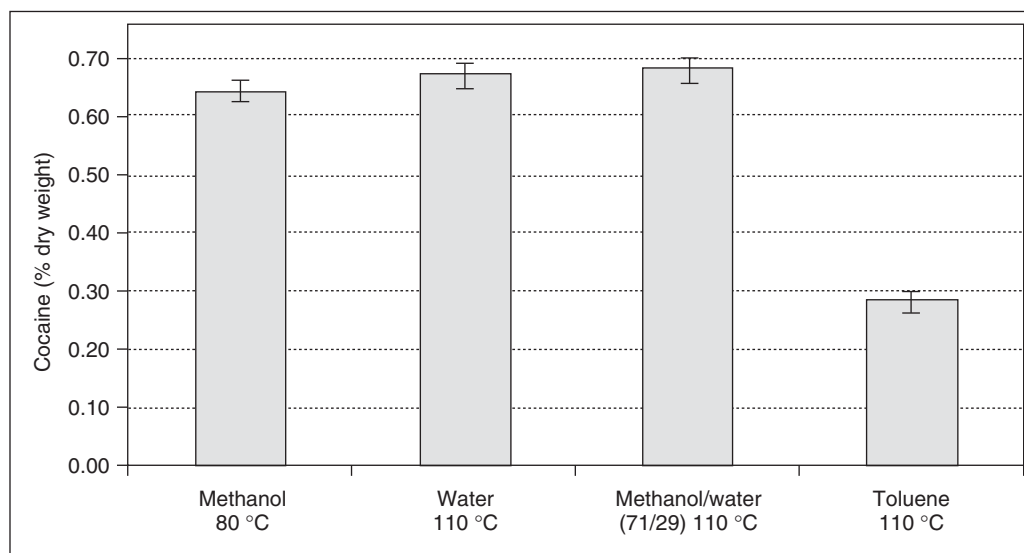
Figure 11 Influence of extracting solvent on cocaine extraction (with a 90–150 μm particle size distribution, 15 MPa, 1 mL min^{-1}, GC-FID analysis). (*Source*: A. Brachet, *et al.* (2001). Reproduced from WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.)

conditions were found to be 200 bar at 80 °C for 10 min. Determination of both alkaloids was carried out by gas chromatography coupled with mass spectrometry (Figure 12) or capillary electrophoresis with UV detection, in less than 4 min for the latter.

3.6   Pressurized Hot Water Extraction

PHWE, also called *subcritical water extraction* (*SWE*) or *superheated water extraction*, is an extraction process that is based on the same principle as PLE. PHWE uses water as a cheap, nontoxic, nonflammable, and environmentally friendly extractant. With the increasing requirement to avoid organic solvents for the extraction of medicinal plants, PHWE may provide an interesting alternative approach. In PHWE, water is heated under pressure to above its boiling point but below its critical temperature (critical point of water: 374 °C and 220 bar). The

physicochemical properties of subcritical water (e.g., polarity, density, and viscosity) significantly change when the temperature is increased, mainly because of breakdown of its hydrogen-bonded structure. Under PLE conditions, the dielectric constant (ε) of water significantly decreases (Ong *et al.*, 2006), passing from approx. 80 at 25 °C at atmospheric pressure to 27 at 250 °C and 50 bar, which falls between those of methanol and ethanol at 25 °C (Mukhopadhyay and Panja, 2009). As explained earlier for PLE, increasing the temperature reduces the surface tension and the viscosity of water, which results in an enhanced solubility of the analytes in this solvent (Ramos *et al.*, 2002). However, the effects of the extraction temperature (up to 300 °C) for the nonpolar compounds may cause the analytes present to degrade. Hence, the temperature must be carefully selected when using PHWE for extraction of constituents present in botanicals and medicinal plants (Ong *et al.*, 2006).
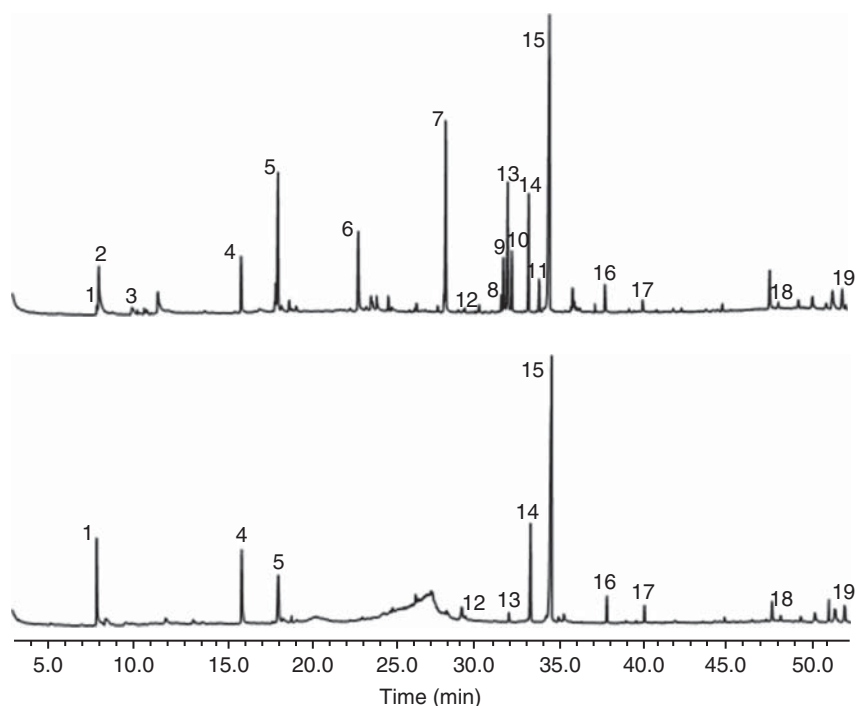
Figure 12   Total-ion current (TIC) GC–MS chromatogram of major compounds in (a) a basified coca extract with $NaHCO_3$ and (b) a nonbasified coca extract. Peaks: 1 = hygrine, 2 = benzoic acid methyl ester, 3 = tropinone, 4 = methylecgonidine, 5 = methylecgonine and pseudo-methylecgonine, 6 = cuscohygrine, 7–11 = esters of fatty acids, 12 = tropacocaine, 13 = phytol, 14 = methadone (GC internal standard), 15 = cocaine, 16 = *cis*-cinnamoylcocaine, 17 = *trans*-cinnamoylcocaine, 18 = vitamin E, 19 = β-amyrin (*Source*: A. Brachet, *et al*. (2001). Reproduced from WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.)

As for PLE, pressure has a limited influence and is mainly used to maintain water as a liquid at the extraction temperature. PHWE can be performed in the static mode or, more frequently, in the dynamic mode. The latter requires a larger volume of solvent. Some examples of applications utilizing both modes have been reviewed by Ong *et al*. (2006). To improve the recovery of some compounds, the addition of ethanol (Ong and Len, 2003b) or surfactants (Triton X-100, sodium dodecyl sulfate) may be of interest (Eng *et al*., 2007; Ong and Len, 2003a), particularly to extract more hydrophobic metabolites.

Unfortunately, no instrument dedicated to PHWE is commercially available. However, the equipment required for PHWE is similar to that used for PLE and most of the applications of PHWE reported in the literature have been performed using laboratory-made systems or Dionex ASE apparatus (Ong *et al*., 2006).

The major advantages of this process over organic solvent extraction are the absence of residual organic solvent, its safety, and the cleanliness of the product. Furthermore, PHWE apparatus can be coupled online to chromatographic systems such as LC (Kronholm *et al*., 2007).

3.7 Applications of PLE

Simple operation procedures combined with high extraction yields resulted in the employment of PLE in various analytical areas, including environmental studies, pharmaceutical, and food industries. PLE was initially used for the extraction of high temperature stable organic compounds from environmental matrices. However, during the last decade, a large number of studies of PLE applications have been published in the field of natural products involving carbohydrates, lipids, phenolics, terpenoids, and alkaloids. A selection of these is presented in Table 2. PLE has also been used to extract triterpenes and sterols in two fungi belonging to the genus *Ganoderma* (Zhao *et al*., 2006b) and mono- and polysaccharides from the edible mushrooms *Cordyceps sinensis* (Berk.) Sacc. and *Cordyceps militaris* (L.) Link. (Guan *et al*., 2010).

Most PLE applications reported in the literature employ the same organic solvents as those commonly used in conventional techniques (Huie, 2002). One extraction cycle is generally applied for 5–20 min at temperatures ranging from 50 to 140 °C in the vast majority of applications.

PLE has also been applied to the extraction of three withanolides from the leaves of *Iochroma gesnerioides* (Kunth) Miers (Kaufmann *et al*., 2001b). Several parameters including extracting solvent, its flow rate, the pressure and temperature, and the particle size of the matrix have been studied and their influence on extraction efficiency has been determined. It was demonstrated that PLE gave a similar level of recovery and selectivity as Soxhlet extraction but that the analysis time and solvent consumption were significantly reduced.

Likewise, the extraction of cocaine and benzoylecgonine has been carried out by PLE. A central composite design was applied to optimize critical parameters (pressure, temperature, and extraction time) (Brachet *et al*., 2001). A total of 10 min at 80 °C and 200 bar was sufficient to extract cocaine quantitatively from coca leaves using methanol as the extracting solvent.

The addition of surfactants may be helpful in improving the extraction process. Water supplemented with sodium dodecyl sulfate or TritonX-100 has been used to extract flavonoids from *Costus speciosus* (Koen.) Sm. (Chang *et al*., 2011). Extraction under acidic or basic conditions is sometimes carried out to alter the pH and thus improve the extraction efficacy. Anthocyanins and phenolic compounds were extracted from fruits of *Myrciaria cauliflora* (Mart.) O. Berg. with water at pH 2.5 (Santos and Meireles, 2011).

Numerous applications of PHWE for medicinal plant extraction have been reviewed (Herrero *et al*., 2006; Kronholm *et al*., 2007; Ong *et al*., 2006), which demonstrate that this technique is valuable for extraction of highly to moderately polar metabolites and for the chemical standardization and quality control of medicinal plants.

4 CONCLUSION

There is a real need for new methods of sample preparation with low solvent consumption in a fast and automatic way. MAE and PLE can fulfill these criteria by providing fast, clean, and cheap procedures that can be used in routine analysis. Both methods are used to extract solid matrices, although liquid or semi-solid matrices can also be extracted using preliminary adsorption on inert supports. However, the difficulty of handling two liquid phases under pressure precluded the use of these processes for liquid

matrices. Unlike SFE, MAE and PLE can be used with all the conventional solvents, thus increasing their range of applicability from polar to nonpolar compounds.

As drastic methods, MAE and PLE have a limited selectivity and require thermal stability of the compounds to be extracted. Nevertheless, because extraction times are drastically shortened, thermal degradation is not necessarily observed.

These two approaches should definitely be considered for fast and efficient extraction of active constituents from plant materials.

REFERENCES

Ananth, V. S., Rajasekhar, K. K., Sivaprasad, Y., *et al*. (2009) *J. Pharm. Res.*, **2**, 1739–1741.

Andreu, V., Blasco, C. and Pico, Y. (2007) *TrAC, Trends Anal. Chem.*, **26**, 534–556.

Anonymous (1995) *US EPA 3545A, Test Methods for Evaluating Solid Waste, USEPA SW-846*, 3rd edn, US Environmental Protection Agency Washington, DC.

Arthur, C. L. and Pawliszyn, J. (1990) *Anal. Chem.*, **62**, 2145–2148.

Asghari, J., Ondruschka, B. and Mazaheritehrani, M. (2011) *J. Med. Plants Res.*, **5**, 495–506.

Bansleben, D., Schellenberg, I. and Wolff, A.-C. (2008) *J. Sci. Food Agric.*, **88**, 1949–1953.

Basalo, C., Mohn, T. and Hamburger, M. (2006) *Planta Med.*, **72**, 1157–1162.

Belanger, J. M. R., Pare, J. R. J., Sanchez Fulvia, N., *et al*. (2008) *Anal. Bioanal. Chem.*, **391**, 929–932.

Bieri, S., Munoz, O., Veuthey, J.-L., *et al*. (2006) *J. Sep. Sci.*, **29**, 96–102.

Bjoerklund, E., Sporring, S., Wiberg, K., *et al*. (2006) *TrAC, Trends Anal. Chem.*, **25**, 318–325.

Bonnington, L., Eljarrat, E., Guillamon, M., *et al*. (2003) *Anal. Chem.*, **75**, 3128–3136.

Brachet, A., Rudaz, S., Mateus, L., *et al*. (2001) *J. Sep. Sci.*, **24**, 865–873.

Camel, V. (2002) *Anal. Bioanal. Chem.*, **372**, 39–40.

Carabias-Martinez, R., Rodriguez-Gonzalo, E., Revilla-Ruiz, P., *et al*. (2005) *J. Chromatogr., A*, **1089**, 1–17.

Chan, C.-H., Yusoff, R., Ngoh, G.-C., *et al*. (2011) *J. Chromatogr., A*, **1218**, 6213–6225.

Chang, Y.-Q., Tan, S.-N., Yong, J. W. H., *et al*. (2011) *J. Sep. Sci.*, **34**, 462–468.

Cheah, E. L. C., Heng, P. W. S. and Chan, L. W. (2010) *Sep. Purif. Technol.*, **71**, 293–301.

Chemat, S., Ait-Amar, H., Lagha, A., *et al*. (2005) *Chem. Eng. Process.*, **44**, 1320–1326.

Chen, S. S. and Spiro, M. (1995) *Flavour Fragrance J.*, **10**, 101–112.

Chen, C.-H., Huang, T.-Y., Lee, M.-R., *et al*. (2007a) *Ind. Eng. Chem. Res.*, **46**, 8138–8143.

Chen, L., Ding, L., Yu, A., *et al*. (2007b) *Anal. Chim. Acta*, **596**, 164–170.

Chen, L., Jin, H., Ding, L., *et al*. (2007c) *J. Chromatogr., A*, **1140**, 71–77.

Chen, L., Song, D., Tian, Y., *et al*. (2008a) *TrAC, Trends Anal. Chem.*, **27**, 151–159.

Chen, X-j., Ji, H., Wang, Y-t., *et al*. (2008b) *J. Sep. Sci.*, **31**, 881–887.

Chen, X. J., Zhao, J., Meng, Q., *et al*. (2009) *J. Chromatogr., A*, **1216**, 7329–7335.

Chi, Y., Zhang, Z., Li, C., *et al*. (2011) *Green Chem*, **13**, 666–670.

Dai, J. and Orsat, V. (2010) *Int. J. Food Eng.*, **6**, article 3. DOI: 10.2202/1556-3758.1636.

Dawidowicz, A. L. and Wianowska, D. (2005) *J. Pharm. Biomed. Anal.*, **37**, 1161–1165.

Dawidowicz, A. L. and Wianowska, D. (2009) *J. Chromatogr. Sci.*, **47**, 914–918.

Dean, J. R. and Ma, R. (2010) *Pressurized fluid extraction*. In: *Handbook of sample preparation*, (Pawliszyn, J. and Lord, H.L., eds), John Wiley & Sons, Inc., Hoboken, New Jersey, pp. 163–179.

Deng, C., Mao, Y., Yao, N., *et al*. (2006a) *Anal. Chim. Acta*, **575**, 120–125.

Deng, C., Yao, N., Wang, B., *et al*. (2006b) *J. Chromatogr., A*, **1103**, 15–21.

Deng, C., Ji, J., Li, N., *et al*. (2006c) *J. Chromatogr., A*, **1117**, 115–120.

Deng, C., Liu, N., Gao, M., *et al*. (2007) *J. Chromatogr., A*, **1153**, 90–96.

Deng, J., Xiao, X., Tong, X., *et al*. (2010) *Sep. Purif. Technol.*, **74**, 155–159.

Destandau, E., Toribio, A., Lafosse, M., *et al*. (2009) *J. Chromatogr., A*, **1216**, 1390–1394.

Du, F.-Y., Xiao, X.-H., Luo, X.-J., *et al*. (2009) *Talanta*, **78**, 1177–1184.

Du, F. Y., Xiao, X. H., Xu, P. P., *et al*. (2010) *Acta Chromatogr.*, **22**, 459–471.

Du, G., Zhao, H. Y., Zhang, Q. W., *et al*. (2010b) *J. Chromatogr., A* **1217** 705–714.

Eng, A. T. W., Heng, M. Y. and Ong, E. S. (2007) *Anal. Chim. Acta*, **583**, 289–295.

Fan, H., Li, S. P., Xiang, J. J., *et al*. (2006) *Anal. Chim. Acta*, **567**, 218–228.

Fan, Z. Q., Wang, S. B., Mu, R. M., *et al*. (2009) *J. Anal. Chem.*, **64**, 289–294.

Fang, X., Wang, J., Zhou, H., *et al*. (2009) *J. Sep. Sci.*, **32**, 2455–2461.

Ferhat, M. A., Tigrine-Kordjani, N., Chemat, S., *et al*. (2007) *Chromatographia*, **65**, 217–222.

Gao, S., Han, W. and Deng, X. (2004) *Flavour Fragrance J.*, **19**, 244–250.

Gao, M., Huang, W., RoyChowdhury, M., *et al*. (2007) *Anal. Chim. Acta*, **591**, 161–166.

Gautam, R., Srivastava, A. and Jachak, S. M. (2011) *Phytochem. Anal.*, **22**, 153–157.

Gong, Y. X., Li, S. P., Wang, Y. T., *et al*. (2005) *Electrophoresis*, **26**, 1778–1782.

Guan, J., Yang, F.-Q. and Li, S.-P. (2010) *Molecules*, **15**, 4227–4241.

Haglund, P. and Spinnel, E. (2011) *LCGC North Am.*, **28**, 66–72.

Hao, J.-y., Han, W., S-d, H., *et al*. (2002) *Sep. Purif. Technol.*, **28**, 191–196.

Hemwimon, S., Pavasant, P. and Shotipruk, A. (2007) *Sep. Purif. Technol.*, **54**, 44–50.

Herrero, M., Cifuentes, A. and Ibanez, E. (2006) *Food Chem.*, **98**, 136–148.

Herrero, M., Plaza, M., Cifuentes, A., *et al*. (2010) *J. Chromatogr., A*, **1217**, 2512–2520.

Hu, Z., Cai, M. and Liang, H.-H. (2008) *Sep. Purif. Technol.*, **61**, 266–275.

Hu, F., Deng, C., Liu, Y., *et al*. (2009) *Talanta*, **77**, 1299–1303.

Hu, J., Guo, Z., Glasius, M., *et al*. (2011) *J. Chromatogr., A*, **1218**, 5765–5773.

Huang, J. and Zhang, Z. (2004) *Anal. Sci.*, **20**, 395–397.

Huang, W., Yang, J., Zhao, J., *et al*. (2010) *J. Pharm. Biomed. Anal.*, **53**, 906–910.

Huang, T., Chen, N., Lai, Y., *et al*. (2011) *J. Med. Plants Res.*, **5**, 1313–1320.

Huie, C. W. (2002) *Anal. Bioanal. Chem.*, **373**, 23–30.

Iriti, M., Colnaghi, G., Chemat, F., *et al*. (2006) *Flavour Fragrance J.*, **21**, 704–712.

Jain, T., Jain, V., Pandey, R., *et al*. (2009) *Asian J. Res. Chem.*, **2**, 19–25.

Japon-Lujan, R., Luque-Rodriguez, J. M. and Luque de Castro, M. D. (2006) *Anal. Bioanal. Chem.*, **385**, 753–759.

Jyothi, D., Khanam, S. and Sultana, R. (2010) *Int. J. Pharm. Pharm. Sci.*, **2**, 46–50.

Karimi, E. and Jaafar, H. Z. E. (2011) *Molecules*, **16**, 6791–6805.

Kaufmann, B. and Christen, P. (2002) *Phytochem. Anal.*, **13**, 105–113.

Kaufmann, B., Christen, P. and Veuthey, J.-L. (2001a) *Phytochem. Anal.*, **12**, 327–331.

Kaufmann, B., Christen, P. and Veuthey, J. L. (2001b) *Chromatographia*, **54**, 394–398.

Kaufmann, B., Souverain, S., Cherkaoui, S., *et al*. (2002) *Chromatographia*, **56**, 137–141.

Kaufmann, B., Rudaz, S., Cherkaoui, S., *et al*. (2007) *Phytochem. Anal.*, **18**, 70–76.

Kerem, Z., German-Shashoua, H. and Yarden, O. (2005) *J. Sci. Food Agric.*, **85**, 406–412.

Kim, S.-J., Murthy, H. N., Hahn, E.-J., *et al*. (2007) *Sep. Purif. Technol.*, **56**, 401–406.

Kim, W.-K., Chae, H.-J. and Kim, J.-H. (2010) *Biotechnol. Bioprocess Eng.*, **15**, 481–487.

Kong, Y., Zu, Y.-G., Fu, Y.-J., *et al*. (2010) *J. Food Compos. Anal.*, **23**, 382–388.

Kronholm, J., Hartonen, K. and Riekkola, M.-L. (2007) *TrAC, Trends Anal. Chem.*, **26**, 396–412.

Lao, S. C., Li, S. P., Kan, K. K. W., *et al*. (2004) *Anal. Chim. Acta*, **526**, 131–137.

Lee, H. K., Koh, H. L., Ong, E. S., *et al*. (2002) *J. Sep. Sci.*, **25**, 160–166.

Lee, G.-D., Lee, S.-Y., Kim, K.-S., *et al*. (2006) *Int. J. Food Sci. Technol.*, **41**, 737–742.

Li, P., Xu, G., Li, S.-P., *et al*. (2008) *J. Agric. Food Chem.*, **56**, 1164–1171.

Li, W., Wang, Z., Chen, L., *et al*. (2010) *J. Sep. Sci.*, **33**, 2881–2887.

Liang, Y.-z., J-y, P., Li, Y., *et al*. (2010) *Med. Plant*, **1**, 71–73.

Liao, Z.-G., Wang, G.-F., Liang, X.-L., *et al*. (2008) *Sep. Purif. Technol.*, **63**, 424–433.

Liu, Y.-b. (2009) *Chem. Res. Chin. Univ.*, **25**, 439–442.

Lo Presti, M., Raqusa, S., Trozzi, A., *et al*. (2005) *J. Sep. Sci.*, **28**, 273–280.

Lu, Y., Yue, X.-F., Zhang, Z.-Q., *et al*. (2007) *Chromatographia*, **66**, 443–446.

Lu, Y., Ma, W., Hu, R., *et al*. (2008) *J. Chromatogr., A*, **1208**, 42–46.

Lu, C., Wang, H., Lv, W., *et al*. (2011) *Chromatographia*, **74**, 139–144.

Luque-Garcia, J. L. and Luque de Castro, M. D. (2004) *TrAC, Trends Anal. Chem.*, **23**, 102–108.

Luthria, D. L. (2008) *Food Chem.*, **107**, 745–752.

Ma, W., Lu, Y., Dai, X., *et al*. (2009) *Sep. Sci. Technol.*, **44**, 995–1006.

Ma, W., Lu, Y., Hu, R., *et al*. (2010) *Talanta*, **80**, 1292–1297.

Majors, R. E. (2006) *LCGC North Am*, **24**, 648, 650, 652, 654, 656–658, 660.

Mandal, S. C. and Mandal, V. 2008. (India). IN patent 2008-KO725 2008KO00725.

Mandal, V. and Mandal, S. C. (2010) *Biochem. Eng. J.*, **50**, 63–70.

Mandal, V., Mohan, Y. and Hemalatha, S. (2007) *Pharmacogn. Rev.*, **1**, 7–18.

Mandal, V., Mohan, Y. and Hemalatha, S. (2008) *J. Pharm. Biomed. Anal.*, **46**, 322–327.

Mao, Y., Li, Y. and Yao, N. (2007) *J. Pharm. Biomed. Anal.*, **45**, 510–515.

Mardarowicz, M., Wianowska, D., Dawidowicz, A. L., *et al*. (2004) *Z. Naturforsch., C. J. Biosci.*, **59**, 641–648.

Martins, S., Aguilar, C. N., de la Garza-Rodriguez, I., *et al*. (2010) *J. Chem. Technol. Biotechnol.*, **85**, 1142–1147.

Mendiola, J. A., Herrero, M., Cifuentes, A., *et al*. (2007) *J. Chromatogr., A*, **1152**, 234–246.

Mendiola, J. A., Rodriguez-Meizoso, I., Senorans, F. J., *et al*. (2008) *EJEAFChe, Electron. J. Environ., Agric. Food Chem.*, **7**, 3301–3309.

Miron, T. L., Plaza, M., Bahrim, G., *et al*. (2011) *J. Chromatogr., A*, **1218**, 4918–4927.

Mirzajani, F., Ghassempour, A., Jalali-Heravi, M., *et al*. (2010) *Phytochem. Anal.*, **21**, 544–549.

Mohn, T., Potterat, O. and Hamburger, M. (2007) *Planta Med.*, **73**, 151–156.

Mohn, T., Suter, K. and Hamburger, M. (2008) *Planta Med.*, **74**, 582–587.

Mroczek, T. (2009) *J. Chromatogr., A*, **1216**, 2519–2528.

Mroczek, T., Glowniak, K. and Kowalska, J. (2006) *J. Chromatogr., A*, **1107**, 9–18.

Mu, R., Wang, X., Liu, S., *et al*. (2007) *Chromatographia*, **65**, 463–467.

Mukhopadhyay, M. and Panja, P. (2009) *Indian Chem. Eng.*, **51**, 311–324.

Mukhopadhyay, S., Luthria, D. L. and Robbins, R. J. (2005) *J. Sci. Food Agric.*, **86**, 156–162.

Mustafa, A. and Turner, C. (2011) *Anal. Chim. Acta*, **703**, 8–18.

Nieto, A., Borrull, F., Marce, R. M., *et al*. (2008) *Curr. Anal. Chem.*, **4**, 157–167.

Nieto, A., Borrull, F., Pocurull, E., *et al*. (2010) *TrAC, Trends Anal. Chem.*, **29**, 752–764.

Nuechter, M., Ondruschka, B., Fischer, B., *et al*. (2005) *Chem. Ing. Tech.*, **77**, 171–175.

Omirou, M., Papastylianou, I., Iori, R., *et al*. (2009) *Phytochem. Anal.*, **20**, 214–220.

Ong, E. S. (2002) *J. Sep. Sci.*, **25**, 825–831.

Ong, E. S. and Len, S. M. (2003a) *J. Sep. Sci.*, **26**, 1533–1540.

Ong, E. S. and Len, S. M. (2003b) *Anal. Chim. Acta*, **482**, 81–89.

Ong, E. S., Len, S. M., Lee, A. C. H., *et al*. (2004) *Rapid Commun. Mass Spectrom.*, **18**, 2522–2530.

Ong, E. S., Cheong, J. S. H. and Goh, D. (2006) *J. Chromatogr., A*, **1112**, 92–102.

Pai, S. R., Nimbalkar, M. S., Pawar, N. V., *et al*. (2011) *Ind. Crops Prod.*, **34**, 1458–1464.

Pan, X., Niu, G. and Liu, H. (2002) *Biochem. Eng. J.*, **12**, 71–77.

Papagiannopoulos, M. and Mellenthin, A. (2002) *J. Chromatogr., A*, **976**, 345–348.

Pare, J. R. J. and Belanger, J. M. R. (2010) Microwave-assisted extraction. In: *Handbook of sample preparation.*, (Pawliszyn, J. and Lord, H.L., eds), John Wiley & Sons, Inc. Hoboken, New Jersey. pp. 197–224.

Peres, V. F., Saffi, J., Melecchi, M. I. S., *et al*. (2006) *J. Chromatogr., A*, **1105**, 148–153.

Pitipanapong, J., Chitprasert, S., Goto, M., *et al*. (2007) *Sep. Purif. Technol.*, **52**, 416–422.

Poerschmann, J., Rauschen, S., Langer, U., *et al*. (2009) *J. Agric. Food Chem.*, **57**, 127–132.

Proestos, C. and Komaitis, M. (2008) *LWT–Food Sci. Technol.*, **41**, 652–659.

Qin, N. Y., Yang, F. Q., Wang, Y. T., *et al*. (2007) *J. Pharm. Biomed. Anal.*, **43**, 486–492.

Qu, C., Bai, Y., Jin, X., *et al*. (2009) *Food Chem.*, **115**, 340–346.

Quennoz, M., Bastian, C., Simonnet, X., *et al*. (2010) *Chimia*, **64**, 755–757.

Raman, G. and Gaikar, V. G. (2002) *Ind. Eng. Chem. Res.*, **41**, 2521–2528.

Ramos, L., Kristenson, E. M. and Brinkman, U. A. T. (2002) *J. Chromatogr., A*, **975**, 3–29.

Richter, B. E., Jones, B. A., Ezzell, J. L., *et al*. (1996) *Anal. Chem.*, **68**, 1033–1039.

Ridgway, K., Lalljie, S. P. D. and Smith, R. M. (2007) *J. Chromatogr., A*, **1153**, 36–53.

Riela, S., Bruno, M., Rosselli, S., *et al*. (2011) *J. Sep. Sci.*, **34**, 483–492.

Romanik, G., Gilgenast, E., Przyjazny, A., *et al*. (2007) *J. Biochem. Biophys. Methods*, **70**, 253–261.

Rostagno, M. A., D'Arrigo, M. and Martinez, J. A. (2010) *TrAC, Trends Anal. Chem.*, **29**, 553–561.

Rouessac, F. and Rouessac, A. (2004) *Analyse Chimique: Méthodes et Techniques Instrumentales Modernes*, 6th ed., ed. Dunod , Paris 462.

Ru, Q. M., Zhang, L. R., Chen, J. D., *et al*. (2009) *Chem. Nat. Compd.*, **45**, 474–477.

Runnqvist, H., Bak, S. A., Hansen, M., *et al*. (2010) *J. Chromatogr., A*, **1217**, 2447–2470.

Sae-Yun, A., Ovatlarnporn, C., Itharat, A., *et al*. (2006) *J. Chromatogr., A*, **1125**, 172–176.

Sanchez-Avila, N., Priego-Capote, F., Ruiz-Jimenez, J., *et al*. (2009) *Talanta*, **78**, 40–48.

Santos, D. T. and Meireles, M. A. A. (2011) *Innov. Food Sci. Emerg. Technol.*, **12**, 398–406.

Schantz, M. M. (2006) *Anal. Bioanal. Chem.*, **386**, 1043–1047.

Seidel, V. (2012) *Methods Mol. Biol.*, **864**, 27–41.

Shah, B. N., Seth, A. K. and Nayak, B. S. (2010) *Int. J. Pharm. Sci. Res.*, **1**, 68–72.

Shen, Y., Han, C., Liu, J., *et al*. (2008) *Chromatographia*, **68**, 679–682.

Shen, Y., Liu, A., Ye, M., *et al*. (2009) *Chromatographia*, **70**, 431–438.

Shen, Y., Han, C., Jiang, Y., *et al*. (2011) *Talanta*, **84**, 1026–1031.

Shi, W., Wang, Y., Li, J., *et al*. (2007) *Food Chem.*, **102**, 664–668.

Sivakumar, G., Bacchetta, L., Gatti, R., *et al*. (2005) *J. Plant Physiol.*, **162**, 1280–1283.

Skalicka-Wozniak, K. and Glowniak, K. (2012) *Molecules*, **17**, 4133–4141.

Smith, L. L. and Strickland, J. R. (2007) *J. Agric. Food Chem.*, **55**, 7301–7307.

Song, J.-Z., Mo, S.-F., Yip, Y.-K., *et al*. (2007) *J. Sep. Sci.*, **30**, 819–824.

Spiro, M. and Chen, S. S. (1995) *Flavour Fragrance J.*, **10**, 259–272.

Sterbova, D., Matejicek, D., Vlcek, J., *et al*. (2004) *Anal. Chim. Acta*, **513**, 435–444.

Sticher, O. (2008) *Nat. Prod. Rep.*, **25**, 517–554.

Sun, C., Xie, Y. and Liu, H. (2007a) *Chin. J. Chem. Eng.*, **15**, 474–477.

Sun, Y., Liao, X., Wang, Z., *et al*. (2007b) *Eur. Food Res. Technol.*, **225**, 511–523.

Sutivisedsak, N., Cheng, H. N., Willett, J. L., *et al*. (2010) *Food Res. Int.*, **43**, 516–519.

Tan, S. N., Yong, J. W. H., Teo, C. C., *et al*. (2011) *Talanta*, **83**, 891–898.

Teng, H., Ghafoor, K. and Choi, Y. H. (2009) *J. Korean Soc. Appl. Biol. Chem.*, **52**, 694–701.

Teo, C. C., Tan, S. N., Yong, J. W. H., *et al*. (2009) *J. Sep. Sci.*, **32**, 613–622.

Thiocone, A., Farmer, E. E. and Wolfender, J.-L. (2008) *Phytochem. Anal.*, **19**, 198–205.

Tigrine-Kordjani, N., Meklati, B. Y. and Chemat, F. (2011) *Phytochem. Anal.*, **22**, 1–9.

Vasu, S., Palaniyappan, V., Kothandam, H. P., *et al*. (2010) *Pharm. Lett.*, **2**, 479–485.

Waksmundzka-Hajnos, M., Petruczynik, A., Dragan, A., *et al*. (2004a) *J. Chromatogr., B Anal. Technol. Biomed. Life Sci.*, **800**, 181–187.

Waksmundzka-Hajnos, M., Petruczynik, A., Dragan, A., *et al*. (2004b) *Phytochem. Anal.*, **15**, 313–319.

Waksmundzka-Hajnos, M., Oniszczuk, A., Szewczyk, K., *et al*. (2007) *Acta Chromatogr.*, **19**, 227–237.

Wan, J. B., Yang, F. Q., Li, S. P., *et al*. (2006a) *J. Pharm. Biomed. Anal.*, **41**, 1596–1601.

Wan, J. B., Lai, C. M., Li, S. P., *et al*. (2006b) *J. Pharm. Biomed. Anal.*, **41**, 274–279.

Wan, J.-B., Li, S-p, Chen, J.-M., *et al*. (2007) *J. Sep. Sci.*, **30**, 825–832.

Wang, J., Shen, P. and Shen, Y. (2003) *Chin. J. Chem. Eng.*, **11**, 231–233.

Wang, Z.-M., Ding, L., Wang, L., *et al*. (2006a) *Chin. J. Chem.*, **24**, 649–652.

Wang, Z., Wang, L., Li, T., *et al*. (2006b) *Anal. Bioanal. Chem.*, **386**, 1863–1868.

Wang, Z., Ding, L., Li, T., *et al*. (2006c) *J. Chromatogr. A*, **1102**, 11–17.

Wang, C., Pan, Y., Fan, G., *et al*. (2010a) *Biomed. Chromatogr.*, **24**, 235–244.

Wang, J., Zhang, J., Zhao, B., *et al*. (2010b) *Carbohydr. Polym.*, **80**, 84–93.

Xia, E.-Q., Wang, B.-W., Xu, X.-R., *et al*. (2011a) *Int. J. Mol. Sci.*, **12**, 5319–5329.

Xia, E.-Q., Cui, B., Xu, X.-R., *et al*. (2011b) *Molecules*, **16**, 7391–7400.

Xiao, W., Han, L. and Shi, B. (2008) *Sep. Purif. Technol.*, **62**, 614–618.

Xiao, X., Guo, Z., Deng, J., *et al*. (2009) *Sep. Purif. Technol.*, **68**, 250–254.

Xie, J.-H., Xie, M.-Y., Shen, M.-Y., *et al*. (2010a) *J. Sci. Food Agric.*, **90**, 1353–1360.

Xie, Y., Zhao, W., Zhou, T., *et al*. (2010b) *Phytochem. Anal.*, **21**, 473–482.

Yahaya, B. S. B., Khalid, K. B. and bin Hj. Md. Sukari, A. (2010) *Solid State Sci. Technol.*, **18**, 209–216.

Yan, M.-M., Liu, W., Fu, Y.-J., *et al*. (2009) *Food Chem.*, **119**, 1663–1670.

Yang, F. Q. and Li, S. P. (2008) *J. Pharm. Biomed. Anal.*, **48**, 231–235.

Yang, F. Q., Li, S. P., Chen, Y., *et al*. (2005) *J. Pharm. Biomed. Anal.*, **39**, 552–558.

Yang, F. Q., Feng, K., Zhao, J., *et al*. (2009) *J. Pharm. Biomed. Anal.*, **49**, 1172–1178.

Yang, Y.-C., Li, J., Zu, Y.-G., *et al*. (2010a) *Food Chem.*, **122**, 373–380.

Yang, Z., Mao, H., Long, C., *et al*. (2010b) *Eur. Food Res. Technol.*, **230**, 779–784.

Yoshida, T., Tsubaki, S., Teramoto, Y., *et al*. (2010) *Bioresour. Technol.*, **101**, 7820–7826.

Yu, Y., Wang, Z.-M., Wang, Y.-T., *et al*. (2007) *Chin. J. Chem.*, **25**, 346–350.

Yuan, Y., Wang, Y., Xu, R., *et al*. (2011) *Analyst (Cambridge, U. K.)*, **136**, 2294–2305.

Zeng, H., Wang, Y., Kong, J., *et al*. (2010) *Talanta*, **83**, 582–590.

Zgorka, G. (2009) *Talanta*, **79**, 46–53.

Zgorka, G. (2011) *J. AOAC Int.*, **94**, 22–31.

Zhang, B., Yang, R. and Liu, C.-Z. (2008) *Sep. Purif. Technol.*, **62**, 480–483.

Zhang, L., Wang, Y., Wu, D., *et al*. (2011) *Molecules*, **16**, 4428–4437.

Zhao, J., Li, S. P., Yang, F. Q., *et al*. (2006a) *J. Chromatogr., A*, **1108**, 188–194.

Zhao, J., Zhang, X.-Q., Li, S.-P., *et al*. (2006b) *J. Sep. Sci.*, **29**, 2609–2615.

Zhao, W., Zhou, T., Fan, G., *et al*. (2007) *J. Sep. Sci.*, **30**, 2370–2381.

Zhong, M., Huang, K.-L., Zeng, J.-G., *et al*. (2010) *J. Sep. Sci.*, **33**, 2160–2167.

Zhou, L., Le Grandois, J., Marchioni, E., *et al*. (2010) *J. Agric. Food Chem.*, **58**, 9912–9917.

Zhu, M., Cao, Y. and Fan, G. (2007) *J. Sep. Sci.*, **30**, 67–73.

5

Solid-Phase Microextraction (SPME) and Its Application to Natural Products

M. Abdul Mottaleb[1,2], Mohammed J. Meziani[2] and M. Rafiq Islam[2]

[1]*Center for Innovation and Entrepreneurship, Northwest Missouri State University, Maryville, MO, USA* and [2]*Department of Natural Sciences, Northwest Missouri State University, Maryville, MO, USA*

1 INTRODUCTION

Sample preparation is a key step in separation that plays a dominant role in choosing analytical methods because of reproducibility, accuracy, selectivity, and sensitivity of the resulting effective time and cost of analysis. Employing suitable preparation methodologies, nowadays, analytical and separation methods can essentially examine all kinds of complex sample mixtures that include volatile, semivolatile, and nonvolatile compounds and/or biological macromolecules, with very low detection limits. In general, the analytical method includes sampling of the samples, preparation of samples (isolation from the matrix, preconcentration or concentration, fractionation, and, if necessary, derivatization), separation, detection/characterization, and data analysis. Surveys display that more than 80% of analysis time is spent on sample collection and sample preparation because majority of analytical equipment cannot handle sample matrices directly. The whole analytical process becomes fruitless if an unsuitable sample preparation method has been used before the sample reaches the instrument and/or the analyzer (Pawliszyn, 2003; Smith, 2003). This chapter focuses on describing fundamental theory and basic operation procedure of solid-phase microextraction (SPME) technique and its application to chemically and biologically important plant extracts.

It is well known that the successes of any analytical methods are dependent on sample preparation technique. Many sample preparation approaches and equipments including SPME are commercially available. SPME is a simple, fast, efficient, and solvent-free sample preparation technique, invented by Pawliszyn and coworkers in 1990 (Arthur and Pawliszyn, 1990). This relatively newer sample extraction tool brings some unique capabilities to the chromatographic analysis of dilute solutions in difficult matrices, both gaseous and liquid states. SPME has three main steps: (i) solute adsorption from the sample matrix onto a thick (relative to conventional capillary columns) layer of chemically modified fused-silica or related adsorptive material, (ii) reaching equilibrium of adsorbed solutes and coating fiber, and (iii) transfer of the analytes into a chromatography inlet system by gaseous or liquid means. This effective adsorption/desorption technique essentially eliminates the need for solvents or complicated apparatus for concentrating volatile, semivolatile, and nonvolatiles compounds, natural products, and chemical and biological macromolecules in liquid samples or headspace (Zhang and Pawliszyn, 1993; Gioti *et al*., 2007; Zhang *et al*. 2011; Zhou *et al*., 2008). SPME is compatible with analyte separation
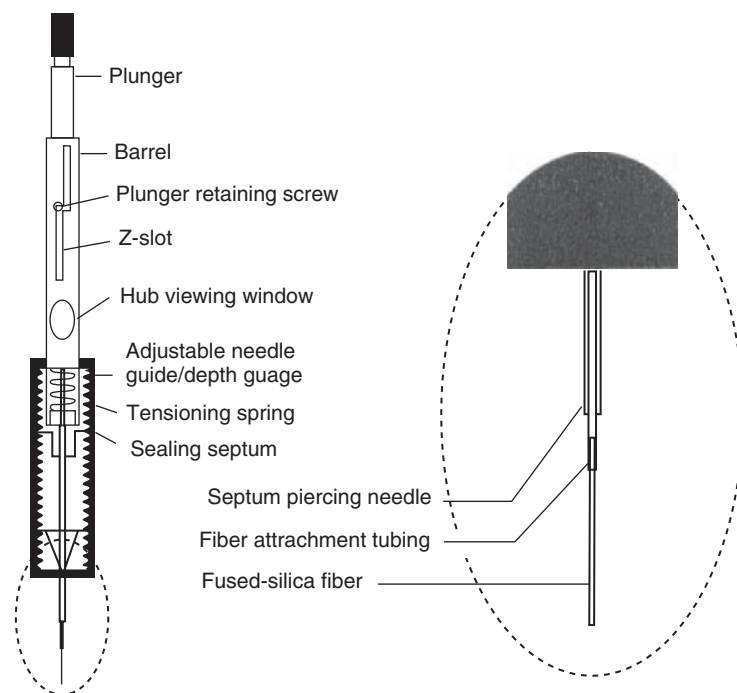
Figure 1   A commercial SPME device. (*Source*: Reprinted with permission from Zhang *et al*. (1994). Solid-phase microextraction. Anal Chem **66**: 844A–854A. Copyright 1994 American Chemical Society.)

and detection by gas chromatography (GC) or high performance liquid chromatography (HPLC or LC) with mass spectrometry (MS) detection and provides linear results for a wide range of concentrations of analytes. By controlling the polarity and thickness of coating on the fiber, maintaining consistent sampling time, and adjusting several other extraction parameters, a researcher can obtain highly consistent, quantifiable results from low concentrations, sub-femtogram range, of analytes with shorter analysis time compared to conventional extraction techniques such as liquid–liquid extraction (LLE), solid-phase extraction (SPE), headspace (HS) analysis, striping, and purge and trap methods. These conventional procedures typically require longer time, complicated equipment, and/or extravagant use of organic solvents. A comparison of SPME and conventional techniques is presented in this section.

1.1   Fundamental Theory of SPME

The basic concept of SPME may have been coined from the idea of an immersed GC capillary column.

The SPME apparatus is a very simple device. A commercial version of SPME device is shown in Figure 1. Its appearance like modified syringe consisting of a fiber holder and a fiber assembly, the latter containing a ~1- to 2-cm-long retractable SPME fiber. The SPME fiber itself is a thin modified fused-silica optical fiber, coated with a thin polymer film (such as polydimethylsiloxane (PDMS)) that is used to isolate and concentrate solute into a range of coating materials, conventionally used as a coating material in chromatography (Zhang *et al*., 1994; Pawliszyn, 2000).

The principle behind the SPME is the partitioning of analytes between the sample matrix and the extraction medium. The capacity of SPME depends on the extraction of analytes from a sample into the SPME adsorptive layer. Numerous SPME fibers and coatings are commercially available and have been used in a wide range of applications and instrumentations (Kataoka, 2002, 2005, 2010; Ouyang and Pawliszyn, 2006). Some of them are presented in Table 1. After a sampling period – the extraction is considered to be complete when it reaches equilibrium. A finite time span is required to reach solute equilibrium between

Table 1 Commercially available SPME fibers, polarity, and operating temperatures with the class of compounds to be analyzed by the analytical tools.

Fiber coating	Film thickness (μm)	Polarity	Maximum operating temperature (°C)	Compatible with technique	Class of compounds to be analyzed
Polydimethylsiloxane (PDMS)	100	Nonpolar	280	GC/HPLC	Volatiles
PDMS	30	Nonpolar	280	GC/HPLC	Nonpolar semivolatiles
PDMS	7	Nonpolar	280	GC/HPLC	Medium to nonpolar semivolatiles
PDMS–divinylbenzene (DVB)	65	Bipolar	270	GC	Polar volatiles
PDMS–DVB	60	Bipolar	270	HPLC	General purposes
PDMS–DVB[a]	65	Bipolar	270	GC	Polar volatiles
Polypyrrole coated	50	Polar	250	HPLC–MS	Catechins and caffeine
Polyacrylate (PA)	85	Polar	320	GC/HPLC	Polar semivolatiles (phenols)
Carboxen–PDMS	75	Bipolar	320	GC	Gases and volatiles
Carboxen–PDMS[a]	85	Bipolar	320	GC	Gases and volatiles
Carbowax–DVB	65	Polar	265	GC	Polar analytes (alcohols)
Carbowax–DVB[a]	70	Polar	265	GC	Polar analytes (alcohols)
Carbowax-templated resin (TPR)	50	Polar	240	HPLC	Surfactants
DVB-PDMS-Carboxen[a]	50/30	Bipolar	270	GC	Odors and flavors
Supel-Q-PLOT	50	Polar	240	HPLC-UV	Isoflavones and phenols

[a] Stableflex type is on a 2-cm fiber.

a sample and the SPME layer, and equilibrium ideally occurs before the extracted solutes are withdrawn for desorption into an instrument. The adsorbed solutes are transferred with the SPME layer into an inlet system that desorbs the solutes into a gas (for GC) or liquid (for LC) mobile phase. Success relies upon choosing conditions such that solutes favor the SPME adsorptive layer as much as possible in the presence of bulk sample and then are subsequently released as quickly and completely as possible for chromatographic analysis by changing the conditions to favor solute release from the adsorptive layer. Secondary trapping and release of desorbed solutes after SPME is sometimes necessary when desorption is too slow to permit full use of column resolving power. This trapping and release can be accomplished using a discrete thermal trapor with column stationary-phase trapping by injection into a GC column and subsequently temperature programming for solute elution.

It has been described earlier that the innovative sample preparation technique SPME needs equilibrium condition after extraction of anlaytes from complex sample mixtures. At equilibrium, the amount of solute extracted or adsorbed by a fiber coating, n_e, is expressed as (Pawliszyn, 1997; Hinshaw, 2003; Pawliszyn and Pedersen-Bjergaard, 2006)

$$n_e = \frac{K_{fs}V_fV_s}{V_s + K_{fs}V_f}C_0 \qquad (1)$$

Equation (1) shows the relationship between the analyte concentration in the sample and the amount extracted by the coated fiber. Where K_{fs} is the partition coefficient for analyte between coating and sample matrix, V_f the volume of the fiber coating, V_s the sample volume, and C_0 the initial concentration of the analyte in the sample.

If the amount of solute extracted onto the fiber is an insignificant portion of that present in the sample, then the product of K_{fs} and V_f is essentially negligible compared to V_s, and Equation (1) becomes

$$n_e = K_{fs}V_fC_0 \qquad (2)$$

Equation (2) indicates that the amount of extracted analyte is independent of the volume of the sample. This means that (i) there is no need to collect a defined amount of sample before analysis, (ii) the fiber can be exposed directly to whatever is being analyzed, and (iii) the amount of extracted analyte will correspond directly to its concentration in the sample matrix. Thus, this allows for the prevention of errors associated with the loss of analyte through decomposition or absorption/adsorption/spills onto sample container walls.

It is noted that the partition coefficients are also influenced by the sample polarity/ionic strength, temperature, and other factors: both between the liquid sample and the HS and between the SPME coatings and the liquid. The influencing parameters should be kept under careful control for good sample-to-sample

consistency. Adding salt to an aqueous sample will often shift the partition coefficients for nonpolar solutes in favor of the SPME layer and decrease the time required to attain equilibrium. The chemistry of the sorptive SPME layer plays a significant role in enhancing or discriminating against classes of compounds. For the most part, SPME coatings for GC column absorb solutes in a manner related to their behavior to GC stationary phases. Polar SPME coatings such as those that contain polyesters or acrylates will enhance polar constituents and discriminate against nonpolar materials. Adsorptive layers with active carbon constituents will retain volatile components more strongly than layers made of nonpolar dimethyl silicones (Hinshaw, 2003). However, analysts should give some thought to desorption because a very strongly held solute might be too difficult to pry off the SPME layer for analysis.

1.2 Basic Operation Procedure of SPME

The simplicity of use, fastness, effectiveness, and solvent-free analyte concentration capability of SPME have drawn special attention to the scientists' community. It has been widely used to analyze target compounds from gaseous and liquid matrices using chromatography coupled to hyphenated detection techniques (Vas and Vekey, 2004; Kataoka, 2005; Zini *et al*., 2002). A simplified schematic diagram of SPME is illustrated in Figure 2 (Sigma-Aldrich Co., 1998). A ∼1- to 2-cm-long fused-silica fiber, coated with a polymer, is bonded to a stainless steel plunger and installed in a holder that looks like a modified microliter syringe (Figure 1). The plunger moves the fused-silica fiber into and out of a hollow needle. To use the unit, an analyst draws the fiber into the needle, passes the needle through the septum that
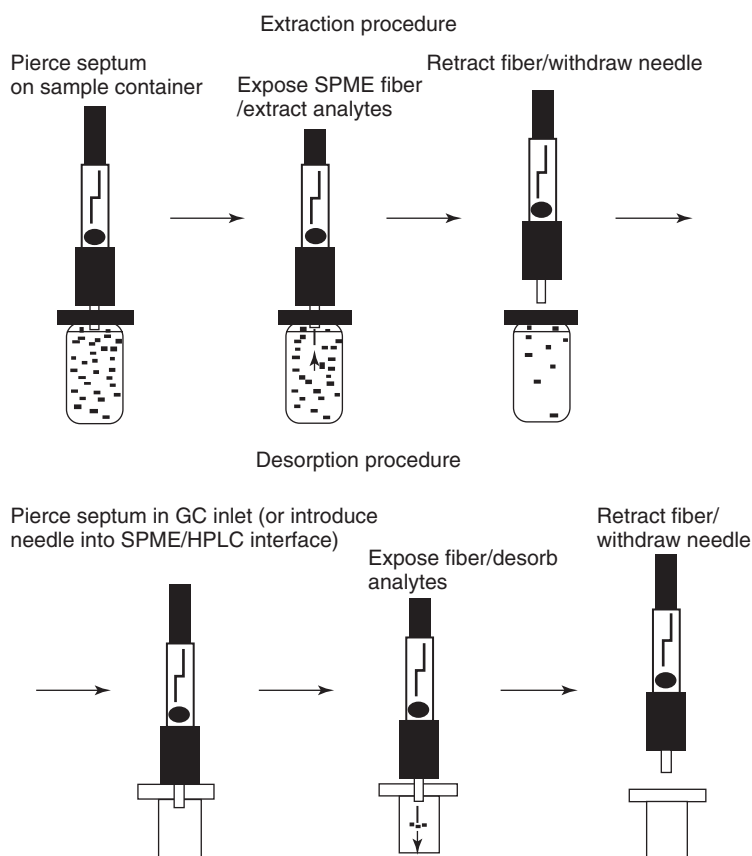
Figure 2   A schematic diagram of SPME adsorption and desorption procedures. (*Source*: Reproduced with permission from Supelco Bulletin, (1998), 923, 1–5. © Sigma-Aldrich Co.)

seals the sample vial, and depresses the plunger to expose the fiber in the liquid sample or HS above the sample. Organic analytes are then adsorbed to the coating on the fiber. After adsorption equilibrium is attained, which can be typically between 2 and 30 min, the fiber is drawn into the needle, and the needle is withdrawn from the sample vial. Finally, the needle is introduced into the GC injector or the SPME/HPLC interface, where the adsorbed analytes are thermally desorbed and delivered to the instrument column. An SPME/HPLC interface allows the technique to be combined with analysis by HPLC, expanding the applications for the extraction technique to detect surfactants in water, pharmaceuticals in biological fluids, chemicals or alkaloids in plant extracts, and many other areas (Kataoka, 2002, 2005, 2010; Togunde *et al.*, 2012; Zhu, 2006).

Selection of the appropriate coating is the first step in SPME method development. The coating type is chosen based on chemical properties such as polarities of the analytes. Different coatings are commercially available (Pragst, 2007; Ouyang and Pawliszyn, 2006; Kataoka, 2005, 2010; Wardencki *et al.*, 2007). The most widely used coating is PDMS. Although PDMS is considered nonpolar and is particularly suitable for the extraction of nonpolar compounds, it is often used for a wide range of screening applications in environmental analysis. Polyacrylate (PA) is a highly polar coating for general use and is ideal for the extraction of phenols. Other available coatings are carboxen/poly(dimethylsiloxane) (CAR/PDMS), carbowax/templated resin (CW/TPR), divinylbenzene/carboxen/poly(dimethylsiloxane) (DVB/CAR/PDMS), and carbowax–polyethylene glycol (PEG). CAR/PDMS is ideal for gaseous/volatile analytes; CW/DVB is suitable for the extraction of polar analytes, especially for alcohols; CW/TPR was developed for HPLC applications; and finally, DVB/CAR/PDMS is the newest coating, which is ideal for a broad range of analyte polarities (suitable for C2–C20 range) (Wardencki *et al.*, 2007). In SPME, equilibriums are established among the concentrations of an analyte in the sample and in the polymer coating on the fused-silica fiber. The amount of analyte adsorbed by the fiber depends on the thickness of the polymer coating and on the distribution constant for the analyte. Extraction time is determined by the length of time required to reach equilibrium for the analytes with the highest distribution constants. The distribution constant

generally increases with increasing molecular weight and boiling point of the analyte. Selectivity can be altered by changing the type of polymer coating on the fiber, or the coating thickness, to match the characteristics of the analytes of interest. In general, volatile solutes require a thick coating, whereas semivolatile compounds are most effectively adsorbed/desorbed with thin coatings. Examples of different fiber coatings together with film thickness, polarities, and compound suitability information have been summarized earlier (Table 1).

1.3 Comparison of SPME with Other Conventional Extraction Techniques

The major problem that keeps analysts from accepting a new analytical method is the expense in terms of both cost and training requirements. An ideal sample preparation technique should be solvent-free, simple, inexpensive, efficient, selective, and compatible with a wide range of analytical methods/instruments and applications. It should be capable of being used simultaneously in separating and concentrating the components and should allow on-site extraction and analysis. SPME is proximate to ideal sample preparation technique that can assimilate sampling, extraction, concentration, and sample introduction into a single step, resulting in high sample throughput.

Sample preparation methods using solvents such as LLE technique are time-consuming and laborious with multistep operations. When concentrating samples, an individual step can introduce errors and lose solutes during sample handling. Waste disposal of solvents is another issue that adds an extra cost to the analytical procedure and additional expense to the environment with health hazards to the laboratory associates. Using SPE cartridges or discs and micro-well plates has reduced many limitations of the conventional LLE method (Mottaleb and Abedin, 1999; Mottaleb *et al.*, 2003). SPE needs less solvent but it is a time-consuming multistep process and often requires a concentration step, which may result in a loss of volatile target compounds. Long sample preparation times are obviously disadvantageous and multistep procedures are prone to loss of analytes. Adsorption of solutes with trace impurities from extraction solvent on the walls of extraction devices can also simultaneously occur. Although SPE requires much less volume of organic

Table 2 Sample preparation steps in LLE, SPE, and SPME.

LLE	SPE	SPME
Addition of organic solvents to the sample	Conditioning of cartridges or membranes	Exposing SPME fiber to the sample
Agitation in a separatory funnel	Sample elution	Desorption of analytes in the analytical instrument
Separation of aqueous and organic phases	Solvent elution to remove interferences and analyte desorption	—
Removal of organic phase	Evaporation/concentration of the organic phase	—
Evaporation/concentration of the organic phase	Injection in the analytical instrument	—
Evaporation/concentration of the organic phase	—	—

solvents compared to LLE, it is still a substantial amount. Evaporation of the SPE extract is more time-consuming than in LLE because polar solvents such as methanol, acetonitrile, or combination of others are commonly used, which usually have lower vapor pressure than that of the nonpolar solvents such as hexane and dichloromethane, primarily used for LLE. Moreover, LLE and SPE are always performed off-line, as automation is rather complex. Automated systems are available, but these did not lead to a breakthrough in the economics of the sample preparation. The newer sample preparation technique, SPME, was aimed to overcome the limitations inherent to SPE and LLE. The SPME is a viable alternative for determining the concentration of target analytes such as pharmaceuticals, volatiles, semivolatiles, nonvolatiles, biological, and chemicals with the benefit of simplicity as it integrates sampling, extraction, concentration, and sample introduction into a single-step solvent-free operation. As the target solutes from the samples are directly extracted and concentrated to the extraction fiber, it improves the performance of the analytical method. Thus, this method saves preparation time and reduces significantly the use of potentially hazardous solvents and waste disposal

costs. Table 2 compares the main steps in the sample preparation among LLE, SPE, and SPME, before analysis.

The primary advantages of SPME devices are their ability to decouple sampling from matrix effects that would otherwise distort the apparent sample composition, their simplicity, and their reduced solvent consumption. These characteristics combine to make SPME an attractive alternative to the classic HS or thermal-desorption sampling, SPE, and classic LLE. A comparison of analytical performance together with cost and simplicity of different sample preparation methodologies is illustrated in Table 3.

1.4 Interfacing of SPME with GC and HPLC

Since introduction, SPME has attracted distinct attention for scientists because of merits over conventional sample handling techniques (Tables 2 and 3). To overcome some of the shortcomings associated with SPME, many groups researched and published a series of reviews and primary articles in literature (Pragst, 2007; Risticevic *et al*., 2009; Lee

Table 3 Comparison of performances of SPME and other conventional techniques.

Detection limit (MS)	Precision (RSD)	Expense	Time	Solvent used	Simplicity
Purge and Trap (ppb)	1–30	High	30 min	None	No
Stripping (ppt)	3–20	High	2 h	None	No
Headspace (ppm)		Low	30 min	None	Yes
Liquid–liquid extraction (ppt)	5–50	High	1 h	1000 mL	Yes
Solid-phase extraction (ppt)	7–15	Medium	30 min	To 100 mL	Yes
SPME (ppt)	<1–12	Low	5 min	None	Yes

Source: Reproduced with permission from Supelco Bulletin, (1998), **923**, 1–5. © Sigma-Aldrich Co.
Table provided by J. Pawliszyn, University of Waterloo, Ontario, Canada.
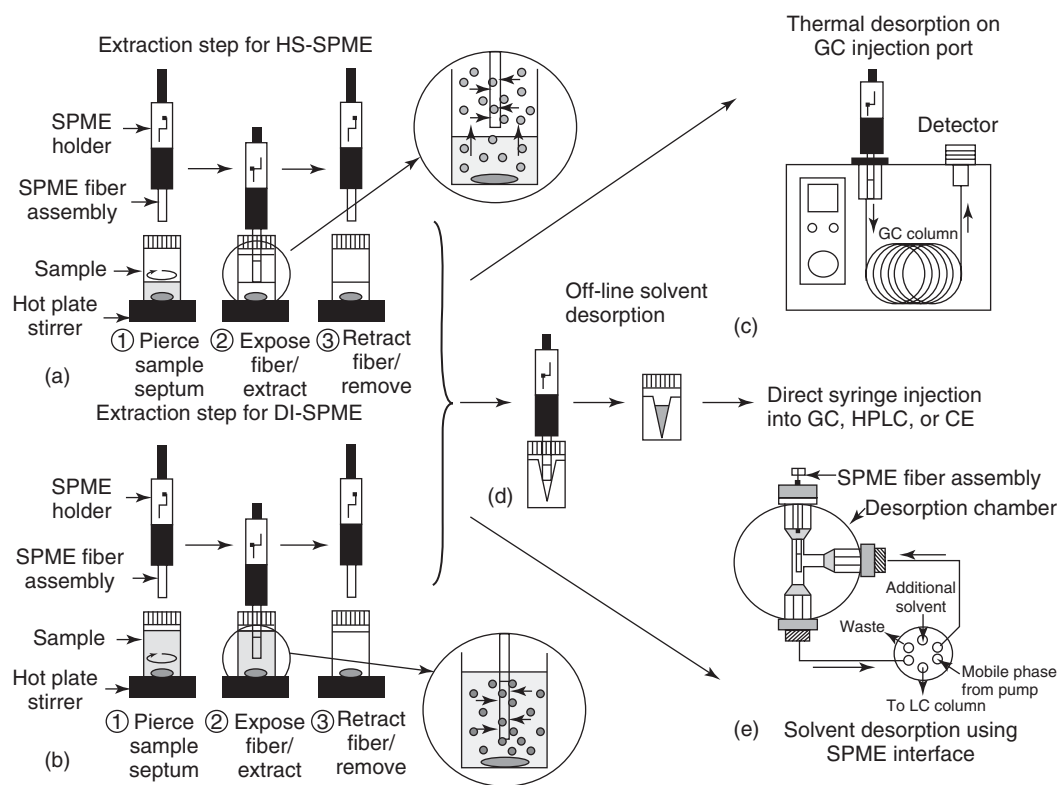
Figure 3 SPME process by HS, DI fiber SPME, and off-line solvent desorption for GC, HPLC, or CE analyses. (*Source*: Kataoka and Saito (2011). Reproduced from Elsevier.)

et al., 2007; Jiang et al., 2010). Nowadays, SPME is a well-established technique routinely used in combination with GC and GC/MS and successfully applied to a wide variety of compounds, including volatile and semivolatile compounds from environmental, biological, and food samples (Ouyang and Pawliszyn, 2006; Koziel *et al*. 2000; Schipilliti *et al*., 2011; Cervera *et al*., 2011; Harizanis *et al*., 2008). SPME has also been used for direct coupling with HPLC to analyze weakly volatile or thermally labile compounds, not amenable by GC or GC/MS (Kataoka, 2002, 2005). A schematic assembly of SPME that extracts analytes from the HS, direct immersion (DI), and off-line solvent desorption for GC, HPLC, or capillary electrophoresis (CE) systems is shown in Figure 3.

An SPME/HPLC interface equipped with a special desorption chamber and switching valve was developed to couple the fiber SPME and HPLC (Figure 3). It is utilized for solvent desorption before liquid chromatographic separation instead of thermal desorption in the injection port for GC applications. In the SPME/HPLC interface, an open-tubular fused-silica capillary column (in-tube SPME) is used instead of the SPME fiber. The in-tube SPME is suitable for automation and can continuously perform extraction, desorption, and injection using a standard autosampler. With the in-tube SPME technique, organic compounds in aqueous samples are directly extracted from the samples into the internally coated stationary phase of a capillary column and then desorbed by introducing a moving stream of mobile phase or static desorption solvent when the analytes are more strongly adsorbed to the capillary coating. Detailed methodologies and applications of SPME/HPLC and SPME/HPLC/MS have been described elsewhere (Zhou *et al*., 2008; Togunde *et al*., 2012; Cudjoe and Pawliszyn, 2012; Zhang *et al*., 2011). A schematic representation of the automated SPME/HPLC/MS system is illustrated in Figure 4.
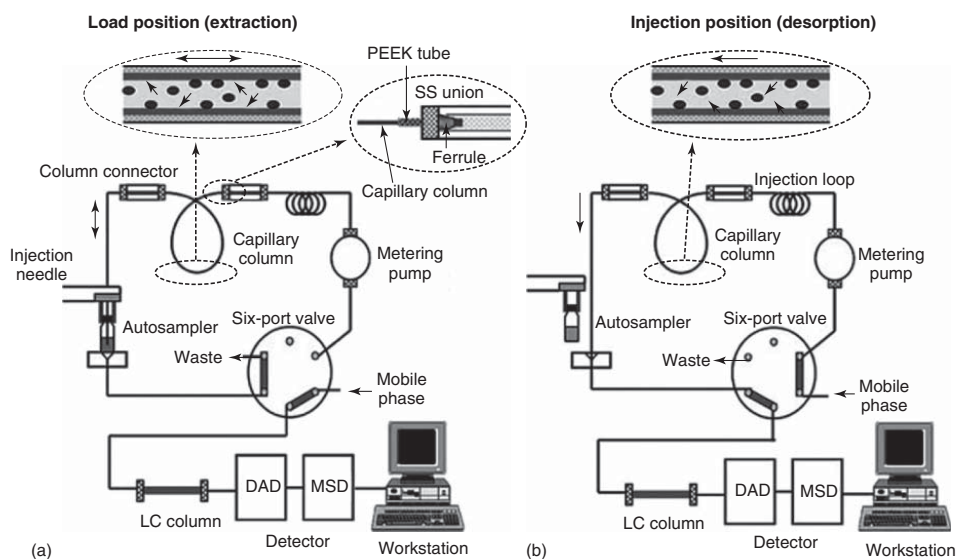
Figure 4   Schematic diagrams of an in-tube SPME/LC/MS system. (a) Load position (extraction) and (b) injection position (desorption). DAD is the photodiode array detector and MSD the mass detector. (*Source*: Kataoka *et al*. (2009c). Reproduced from Elsevier.)

1.5   Comparison of Fiber SPME and In-Tube SPME for HPLC or HPLC/MS

While the principles and theories of both fiber and in-tube SPME approaches are similar, a significant difference between these techniques is that the extraction of analytes is performed on the outer surface of the fiber for fiber SPME by agitation and on the inner surface of the capillary column for in-tube SPME by flow of the sample solution. The transfer of the target solutes in the extraction process of

the fiber and in-tube SPME is displayed in Figure 5 and the basic feature of each procedure is summarized in Table 4. It is necessary to prevent plugging of the capillary column and flow lines during extraction with the in-tube SPME, and particles must be removed from samples by filtration before extraction. For the fiber SPME, it is not necessary to remove particles before extraction, because they can be removed by washing the fiber with water before insertion into the desorption chamber of the SPME–HPLC interface. However, the fibers should be carefully handled,
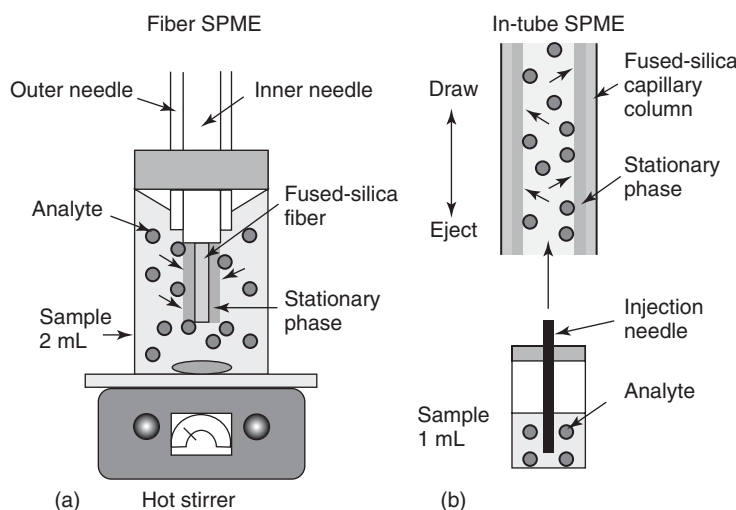
Figure 5   Extraction of analytes by fiber SPME (a) and in-tube SPME (b). (*Source*: Reproduced from Anal Bioanal Chem, **373**, (2002), 31–45, Automated sample preparation using in-tube solid-phase microextraction and its application – a review, Kataoka H., with kind permission from Springer Science and Business Media.

Table 4   Comparison of fiber and in-tube SPME techniques for combination with HPLC or LC–MS.

	Fiber SPME	In-tube SPME
SPME device	Commercially available SPME fibers are limited	Commercial capillary columns with a vast array of stationary phases are available
Field	Outer surface of fiber	Inner surface of capillary column
Extraction	Immerse fiber in sample solution under agitation	Repeatedly draw and eject sample solution into capillary column
Equilibration time	30–60 min (depending on the compounds)	10–15 min (depending on the compounds)
Desorption	Expose fiber in desorption chamber filled with mobile phase or additional solvent	Draw desorption solvent or mobile phase into capillary column
Carryover	~10% (depending on the compounds)	Negligible
Applicable sample	Clear and cloudy samples	Clear sample only
Operation	Manual	Automatic
Precaution	Fiber must be carefully handled because the coating is prone to strip off from the needle during insertion and removal from desorption chamber	Sample solution must be miscible with mobile phase and not contain insoluble matters because the flow-line is prone to stop

Source: Reproduced from Anal Bioanal Chem, **373**, (2002), 31–45, Automated sample preparation using in-tube solid-phase microextraction and its application – a review, Kataoka H., with kind permission from Springer Science and Business Media.

because they are fragile and can be easily broken, and the fiber coating can be damaged during insertion and agitation. Furthermore, high molecular weight compounds such as proteins can be irreversibly adsorbed by the fiber, thus changing the properties of the stationary phase and rendering it unusable. On the other hand, open-tubular GC capillary columns are very stable and useful for in-tube SPME coupled with HPLC or LC–MS. Another important difference between in-tube SPME and manual fiber SPME–HPLC is presumably decoupling of desorption and injection with in-tube SPME. In fiber SPME, analytes are desorbed during injection as the mobile phase passes over the fiber. In in-tube SPME, analytes are desorbed either by mobile phase flow or by aspirating desorption solvent from a second vial, which is then transferred to the HPLC column by mobile phase flow. The fiber SPME–HPLC method also has the advantage of eliminating the solvent front peak from the chromatogram, but peak broadening is sometimes observed because analytes can be slowly desorbed from the fiber. With in-tube SPME, peak broadening is comparatively small, because analytes are completely desorbed before injection. The coating thickness is only on the order of $0.1 \, \mu M$ with fast desorption. If analytes are sufficiently solvated by the mobile phase, then there is no need to use additional solvent for desorption. On the other hand, in-tube SPME can be easily automated because extraction and subsequent desorption can be continuously carried out without detaching the capillary column. Furthermore, carryover in in-tube SPME is diminished or eliminated in comparison with fiber SPME.

The extractions of analytes in fiber and in-tube SPMEs are based on the distribution coefficient between sample solution phase and SPME stationary phase. Thus, the time in which the analyte reaches distribution equilibrium between two phases becomes the extraction time. Although SPME has a maximum sensitivity at the partition equilibrium, a proportional relationship is obtained between the amount of analyte extracted by SPME and its initial concentration in the sample matrix before reaching partition equilibrium. Therefore, full equilibration is not necessary for quantitative analysis by SPME. Generally, SPME is affected by various factors such as fiber coating, pH of the sample solution, salt level, warming, and agitation and in-tube SPME is influenced by the type of capillary coating, pH of the sample solution, length of the capillary column, draw/eject volume of the sample, and their cycles

and speeds. The applications of fiber and in-tube SPMEs to environmental, food, biological, and clinical samples are well documented in literature (Guan *et al.*, 1998; Kataoka, 2002, 2010; Zhang *et al.*, 2011; Degel, 1996; Gallardo *et al.*, 2006).

Overall, either SPME or in-tube SPME is proven to be efficient sample preparation techniques capable of coupling with GC or GC–MS, HPLC, or LC–MS for analyzing a wide range of samples (gaseous to biological macromolecules) with shortened analysis times providing reproducible and robust data. These techniques have shown good analytical performance combined with simplicity and low cost and produced relatively clean and concentrated extracts.

1.6 Application of SPME

SPME technique has been regarded as the most attractive for the pretreatment of complex samples before chromatographic processes because it enables rapid analysis at low operating costs with no environmental pollution. This application includes preparing samples for analysis of volatile, semivolatile, and nonvolatile compounds and biological macromolecules from environmental composite mixtures. SPME methodologies are also frequently employed to extract fragrances and biochemical and biological samples before analysis. The advanced SPME coupled with GC–MS or HPLC–MS can effectively be used to analyze antioxidants, pharmaceuticals, personal care products, metabolites, flavors, and aroma from natural products, food plants, and organic compounds in media containing nanoparticles (Mester *et al.*, 2001; Masuck *et al.*, 2010; Zhu, 2006; Augusto and Valente, 2002; Chen *et al.*, 1998; Yang and Peppard, 1994). More applications of SPME with capillary electrophoresis were also reported (Liu and Pawliszyn, 2006). This section will explore the applications of SPME in the analysis of natural products and organic compounds in media containing potentially sorbing nanoparticles.

1.6.1 Analysis of Natural Products

The plant and fruit extracts that contain molecules with antioxidant and antibacterial properties, which are beneficial for human health, have increasingly been used in the past years because of their safety compared to artificial products. In general, plant

extracts are complex mixtures composed of a variety of chemicals/biochemical constituents. Some of them have medicinal values, whereas some may be potentially harmful to human health or living organisms. Thus, chemical composition and characterization of medicinal plants and their extracts are very important.

Employing SPME/GC–MS, Piccirillo *et al.* (2013) identified the composition and the antibacterial properties of the extracts from the stems and leaves of native Portuguese Ginja cherry (*Prunus cerasus* L., Rosaceae). These cherries are used to make liquor (Ginjinha); both stems and leaves are a by-product of this process. Thirty-six volatile compounds were identified from the extracts; terpenes – both hydrocarbons and oxygenated – were the main class of compounds. The extracts made from the stems with ethyl acetate were particularly rich in linalool, α-pinene, 4-terpineol, and cedrene with concentrations >10 times higher than that in other extracts. A volume of 2 mL of each solution of the redissolved extracts were placed into a 10-mL flask with a gas-tight seal. The flask was then placed into a water bath at 40 °C. The volatiles were adsorbed on a DVB/CAR/PDMS-coated SPME fiber surface after 10 min of exposure in the flask. Afterward, the fiber was placed into the injector of a GC and left there

for 10 min – during which desorption of the chemical molecules was performed at 220 °C and directly transfer into the analytical column. The medicinal compounds identified from stems and leaves extracts by SPME/GC–MS are listed in Table 5.

Martendal *et al.* (2011) extracted and collected the volatile aroma fractions from plant matrices using HS–SPME and analyzed the collected extracts by injecting them into GC-flame ionization detection (FID) and GC–MS systems. They optimized the extraction time (10–60 min) and temperature (5–60 °C) using DVB/CAR/PDMS-coated SPME fiber using 100 mg of a mixture of plant matrices. The chromatogram was divided into four groups of peaks based on the elution temperature to provide a better understanding of the influence of the extraction parameters on the extraction efficiency considering compounds with different volatilities/polarities. The authors proposed a new SPME method and compared the obtained results with the optimized conventional methods based on a single extraction temperature (45 min of extraction at 50 °C). The proposed SPME method led to better results in all cases, considering both peak area and the number of identified peaks, and provided an excellent alternative to extract analytes with quite different volatilities. Comparisons
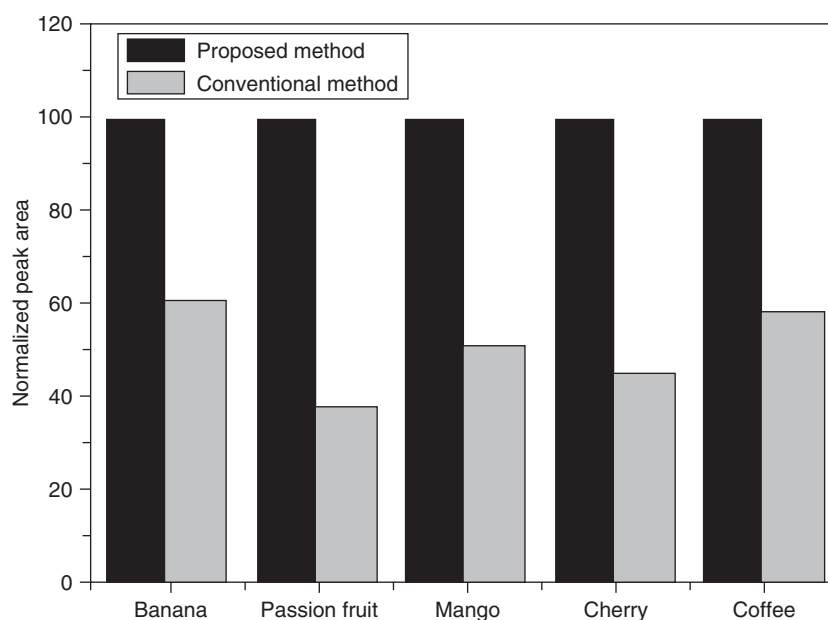
Figure 6 Comparison obtained between the SPME (proposed) and the conventional methods for the extraction of the volatile fraction of selected samples using the summed area of the compounds identified in the GC–MS analysis. (*Source*: Martendal *et al*. (2011). Reproduced from Elsevier.)

Table 5 List of the compounds identified in stems and leaves extract by SPME/GC–MS method: integrated peak areas (unit: counts).

Number	Compounds	Stems			Leaves		
		Ethyl acetate	Ethanol	Acetone	Ethyl acetate	Ethanol	Acetone
Terpenes hydrocarbons							
1	α-Pinene	1326	—	—	—	—	—
2	Camphene	1515	1014	—	360	527	—
3	3-Carene	587	—	—	396	—	—
4	β-Ocimene	1075	389	—	—	—	—
5	Aromadendrene	1182	158	208	—	187	—
Oxygenated terpenes							
6	5-Hepten-2-one, 6-methyl	20,187	1331	541	1244	952	194
7	Verbenyl ethyl ether	—	359	167	—	—	—
8	Lylac aldehydes A	—	1022	805	656	257	514
9	Lylac aldehydes B	—	—	—	657	353	—
10	Linalool	37,451	4949	2202	1167	2590	1519
11	4-Terpineol	1230	405	216	546	995	50
12	α-Terpineol	9162	1434	394	884	763	176
13	*trans*-β-Ionone	190	189	79	439	462	126
Aldehydes							
14	2-Octenal	787	—	—	—	—	—
15	Benzene acetaldehyde	—	1066	—	713	1052	—
Ketones							
16	5,9,9-Trimethyl-spiro[3.5]non-5-en-1-one	262	242	105	195	418	70
17	Ethanone-1,2-hydroxyphenyl	1507	—	—	—	—	—
Esters							
18	Benzyl acetate	3747	—	—	—	—	—
19	Methyl salicylate	9585	—	—	—	—	406
20	Diethyl phthalate	—	1693	525	1625	368	—
21	Propanoic acid, 3-chloro,4-formylphenyl ester	1544	—	58	104	165	1083
22	2(4*H*)-Benzofuranone, 5,6,7,7a-tetrahydro-4,4,7a-trimethyl	—	—	—	400	314	87
23	Acetic acid, 2-phenylethyl ester	808	—	—	—	—	32
24	Anisyl acetate	992	—	—	—	—	—
Alcohols							
25	5-Octen-2-yn-ol	427	—	—	—	—	—
26	Octanol, 2-methyl	3386	—	—	—	—	69
27	Phenyl ethyl alcohol	2107	444	244	225	378	490
28	Eugenol	2816	—	—	—	120	437
29	2-Decyn-1-ol	528	158	123	392	422	94
30	Benzene methanol α,α,α-dimethyl	785	166	58	127	124	—
Acids							
31	Acetic acid	3151	—	195	—	295	94
32	Spiro[2.2]pentane-1-carboxylic acid, 2-cyclopropyl-2-methyl	645	—	—	—	—	—
Hydrocarbons							
33	1,5,5-Trimethyl-6-methylene-cyclohexene	968	—	—	—	160	—
34	α-Caryophyllene	—	—	—	—	336	—
35	Cedrene	2021	695	315	1075	2793	107
Amino acids							
36	DL-Phenylalanine, *N*-[(phenylmethoxy) carbonyl]	4414	—	—	—	—	—

Source: Piccirillo *et al*. (2013). Reproduced from Elsevier.

between the proposed and other conventional methods are illustrated in Figure 6.

Guan *et al*. (1998) employed SPME to separate dinitroaniline herbicides from complicated biological matrices, human urine and blood, extending its application to biomedical analysis by GC-electron capture detection (ECD). In their work, urine or water spiked with herbicides and anhydrous sodium sulfate was preheated at 70 °C for 10 min, and a PDMS-coated SPME fiber was exposed to the HS at 70 °C for another 30 min. Spiked diluted blood was treated at 90 °C in the same way. The HS SPME

$R_1 = CH_3$ $R_2 = H$ $R_3 = H$

1. Ethalfluralin
$R_4 = CH_2\overset{\displaystyle CH_3}{\underset{\displaystyle |}{C}}=CH_2$
$R_5 = CH_2CH_3$

2. Benfluralin
$R_4 = CH_2CH_3$
$R_5 = CH_2CH_2CH_2CH_3$

Internal standard, profluralin
$R_4 = CH_2-\triangleleft$
$R_5 = CH_2CH_2CH_3$

3. Fluchloralin
$R_4 = CH_2CH_2CH_3$
$R_5 = CH_2CH_2Cl$

$R_1 = CF_3$ $R_2 = NH_2$ $R_3 = H$

4. Prodiamine
$R_4 = CH_2CH_2CH_3$
$R_5 = CH_2CH_2CH_3$

$R_1 = CH(CH_3)_2$ $R_2 = H$ $R_3 = H$

5. Isopropalin
$R_4 = CH_2CH_2CH_3$
$R_5 = CH_2CH_2CH_3$

$R_1 = CH_3$ $R_2 = H$ $R_3 = CH_3$

6. Pendimethalin
$R_4 = H$
$R_5 = CH(CH_2CH_2CH_3)_2$
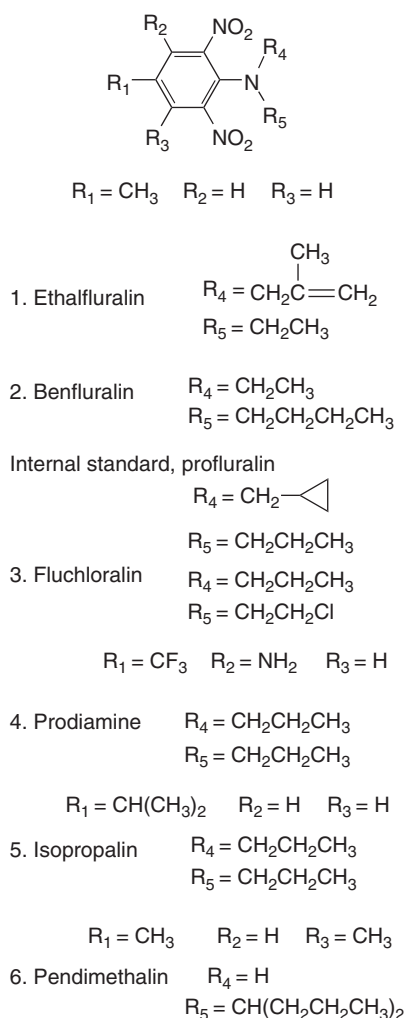
Scheme 1 Chemical structure of the herbicides. (*Source*: Guan *et al.* (1998). Reproduced from Elsevier.)

yielded clean extracts of dinitroaniline herbicides from urine, blood, or water, which could be directly analyzed by GC–ECD without further purification. The chemical structure of the herbicides is shown in Scheme 1. Authors used a Supelco SPME device with a 10-mm-long fiber, 100 μm in diameter, and coated with 100-μm-thick PDMS. This fiber was conditioned for 1 h in a GC injection port at 250 °C before sampling and used immediately to avoid contamination by inserting the SPME syringe in GC port. An aliquot volume of river water or urine was placed in a 4-mL glass vial with a silicone-septum cap, spiked with internal standard (I.S.) herbicides, and

mixed with anhydrous sodium sulfate. The vial was capped and placed in a Reacti-Therm heater at 70 °C. After 10 min, the needle of the SPME device pierced the septum of the vial, and the fiber was exposed for 30 min to the HS above the spiked urine or water and stirred by a small PTFE-coated bar. Finally, the fiber was retracted into the needle, pulled out from the vial, and immediately inserted into the GC injection port at 270 °C. It was exposed there for 5 min for complete desorption of analytes. The SPME of herbicides from human blood was conducted in the same way except that 0.5 mL of blood spiked with herbicides. Then, the blood was mixed with 0.5 mL of distilled water and was preheated at 90 °C for 10 min and HS-extracted at 90 °C for 30 min. Typical GC–ECD chromatograms of SPME that processed urine, blood, and water samples for herbicides are shown in Figure 7.

Gioti *et al.* (2007) assessed the antioxidant activity of plant extract by HS-SPME and GC–ECD. They monitored malondialdehyde (MDA), pentanal, and hexanal simultaneously as final products of lipid peroxidation using HS-SPME with on-fiber derivatization. The aldehyde compounds are extracted and subjected to on-sorbent derivatization into stable hydrazones with 2,4,6-trichlorophenyl-hydrazine (TCPH) for analysis. Aldehydes are polar compounds but their derivatization leads to a decrease in their polarity and therefore, the use of an apolar coating/fiber is feasible. The TCPH-containing PDMS represents a favorable phase that can accommodate the derivatized analytes and shows greater yield and reproducibility than the PDMS/Carbowax (75-μm thickness) fiber. A schematic of the reaction between carbonyls and the TCPH-derivatizing agent occurring on the SPME fiber is given in Scheme 2. The degree of inhibition of oxidation was performed by monitoring the chlorinated hydrazones after thermal desorption. The procedure was employed to evaluate *in vitro* the antioxidant activity of *Hypericum perforatum* L. extracts and vitamin E following induction of oxidation of sunflower oil, as a model lipid system. Before the measurement of antioxidant activity, the optimal process conditions such as HS volume, temperature, agitation, extraction/derivatization time, and desorption time and temperature were properly established. Aqueous extracts of *H. perforatum* L. exhibited the highest antioxidative effect. The method was shown to be promising for screening purposes for antioxidant substances and natural extracts. A typical
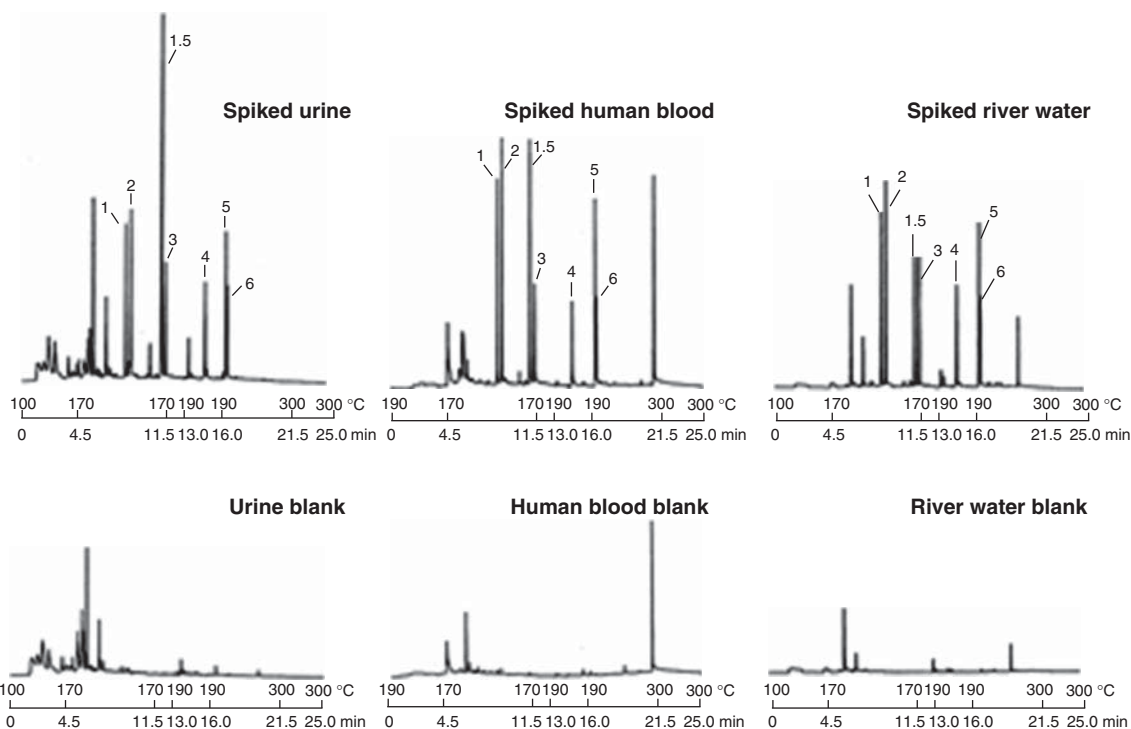
Figure 7   GC–ECD chromatograms of the herbicides in SPME extracted from human urine and blood and river water. Urine (1.0 mL) spiked with 1.0 ng mL^{-1} of each herbicide and 10 ng mL^{-1} of I.S. was extracted at 70 °C for 30 min, whereas human blood (0.50 mL, diluted with 0.50 mL of water) spiked with 6.0 ng of each herbicide and 10 ng of I.S. was extracted at 90 °C for 30 min. River water (1.0 mL) spiked with 1.0 ng mL^{-1} of each herbicide and I.S. was extracted at 70 °C for 30 min. Peak identities are ethalfluralin (1), benfluralin (2), profluralin (I.S.), fluchloralin (3), prodiamine (4), isopropalin (5), and pendimethalin (6). (*Source*: Guan *et al*. (1998). Reproduced from Elsevier.)
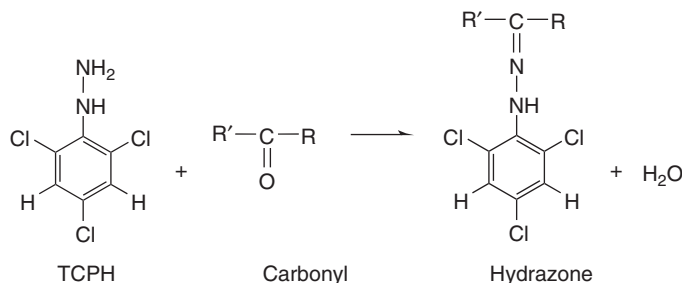
Scheme 2   A schematic of the reaction between the carbonyls and the TCPH-derivatizing agent occurring on the SPME fiber. (*Source*: Gioti *et al*. (2007). Reproduced from Elsevier.)

example of HS volume optimization is illustrated in Figure 8.

Gas chromatography-combustion-isotope ratio mass spectrometry (GC-C-IRMS) enhanced by enantio-selective measurements allowed to attain a more complete characterization of the fresh food and plant extracts. Many of the chemical compounds that contribute flavor are naturally occurring. Sometimes,

these are too unstable to be used in commercial flavorings because of the required shelf-life of the final industrial product. Authenticity assessment of flavored strawberry foods was performed using HS-SPME coupled to GC-C-IRMS by Schipilliti *et al.* (2011). A validity range was achieved by investigating the carbon isotope ratio of numerous selected aroma active volatile components.
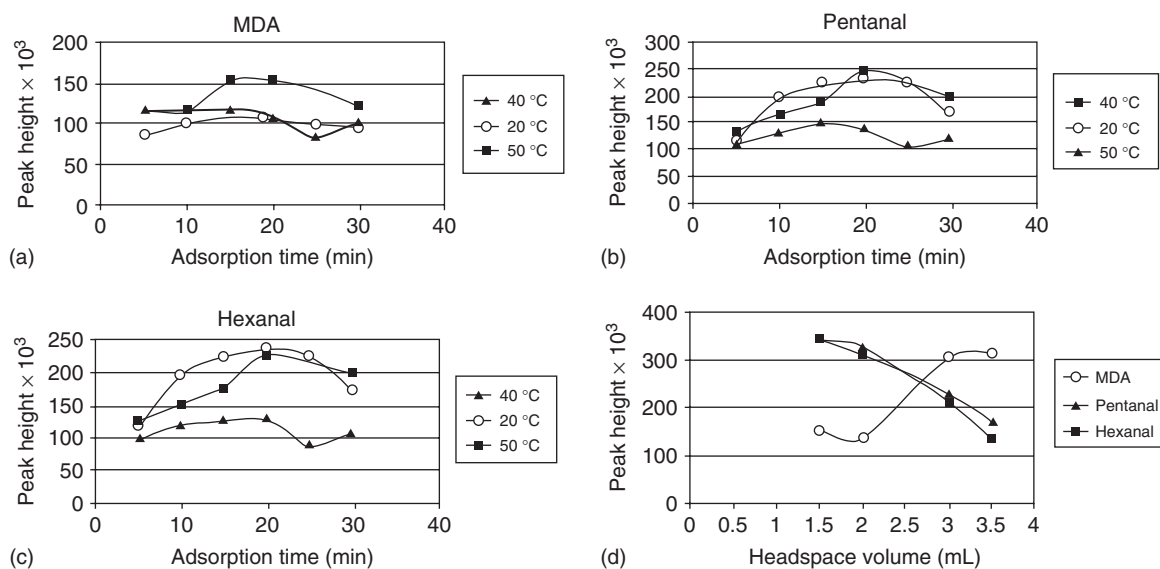
Figure 8 Adsorption and derivatization profiles for (a) MDA, (b) pentanal, (c) hexanal as a function of temperature and time, and (d) effect of headspace volume on extraction/derivatization. (*Source*: Gioti *et al.* (2007). Reproduced from Elsevier.)
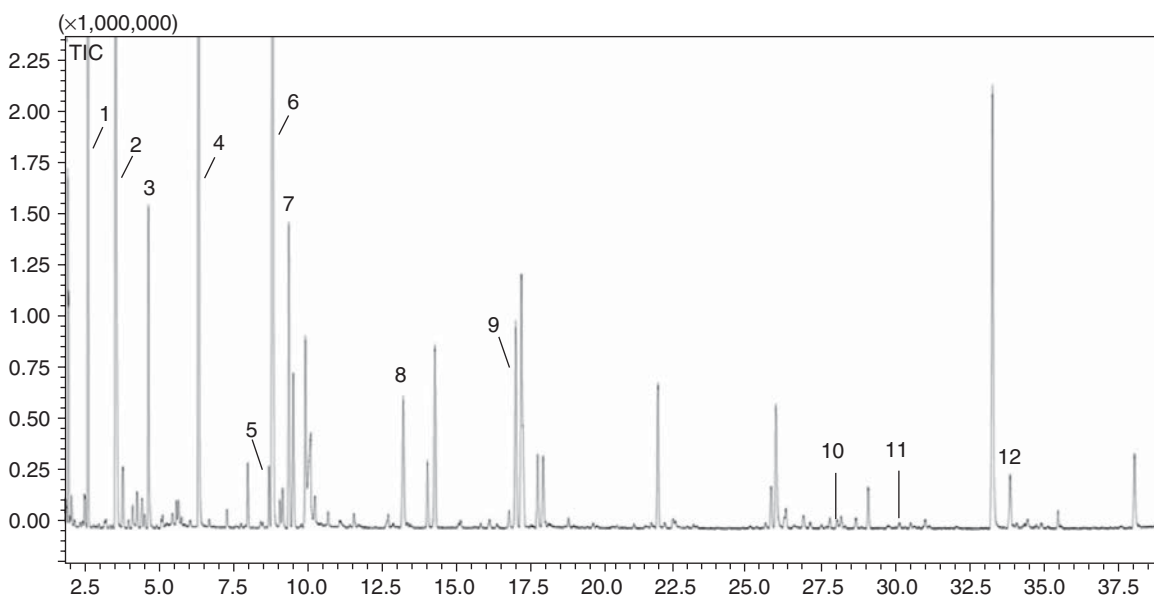
Figure 9 TIC gas chromatogram of the HS-SPME of organic strawberry. Peak identification – (1) methyl butanoate, (2) ethyl butanoate, (3) hex-(2*E*)-enal, (4) methyl hexanoate, (5) buthyl butanoate, (6) ethyl hexanoate, (7) hexyl acetate, (8) linalool, (9) hexyl butanoate, (10) octyl isovalerate, (11) γ-decalactone, and (12) octyl hexanoate. (*Source*: Schipilliti *et al.* (2011). Reproduced from Elsevier.)

These included methyl butanoate, ethyl butanoate, hex-(2*E*)-enal, methyl hexanoate, butyl butanoate, ethyl hexanoate, hexyl acetate, linalool, hexyl butanoate, octylisovalerate, γ-decalactone, and octyl hexanoate in organic Italian fresh strawberries, Kenyan pineapple, and fresh Italian peach extracts. A typical separation of GC chromatogram of HS-SPME of organic strawberry extract is shown in Figure 9.

Salvia plants and their essential oils have immense economic values in food, pharmaceuticals, and cosmetic products. Ozek *et al*. (2010) did a comprehensive study using four different isolation techniques: conventional hydrodistillation (HD), microwave-assisted hydrodistillation (MWHD), microdistillation (MD), and micro-steam distillation-solid-phase microextraction (MSD-SPME) for identification and composition analysis of volatile chemical constituents from the aerial parts of *Salvia rosifolia* Sm. by GC-FID and GC–MS. In the case of MSD-SPME, the dried and ground plant material (1.0 g) was placed in 10-mL water inside a 25-mL flask. The flask was fitted with a Claisen distillation head with a plug and a condenser set up for refluxing rather than distillation. Heating was achieved using electric heater, and threaded plug was used for SPME fiber assembly. A manual SPME holder and the 65-μm, "blue-type" PDMS–DVB fiber were used for SPME procedure of volatiles. Fiber was conditioned at 250 °C for 30 min before the experiment. After the SPME needle pierced the plug, the fiber was exposed to the HS above a plant sample. MSD-SPME procedure was carried out at the boiling temperature of water. The suitable equilibrium time was 1 min between loading of SPME fiber into flask and starting of the extraction. The extraction times of 3, 1, and 0.5 min were tested to ensure the complete composition of the volatiles in comparison with conventional HD. The minimum extraction time giving the most suitable results was estimated as 0.5 min. After trapping of the volatile, the loaded SPME fiber was withdrawn into the needle, and then the needle was removed from the plug and subsequently used for GC-FID and GC–MS analyses. Desorption of analytes from the fiber coating was performed by heating the fiber in the injection port to 250 °C for 5 min. The analytes were then transferred directly into the chromatographic column for analysis. Afterward, the SPME fiber was reconditioned at 250 °C for the next extraction experiment for 30 min. The fiber was subjected to a blank injection to ensure fiber integrity and the absence of any analytes after each reconditioning period. Table 6 represents the oil/chemical composition of *S. rosifolia*, endemic to Turkey. The list of detected compounds with their relative percentages composition, retention indices, and percentages of compound classes are given in Table 6 in the order of elution on the HP-Innowax FSC column. All the techniques applied to *S. rosifolia* resulted in the volatiles being richer in monoterpene alkaloids (Table 6). Typical chromatographic profiles are also provided in Figure 10. Other abundant alkaloids such as nicotine, nornicotine, anabasine, and anatabine in tobacco plant have been investigated by employing SPME combined with GC–MS (Wu *et al*., 2002). The limit of detection for nicotine, nornicotine, anabasine, and anatabine was estimated as 1.80, 0.134, 0.475, and 0.209 μg, respectively. Their recoveries ($n = 3$) were determined to be 108%, 100%, 103%, and 108%, with relative standard deviations of 6.7%, 9.9%, 17.6%, and 20.6%, respectively. Tobacco samples collected from the United States, China, India, Indonesia, Brazil, Pakistan, the Russian Federation, Japan, Bangladesh, Nigeria, Mexico, Germany, Egypt, and Kenya were analyzed. Alkaloid levels detected in various tobacco brands collected from these countries ranged from 8.1 to 22.1 mg g^{-1} for nicotine, 0.36 to 1.50 mg g^{-1} for nornicotine, 0.51 to 0.16 mg g^{-1} for anabasine, and 0.13 to 1.12 mg g^{-1} for anatabine. Additional applications in detection and analysis of alkaloid natural products were also reported elsewhere (Tzanetou and Kasiotis, 2013; Czerwinski *et al*., 1996; Tellez *et al*., 2004).

Lord *et al*. (2004) confirmed the uses of SPME to monitor herbicide levels in living tomato, reeds (class: Juncaceae; *Juncus effusus*), and onion plants at various locations within the plants. Subsequent to extraction, toxic chemicals were quantified by LC–MS/MS. The ability to monitor selected forms of an herbicide in near real time as they move through an *in vivo* living plant provided a faster and more accurate means of assessing both herbicide mode of action and the plant's physiological response to atrazine, simazine, and prometryn herbicides. Table 7 reveals the concentration of pesticides in different heights of tomato stems. In this study, plants were watered regularly and exposed to indirect sunlight for approximately 8 h per day and maintained at ambient temperature (18–22 °C). Pesticides were applied to the root area in water over the 21 days to avoid direct contact between the plant stalks and the pesticide solution. The total mass of each pesticide applied to the plants was 20 mg for reed and onion and 50 mg for tomato. Before analysis, SPME fibers were monitored for absence of pesticide residues by performing a fiber injection, and then preconditioned in phosphate buffered saline (PBS) for 30 min. For analysis, a small 1.5-cm-long hole was prepared in the plant material and filled with PBS, by means of a 22 G hypodermic needle mounted on a buffer

Table 6 The composition of *Salvia rosifolia* volatiles obtained by hydrodistillation (HD), microwave-assisted hydrodistillation (MWHD), microdistillation (MD), and microsteam distillation-solid phase microextraction (MSD-SPME) techniques.

No.	RRI[a]	RRI[b]	Compound	Composition (%)[c]				ID method
				HD	MWHD	MSD-SPME	MD	
1	966		2-Ethyl furan	t	—	—	—	d, e
2	1014	1014	Tricyclene	0.2 ± 0.05	0.1 ± 0.0	0.1 ± 0.0	0.1 ± 0.0	d, e
3	1032	1032	α-Pinene	24.3 ± 1.7	21.4 ± 0.2	15.7 ± 0.7	34.8 ± 0.3	d, e
4	1035	1035	α-Thujene	0.5 ± 0.1	0.7 ± 0.1	0.8 ± 0.1	1.3 ± 0.1	d, e
5	1076	1085	Camphene	3.1 ± 0.4	2.5 ± 0.0	1.5 ± 0.1	2.7 ± 0.0	d, e
6	1118	1118	β-Pinene	6.9 ± 0.9	8.8 ± 0.1	6.8 ± 1.1	13.6 ± 0.0	d, e
7	1132	1132	Sabinene	1.5 ± 0.3	2.1 ± 0.0	1.5 ± 0.3	2.0 ± 0.0	d, e
8	1136	1117	Thuja-2,4(10)-diene	0.1 ± 0.0	t	t	0.1 ± 0.0	d, e
9	1174	1156	Myrcene	1.1 ± 0.0	1.3 ± 0.0	1.6 ± 0.1	1.1 ± 0.0	d, e
10	1183	1183	Pseudolimonene	t	t	t	t	d, e
11	1188	1179	α-Terpinene	t	0.1 ± 0.0	0.1 ± 0.0	0.2 ± 0.1	d, e
12	1203	1205	Limonene	2.7 ± 0.6	3.9 ± 0.1	4.5 ± 0.7	2.9 ± 0.0	d, e
13	1213	1210	1,8-Cineole	16.6 ± 2.0	19.2 ± 0.0	25.1 ± 0.6	18.4 ± 0.1	d, e
14	1225		(Z)-3-Hexenal	t	0.1 ± ± 0.0	0.8 ± 0.2	0.4 ± 0.0	d, e
15	1244	1237	2-Pentyl furan	t	t	—	—	d, e
16	1246	1230	(Z)-β-Ocimene	t	t	t	t	d, e
17	1255	1256	γ-Terpinene	2.3 ± ± 0.3	2.6 ± 0.0	3.3 ± 0.3	1.7 ± 0.0	d, e
18	1266	1252	(E)-β-Ocimene	t	0.1 ± 0.0	0.2 ± 0.0	0.1 ± 0.0	d, e
19	1269		5-Methyl-3-heptanone	t	0.1 ± 0.0	0.2 ± 0.0	0.1 ± 0.0	d, e
20	1280	1279	p-Cymene	6.7 ± 0.2	3.1 ± 0.1	4.3 ± 0.4	1.8 ± 0.0	d, e
21	1290	1283	Terpinolene	t	0.1 ± 0.0	0.1 ± 0.0	0.1 ± 0.0	d, e
22	1304	1305	1-Octen-3-one	t	t	t	0.1 ± 0.0	d, e
23	1391	1393	(Z)-3-Hexenol	t	-	0.1 ± 0.0	0.1 ± 0.0	d, e
24	1393	1368	3-Octanol	0.1 ± 0.0	0.1 ± 0.0	0.2 ± 0.0	0.1 ± 0.0	d, e
25	1439		γ-Campholene aldehyde	t	t	t	0.1 ± 0.0	d, e
26	1452	1459	α,p-Dimethylstyrene	0.2 ± 0.0	0.2 ± 0.0	0.2 ± 0.0	0.1 ± 0.0	d, e
27	1453	1445	1-Octen-3-ol	0.2 ± 0.0	0.1 ± 0.0	0.9 ± 0.1	0.3 ± 0.0	d, e
28	1466	1458	α-Cubebene	0.3 ± 0.0	0.4 ± 0.0	1.0 ± 0.0	0.2 ± 0.0	d, e
29	1474	1474	trans-Sabinene hydrate	0.1 ± 0.0	0.4 ± 0.0	0.1 ± 0.0	0.2 ± 0.0	d, e
30	1477		4,8-Epoxyterpinolene	t	t	0.1 ± 0.0	t	d, e
31	1497	1497	α-Copaene	0.9 ± 0.1	0.5 ± 0.0	t	0.1 ± 0.0	d, e
32	1528		α-Bourbonene	t	t	t	t	d, e
33	1532	1532	Camphor	3.9 ± 0.6	4.7 ± 0.2	6.5 ± 0.1	3.0 ± 0.0	d, e
34	1535	1535	β-Bourbonene	0.6 ± 0.1	t	1.6 ± 0.2	0.4 ± 0.0	d, e
35	1542		4(15),5-Muuroladiene	t	t	1.4 ± 0.4	0.3 ± 0.0	f
36	1553	1556	Linalool	0.2 ± 0.0	0.2 ± 0.0	0.3 ± 0.0	0.2 ± 0.0	d, e
37	1556	1556	cis-Sabinene hydrate	0.1 ± 0.0	0.1 ± 0.0	0.2 ± 0.0	0.1 ± 0.0	d, e
38	1562	1565	Octanol	0.2 ± 0.0	0.8 ± 0.0	0.8 ± 0.0	0.2 ± 0.0	d, e
39	1586	1585	Pinocarvone	0.9 ± 0.5	0.2 ± 0.0	0.1 ± 0.0	0.1 ± 0.0	d, e
40	1589		Aristolene	0.3 ± 0.1	0.1 ± 0.0	0.2 ± 0.0	0.1 ± 0.0	d, e
41	1590	1571	Bornyl acetate	0.1 ± 0.0	0.1 ± 0.0	0.1 ± 0.0	0.1 ± 0.0	d, e
42	1600	1594	β-Elemene	t	t	t	t	d, e
43	1610		Calarene (β-gurjunene)	0.8 ± 0.3	0.5 ± 0.0	t	0.3 ± 0.0	d, e
44	1611	1616	Terpinen-4-ol	t	0.1 ± 0.0	1.7 ± 0.2	0.4 ± 0.1	d, e
45	1612	1604	β-Caryophyllene	5.1 ± 0.1	2.1 ± 0.0	4.9 ± 0.2	1.4 ± 0.0	d, e
46	1628	1653	Aromadendrene	0.1 ± 0.0	0.1 ± 0.0	0.1 ± 0.0	t	d, e
47	1648	1645	Myrtenal	t	01 ± 0.0	0.1 ± 0.0	0.1 ± 0.0	d, e
48	1668	1632	(Z)-β-Farnesene	0.9 ± 0.4	0.5 ± 0.0	0.4 ± 0.2	0.4 ± 0.0	d, e
49	1683	1680	trans-Verbenol	0.1 ± 0.1	t	0.4 ± 0.1	0.2 ± 0.0	d, e
50	1687	1687	α-Humulene	1.3 ± 0.4	0.7 ± 0.0	0.9 ± 0.2	0.7 ± 0.1	d, e
51	1704	1695	γ-Muurolene	0.1 ± 0.0	t	0.1 ± 0.0	t	d, e
52	1706	1706	α-Terpineol	t	0.2 ± 0.0	0.2 ± 0.0	0.2 ± 0.0	d, e
53	1719	1705	Borneol	1.8 ± 0.0	2.9 ± 0.1	2.4 ± 0.2	1.4 ± 0.0	d, e

(*continued overleaf*)

Table 6   (*Continued*)

No.	RRI[a]	RRI[b]	Compound	Composition (%)[c]				ID method
				HD	MWHD	MSD-SPME	MD	
54	1722		Verbenone	t	t	0.1 ± 0.0	0.2 ± 0.0	d, e
55	1726	1716	Germacrene D	1.2 ± 0.5	0.9 ± 0.0	1.3 ± 0.3	0.7 ± 0.0	d, e
56	1737	1728	(Z,E)-α-Farnesene	0.1 ± 0.1	0.1 ± 0.0	0.1 ± 0.0	0.1 ± 0.0	d, e
57	1755	1742	Bicyclogermacrene	0.1 ± 0.0	0.2 ± 0.0	0.1 ± 0.0	0.1 ± 0.0	d, e
58	1758	1749	(E,E)-α-Farnesene	0.2 ± 0.0	0.1 ± 0.0	0.1 ± 0.0	t	d, e
59	1773	1764	δ-Cadinene	0.2 ± 0.1	0.2 ± 0.0	0.4 ± 0.0	0.2 ± 0.0	d, e
60	1776	1766	γ-Cadinene	0.1 ± 0.0	0.1 ± 0.0	0.2 ± 0.0	0.1 ± 0.0	d, e
61	1786	1781	ar-Curcumene	0.1 ± 0.0	t	0.1 ± 0.0	0.1 ± 0.0	d, e
62	1853	1827	cis-Calamenene	t	0.1 ± 0.0	0.3 ± 0.0	0.4 ± 0.0	d, e
63	1900		epi-Cubebol	t	t	0.1 ± 0.1	0.1 ± 0.0	d, e
64	2008	2008	Caryophyllene oxide	4.3 ± 0.2	4.5 ± 0.12	1.4 ± 0.1	1.5 ± 0.1	d, e
65	2037	2016	Salvial-4(14)-en-1-one	0.2 ± 0.0	0.2 ± 0.0	0.1 ± 0.0	0.1 ± 0.0	d, e
66	2071	2071	Humulene epoxide II	0.5 ± 0.1	0.5 ± 0.0	0.1 ± 0.0	0.5 ± 0.0	d, e
67	2144	2150	Spathulenol	1.1 ± 0.1	2.3 ± 0.1	0.2 ± 0.1	0.7 ± 0.0	d, e
68	2324	2324	Caryophylladienol II	0.3 ± 0.0	0.1 ± 0.0	0.1 ± 0.0	0.4 ± 0.0	d, e
69	2366	2396	Eudesma-4(15),7-dien-β-ol	1.7 ± 0.6	0.3 ± 0.0	t	0.2 ± 0.0	d, e
70	2392	2392	Caryophyllenol II	0.2 ± 0.1	0.8 ± 0.0	t	0.30 ± 0.0	d, e
			Total	93.7 ± 1.6	91.8 ± 0.5	95.7 ± 0.1	97.6 ± 0.3	—
			Monoterpene hydrocarbons	48	46.5	40.5	63	—
			Oxygenated monoterpenes	24	28.5	37.5	24	—
			Sesquiterpene hydrocarbons	12.4	6.2	12	5.6	—
			Oxygenated sesquiterpenes	8.4	9	1.5	3.8	—

[a] RRI, relative retention indices calculated against *n*-alkanes (C_9–C_{20}) on HP-Innowax column.
[b] RRI, relative retention indices on polar column reported in literature.
[c] %, calculated from flame ionization detector data.
t, trace (<0.1%); d, identification based on retention index of genuine compounds on the HP-Innowax column; e, identification on the basis of computer matching of the mass spectra from Baser, Adams, MassFinder, Wiley, and NIST libraries; f, correct isomer not identified; ID, identification method.
Source: Ozek *et al.* (2010). Reproduced from Elsevier.

Table 7   Total concentrations of pesticides in tomato stem as nanograms per gram of the plant, fresh weight.

Stem height (cm)	Attrazine	Simazine	Prometryn
8–10	31.6	—	2.8
6–8	103.1	—	5.5
4–6	176.5	150.5	3.0
2–4	1940.8	2122.7	43.5
0–2	3105.4	2275.7	513.3

—, not detected.
Source: Reproduced from Lord *et al.* (2004) with permission from The Royal Society of Chemistry.

filled syringe. The retracted SPME fiber assembly (without syringe or spring) was placed into the plant to the end of the prepared hole. The outer needle was retracted while holding the fiber support rod stationary to expose the fiber to the plant material. The process was reversed to remove the fiber. After removal, the fiber was rinsed briefly with a stream of bidistilled water to eliminate any plant material adhered to the outer surface of the fiber. The assembly was then introduced into the SPME desorption interface for pesticide desorption, analysis, and quantification using triple quadrupole operated in positive ionspray mode of LC–MS/MS. Equilibration time profiles were evaluated after extraction of buffer samples (3 mL) containing 20 ng mL^{-1} of pesticides. Simazine reached equilibrium at 2 h. Atrazine and prometryn did not reach equilibrium in the 3-h duration of the test. For practical purposes, it was desirable to use extraction and equilibrium times shorter than 1 h. More applications of SPME/LC–MS have been demonstrated in literature (Kataoka *et al.*, 2009a, 2009b; Miller *et al.*, 2010; Kataoka, 2010). A typical example for analysis of biological samples with in-tube SPME/LC–MS chromatogram is shown in Figure 11.

Flavonoids, polyphenolic compounds, are a large group of thousands of secondary plant metabolites, which can be separated into five subclasses: flavonols, flavones, anthocyanins, catechins, and
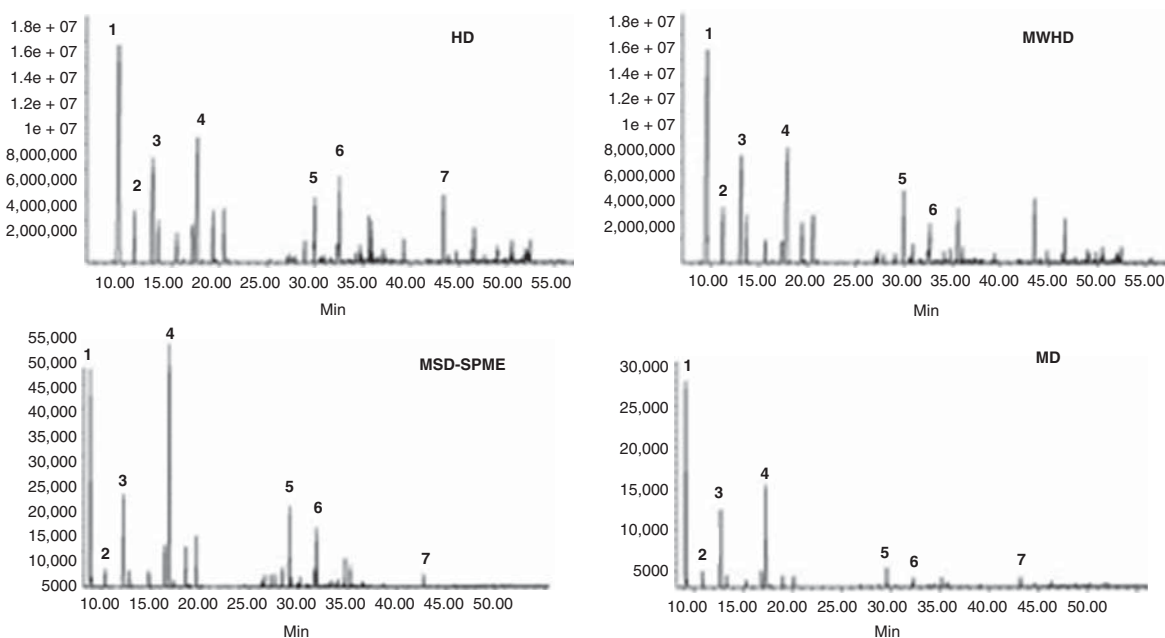
Figure 10 Typical chromatographic profiles of the oils of *Salvia rosifolia* obtained by different isolation techniques. HD, hydrodistillation; MWHD, microwave-assisted hydrodistillation; MSD-SPME, microsteam distillation-solid-phase microextraction; MD, microdistillation. Peaks identified are (1) α-pinene, (2) camphene, (3) β-pinene, (4) 1,8-cineole, (5) camphor, (6) β-caryophyllene, and (7) caryophyllene oxide. (*Source*: Ozek *et al*. (2010). Reproduced from Elsevier.)

flavonones (Wang and Huang, 2004). Flavonoids are widely used as remedies because of their spasmolytic, antiphlogistic, antiallergic, and diuretic properties (Cao *et al*., 1997). Flavonols (e.g., quercetin and myricetin) are a group of flavonoids that occur in foods or plant extracts as O- and C-glycosides. Kumar *et al*. (2009) determined myricetin and quercetin flavanols using SPME/HPLC-UV/VIS system from fruit and vegetable specimens. The method involves adsorption of flavonoids on carboxen/templated resin (CAR/TPR) fiber followed by desorption in the desorption chamber of SPME/HPLC interface using citrate buffer (0.001 M):acetonitrile (70 : 30) as mobile phase and UV detection at 372 nm. The detection limits for myricetin and quercetin were 48.3 and 24.7 pg mL^{-1}, respectively. The proposed method was validated by determining myricetin and quercetin in tomato, onion, grapes, and red wine samples. Mitani *et al*. (2003) has developed a method for determination of isoflavones, daidzein, and genistein in soybean plant by automated in-tube SPME that was coupled to the HPLC technique.

Daidzein, genistein, and their glucosides tested in this study were clearly separated within 8 min by HPLC using an XDB-C column with diode array detection. The glucosides daidzin and genistin were analyzed as aglycones after hydrolysis because the glucosides were not concentrated by in-tube SPME. The optimum extraction conditions for daidzein and genistein were obtained with 20 draw/eject cycles of 40 mL of sample using a Supel-Q porous layer open-tubular capillary column. The extracted compounds were easily desorbed from the capillary by mobile phase flow, and carryover was not observed. Using the in-tube SPME–HPLC method, the calibration curves of these compounds were linear in the range 5–200 ng mL^{-1}, with a correlation coefficient above 0.9999 ($n = 18$), and the detection limits (S/N = 3) were 0.4–0.5 ng mL^{-1}. This method was successfully applied to the analysis of soybean foods without interference peaks. The recoveries of aglycones and glucosides spiked into food samples were above 97%. Wu *et al*. (2000) focused on the analysis of five catechins (epigallocatechin, (+)-catechin, epicatechin, epigallocatechin gallate,
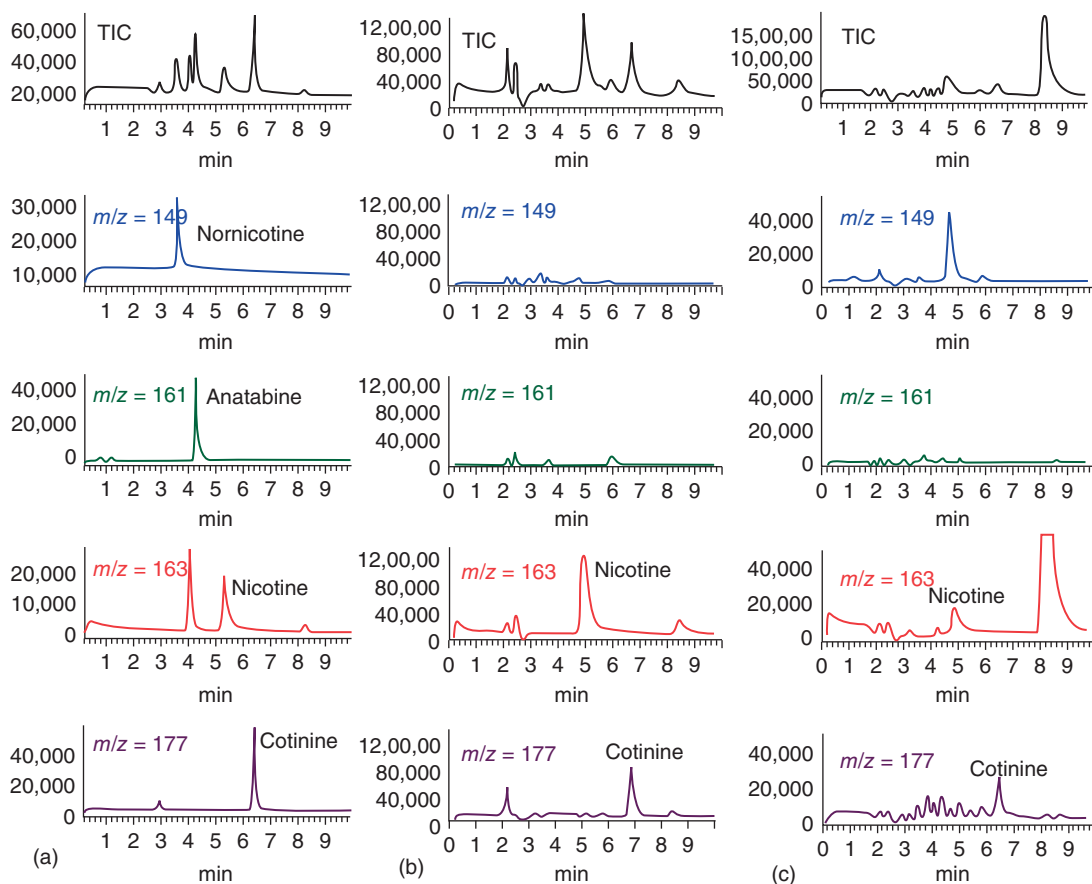
Figure 11 LC–MS chromatograms obtained from (a) 100 ng mL^{-1} standard compounds (direct injection), (b) urine (0.2 mL), and (c) saliva (0.2 mL) samples (in-tube SPME) after nicotine intake. In-tube SPME was carried out by 25 draw/eject cycles of 40 μL of sample solution at a flow rate of 150 μL min^{-1} using a CP-Pora PLOT amine capillary (60 cm × 0.32 mm inner diameter). The LC–MS conditions were as follows: LC column, Synergi 4u POLAR-RP 80A (150 mm × 4.6 mm inner diameter); mobile phase, 5 mM ammonium formate/methanol (55 : 45, v/v) and flow rate of 0.8 mL min^{-1}; ionization, electrospray ionization positive mode. (*Source*: Kataoka *et al*. (2009b). Reproduced from Elsevier.)

and epicatechin gallate) and caffeine in green tea using in-tube SPME–HPLC with electrospray ionization mass spectrometric (ESI-MS) detection. All samples have been separated and evaluated using a polypyrrole-coated capillary and several commercially available capillaries (capillary GC columns). Catechins were determined in both positive and negative ion detection modes. The detection limit (S/N = 3) for each of the studied catechins was <0.5 ng mL^{-1}. Caffeine was only determined under positive ES-MS detection modes and its detection limit was 0.01 ng mL^{-1}. Small amounts of catechins were also detected in grape juice and wine samples.

1.6.2 Analysis of Organic Compounds in Media Containing Nanoparticles

In recent investigations, the SPME was found to have a strong ability for dynamic speciation analysis of organic compounds in media containing potentially sorbing nanoparticles. Natural and engineered nanoparticles have recently attracted a lot of attention because of their unique properties and widespread applications. Particles in the nanosized range (1–100 nm) are ubiquitous in the environment and encompass both permeable and impermeable entities, for example, enzymes, protein molecules,

clays, and soot. However, their reactivities and fluxes in environmental compartments and consequent risk posed to biological systems are still largely unknown. In the aquatic environment, for example, natural and engineered nanoparticles are known to affect the binding, transport, and fate of vital and toxic compounds such as organics and heavy metals (Buffle and van Leeuwen, 1992; Hofmann and von der Kammer, 2008). Depending on the system, both enhancement and reduction in the mobility, bioavailability, and degradation kinetics of organics in the presence of nanoparticles have been reported. An extraction methodology such as SPME for speciation analysis of organic compounds in the presence of nanoparticles will allow establishment of a sound knowledge of compound speciation into predictions of their bioavailability. Town and coworkers (Benhabib *et al.*, 2009) applied SPME in the dynamic speciation analysis of the pesticide atrazine in an aqueous medium containing sorbing latex nanoparticles. For comparison, the atrazine was distributed over its freely dissolved and particle-bound forms. A PDMS fiber coating was applied as the solid phase for the extraction and was exposed to the sample solutions under mild stirring on a rock-and-roller shaker at ambient temperature. The overall rate of extraction of the analyte was faster in the presence of nanoparticles, which was explained on the basis of coupled diffusion of free and particle-bound atrazine toward the solid/sample solution interface. The enhanced diffusive flux for atrazine in dispersions of carboxylated latex particles demonstrated the presence of labile particle-bound atrazine species that do not enter the solid phase as such. In the eventual equilibrium, the total atrazine concentration in the solid phase was dictated by the solid phase/water partition coefficient (K_{sw}) and the concentration of the free atrazine in the sample solution. These observations demonstrated that the nanoparticles do not enter the solid phase. The results also pointed out to the important role of particle size in determining the contribution of a particle-bound organic compound to the steady-state diffusive transport flux: for colloids of radius >100 nm, the contribution might be negligible, whereas the contribution from an organic compound bound to smaller nanoparticles might be significant. The same group (Zielinska *et al.*, 2012) also recently explored the chemodynamic sorption of diclofenac by two model nanoparticles, silica (impermeable) and bovine serum albumin (BSA) (permeable) and their speciation analysis by

SPME using the same strategy. The study included aspects such as the impact of charged (unprotonated) diclofenac species and the lability of the nanoparticulate diclofenac complex. Diclofenac, a nonsteroidal anti-inflammatory drug prescribed for human and veterinary, is recalcitrant to standard wastewater treatment. Because diclofenac is a fairly polar molecule, less hydrophobic solid-phase polymer fibers, PA film and polyethylene glycol (PEG), were selected for the SPME analyses in the absence and presence of SiO_2 and BSA nanoparticles, respectively. Here too, the fibers were exposed to the sample solutions under mild stirring on a rock-and-roller shaker or magnetic bar stirring at ambient temperature. It was shown that only the protonated neutral form of diclofenac was accumulated in the solid phase, and hence, this specie governs the eventual partition equilibrium. The temporal profiles for the SPME accumulation of diclofenac in the absence and presence of SiO_2 nanoparticles are shown in Figure 12. These results demonstrated that the rate of the solid/water partition equilibration was enhanced in the presence of the sorbing nanoparticles of SiO_2 and BSA and these nanoparticles themselves do not enter the solid phase to any appreciable extent. The enhanced rate of attainment of equilibrium was attributed to a shuttle type of contribution from the nanoparticle species to the diffusive supply of diclofenac to the water/solid interface. The experimental accumulation rate constants were found to be in good agreement with the theoretical values confirming the dynamic equilibrium between the bound and the free diclofenac species in the bulk medium and labile behavior of particle-bound diclofenac on the effective timescale of diffusion toward the solid phase. Consequently, the rate-limiting step of the extraction process was determined by the coupled diffusion of the free and bound forms of the analyte in the aqueous phase. Similar conclusions were reached for atrazine in the presence of sorbing latex nanoparticles (Benhabib *et al.*, 2009).

1.7 Limitation of SPME

Since introduction, SPME has been used for a wide range of air and water monitoring applications and a very few limitations have been reported in the literature. The technique offers a simple and quick alternative to the traditional methods of sample handling techniques, but fused-silica-type column fiber
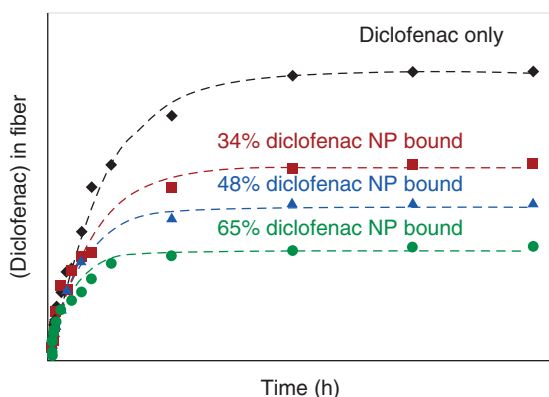
Figure 12 Mean concentration of diclofenac in the PA solid-phase film as a function of extraction time, *t*, in the absence and presence of different concentrations of SiO_2 nanoparticles at pH 5. (*Source*: Reprinted with permission from Zielinska K, van Leeuwen HP, Thibault S, Town RM. (2012). Speciation analysis of aqueous nanoparticulate diclofenac complexes by solid-phase microextraction. Langmuir **28**: 14672–14680. Copyright 2012 American Chemical Society.)

coatings sometimes restrict to analyze nonvolatile biological fluids. Murray (2001) demonstrated limitations of determination of volatile organic sulfur compounds produced by brassica plants to the use of Carboxen/PDMS fibers for their analysis. In general, the better performance of SPME fibers is also dependent on the polarities of the fiber coating materials and analytes. Moreover, parameters such as sampling time, temperatures, and equilibrium time influence the analysis. Polymer coatings on the fiber are fragile, easily broken, and have limited lifetime. Thus, careful handling of the device is important to obtain longer lifetime of the SPME fibers with reproducible and high throughput results.

CONCLUSION

SPME is an innovative technique for extraction and concentration of organic compounds from gaseous, aqueous, and solid matrices. It is rapid, simple, and ideal for automation and for *in situ* applications requiring no harmful solvents, as a result of significant reduction of analysis and waste disposal costs with minimal health hazard risk to users. The fundamental principle of SPME involves equilibration of the analytes between the sample matrix and an organic polymeric phase coated on a fused-silica

fiber. Coupling it to GC or GC–MS has documented sensitive, accurate, and reproducible quantitative analysis of different classes of volatile compounds. Both fiber and in-tube SPME interface are capable of coupling to HPLC, LC–MS, and CE and thus widen its range of applications to polar, nonpolar, nonvolatile, and thermally unstable compounds. The main advantages of in-tube SPME are its simplicity, rapidity, solvent elimination features, high sensitivity, small sample volume, lower cost, and simple automation. Because of the nature of selectivity and sensitivity, SPME has shown potential to analyze specimens ranging from drug screening, forensic toxicology, proteins, lipids, enzymes, and biologically and chemically valuable medicinal plant extracts. This technique has also demonstrated a strong ability for dynamic speciation analysis of natural products in media containing adsorbing nanoparticles.

REFERENCES

Arthur, C. L. and Pawliszyn, J. (1990) *Anal. Chem.*, **62**, 2145–2148.

Augusto, F. and Valente, A. L. P. (2002) *Trends Anal. Chem.*, **21**, 428–438.

Benhabib, K., Town, R. M. and van Leeuwen, H. P. (2009) *Langmuir*, **25**, 3381–3386.

Buffle, J and van Leeuwen, H. P., eds. (1992) Environmental Particles, Lewis Publishers, MI, vol. **I**, Chapters 10–13.

Cao, G., Sofic, E. and Prior, R. L. (1997) *Free Radic. Biol. Med.*, **22**(5), 749–760.

Cervera, M. I., Beltran, L. F. J. and Hernandez, F. (2011) *Anal. Chim. Acta*, **704**, 87–97.

Chen, W., Poon, K. F. and Lam, M. H. W. (1998) *Environ. Sci. Technol.*, **32**, 3816–3820.

Cudjoe, E. and Pawliszyn, J. (2012) *J. Chromatogr., A*, **1232**, 77–83.

Czerwinski, J., Zygmunt, B. and Namiesnik, J. (1996) *Fresenius J. Anal. Chem.*, **356**, 80–83.

Degel, F. (1996) *Clin. Biochem.*, **29**, 529–540.

Gallardo, E., Barroso, M., Margalho, C., *et al*. (2006) *J. Chromatogr., B*, **832**, 162–168.

Gioti, E. M., Fiamegos, Y. C., Skalkos, D. C., *et al*. (2007) *J. Chromatogr., A*, **1152**, 150–155.

Guan, F., Watanabe, K., Ishii, A., *et al*. (1998) *J. Chromatogr., B*, **714**, 205–213.

Harizanis, P. C., Alissandrakis, E., Tarantilis, P. A., *et al*. (2008) *Food Addit. Contam.*, **25**, 1272–1277.

Hinshaw, J. V. (2003) *LC-GC Europe*, December, 2–5.

Hofmann, T. and von der Kammer, F. (2008) *Environ. Pollut.*, **157**, 1117–1126.

Jiang, C., Sun, Y., Zhu, X., *et al*. (2010) *J. Sep. Sci.*, **33**, 2784–2790.

Kataoka, H. (2002) *Anal. Bioanal. Chem.*, **373**, 31–45.

Kataoka, H. (2005) *Curr. Pharm. Anal.*, **1**, 65–84.

Kataoka, H., Itano, M., Ishizaki, A., *et al*. (2009a) *J. Chromatogr., A*, **1216**, 3746–3750.

Kataoka, H., Inoue, R., Yagi, K., *et al*. (2009b) *J. Pharm. Biomed. Anal.*, **49**, 108–114.

Kataoka, H., Ishizaki, A., Nonaka, Y., *et al*. (2009c) *Anal. Chim. Acta*, **655**, 8–29.

Kataoka, H. (2010) *Anal. Bioanal. Chem.*, **396**, 339–364.

Kataoka, H. and Saito, K. (2011) *J. Pharm. Biomed. Anal.*, **54**, 926–950.

Koziel, J., Jia, M. and Pawliszyn, J. (2000) *Anal. Chem.*, **72**, 5178–5188.

Kumar, A., Malik, A. K. and Tewary, D. K. (2009) *Anal. Chim. Acta*, **631**, 177–181.

Lee, M. R., Chang, L. Y. and Dou, J. (2007) *Anal. Chim. Acta*, **582**, 19–23.

Liu, Z. and Pawliszyn, J. (2006) *J. Chromatogr. Sci.*, **44**, 266–373.

Lord, H. L., Moder, M., Popp, P., *et al*. (2004) *Analyst*, **129**, 107–108.

Martendal, E., de Souza Silveira, C. D., Nardini, G. S., *et al*. (2011) *J. Chromatogr., A*, **1218**, 3731–3736.

Masuck, I., Hutzler, C. and Luch, A. (2010) *J. Chromatogr., A*, **1217**, 3136–3143.

Mester, Z., Sturgeon, R. and Pawliszyn, J. (2001) *Spectrochim. Acta Part B: Atomic Spectrosc.*, **56**, 233–260.

Miller, E. I., Hye-Ryun, K., Norris, H. R. K., *et al*. (2010) *J. Chromatogr., B Analyt. Technol. Biomed. Life Sci.*, **878**, 725–737.

Mitani, V., Narimatsu, S. and Kataoka, H. (2003) *J. Chromatogr., A*, **986**, 169–177.

Mottaleb, M. A. and Abedin, M. Z. (1999) *Anal. Sci.*, **15**, 283–288.

Mottaleb, M. A., Abedin, M. Z. and Islam, M. S. (2003) *Anal. Sci.*, **19**, 1365–1369.

Murray, R. A. (2001) *Anal. Chem.*, **73**, 1646–1649.

Ouyang, G. and Pawliszyn, J. (2006) *Anal. Bioanal. Chem.*, **386**, 1059–1073.

Ozek, G., Demirci, F., Ozek, T., *et al*. (2010) *J. Chromatogr., A*, **1217**, 741–748.

Pawliszyn, J. (1997) *Solid-Phase Microextraction – Theory and Practice*. Wiley-VCH; New York, pp. 3–227.

Pawliszyn, J. (2000) *J. Chromatogr. Sci.*, **38**, 270–278.

Pawliszyn, J. (2003) *Anal. Chem.*, **75**, 2543–2558.

Pawliszyn, J. and Pedersen-Bjergaard, S. (2006) *J. Chromatogr. Sci.*, **44**, 291–306.

Piccirillo, C., Demiray, S., Silva-Ferreira, A. C., *et al*. (2013) *Ind. Crops Prod.*, **43**, 562–569.

Pragst, F. (2007) *Anal. Bioanal. Chem.*, **388**, 1393–1414.

Risticevic, S., Niri, V. H., Vuckovic, V., *et al*. (2009) *Anal. Bioanal. Chem.*, **393**, 781–795.

Schipilliti, L., Dugo, P., Bonaccorsi, I., *et al*. (2011) *J. Chromatogr., A*, **1218**, 7481–7486.

Smith, R. M. (2003) *J. Chromatogr., A*, **1000**, 3–27.

Sigma-Aldrich Co (1998) *Supelco Bull.*, **923**, 1–5.

Tellez, M. R., Khan, I. A., Schaneberg, B. T., *et al*. (2004) *J. Chromatogr., A*, **1025**, 51–56.

Togunde, O. P., Oakes, K. D., Servos, M. R. and Pawliszyn, J. (2012) Determination of pharmaceutical residues in fish bile by solid-phase microextraction couple with liquid chromatography-tandem mass spectrometry (LC–MS/MS). *Environ. Sci. Technol.* **46**: 5302−5309.

Tzanetou, E. N. and Kasiotis, K. M. (2013) *World J. Anal. Chem.*, **1**, 14–17.

Vas, G. and Vekey, K. (2004) *J. Mass Spectrom.*, **39**, 233–254.

Wang, S. P. and Huang, K. J. (2004) *J. Chromatogr., A*, **1032**, 273–279.

Wardencki, W., Curylo, J. and Namiesnik, J. (2007) *J. Biochem. Biophys. Meth.*, **70**, 275–288.

Wu, J., Xie, W. and Pawliszyn, J. (2000) *Analyst*, **125**, 2216–2222.

Wu, W., Ashley, D. L. and Watson, C. H. (2002) *Anal. Chem.*, **74**, 4878–4884.

Yang, X. and Peppard, T. (1994) *J. Agric. Food Chem.*, **42**, 1925–1930.

Zhang, X., Oakes, K. D., Luong, D., *et al*. (2011) *Anal. Chem.*, **83**, 6532–6538.

Zhang, Z. and Pawliszyn, J. (1993) *Anal. Chem.*, **65**, 1843–1852.

Zhang, Z., Yang, M. J. and Pawliszyn, J. (1994) *Anal. Chem.*, **66**, 844A–854A.

Zhou, S. N., Okes, K. D., Servos, M. R., *et al*. (2008) *Environ. Sci. Technol.*, **42**, 6073–6079.

Zhu, M. (2006) *Anal. Sci.*, **22**, 1249–1251.

Zielinska, K., van Leeuwen, H. P., Thibault, S., *et al*. (2012) *Langmuir*, **28**, 14672–14680.

Zini, C. A., Lord, H., Christensen, E., *et al*. (2002) *J. Chromatogr. Sci.*, **40**, 140–146.

Part Two

Instrumentation for Chemical Analysis

6

Microscopic Analysis

Johannes Saukel and Elisabeth Ginko

University of Vienna, Vienna, Austria

1 INTRODUCTION

"*Many plant constituents* (=botanicals) *are directly recognizable by means of the compound microscope, and when micro-chemical tests are employed in addition there are few plant products* (=botanicals) *that cannot be detected, even with more certainty than by chemical analysis*" (Jelliffe and Rusby, 1895). This quotation remains significant even today. The increasing use of botanicals especially from Asia and other foreign countries demands adequate methods for the proof of identification and purity in addition to the careful investigation of newly introduced drugs. In the first instance, the identification of herbal drugs is based on morphological and anatomical properties including some types of secondary chemical compounds of the source species. In pharmacognosy, microscopic analyses of biological material (medicinal and food plants or products of plants, fungi, lichen, or algae) have a long history. Since the middle of the nineteenth century, a range of books on this matter have been published providing the basis for microscopic identification of botanicals. Monographs providing detailed plant description, often accompanied by wonderful drawings, made to identify species, can be particularly found in books published between the end of the nineteenth century and World War II. Some hot spots of activity are Austria (Vienna, Graz, and Innsbruck (Mitlacher, 1904; Moeller, 1886, 1889, 1892; Schroff, 1853; Vogl, 1872, 1892; Wasicky *et al.*, 1936; Wiesner, 1900)), Germany (Göttingen (Wiggers, 1847), Berlin

(Berg, 1865), Marburg (Wigand, 1887), Munich (Solereder, 1908a, 1908b)), Switzerland (Bern (Flückiger, 1867; Tschirch and Oesterle, 1900), Zürich (Flück, Schlumpf, and Siegried, 1935)).

However, intensified usage of botanicals from other cultures (e.g., traditional Chinese medicine (Stöger, 2005) and Ayurveda medicine (Government of India, unknown), recently broadened our demand for monographs and basic education it is necessary to employ trained personnel trained personnel for the inspection of incoming botanical material. Since the early days, microscopic analyses have constituted the most rapid and cost-effective approach for identity and purity control of biological crude drugs, but only a trained pharmacognosist benefits from this technique. In Europe and the United States, nowadays, research in and teaching of morphology and anatomy of herbals are critically endangered because of other gainful research fields in pharmacy (e.g., molecular targets). In the United States, a remedial action for the conservation of the knowledge available in Europe resulted in the publication of "*American Herbal Pharmacopoeia*" (Upton *et al.*, 2011).

The essentials for the identification of botanical crude drugs by means of microscopic analyses are provided in the following section.

2 BASICS IN OPTICAL MICROSCOPY

Optical microscopy was and still is one of the most important promoters of life sciences. The past

decades have witnessed enormous efforts in applying this technique (e.g., phase contrast microscopy or confocal laser scanning microscopy). In addition, advances in digital imaging and analyses have opened new dimensions in documentation and measurements of various types of specimens. In the following sections, the interested reader gets information and guidance as to how to use an optical microscope and other necessary equipment built on our traditional knowledge accumulated over time. For more detailed information, one can either reach out to informative books that are available (Gabel, 1912; Mohl, 1846; Schacht, 1862) or find current information in the field of microscopic techniques on the Web sites of some providers of microscopes.

2.1   Short Introduction to Optics

Image formation in the light microscope – as well as in our eyes or in a magnifying glass – depends on the interactions of light with matter through six basic phenomena: reflection, absorption, refraction (inclusive double refraction), dispersion, interference, and diffraction (Figures 1, 2, and 4):

- *Reflection* occurs on all reflective surfaces depending on the angle of the oncoming light. Sometimes
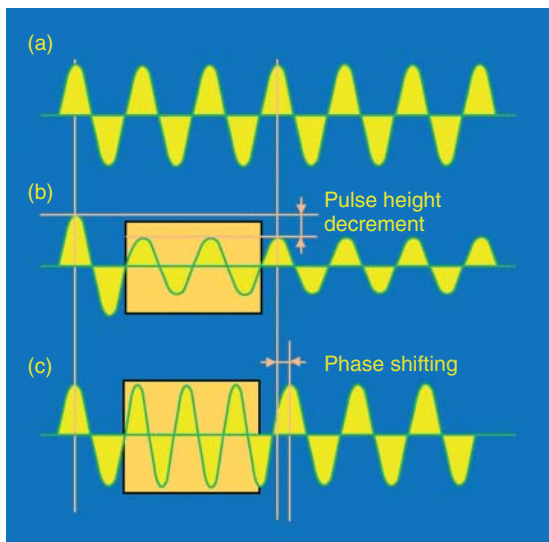
Figure 1   (a) Unaffected light; (b) light passing through an opaque object with a height decrement of the amplitude (*amplitude specimen*); and (c) light passing through an opaque object causing phase shifting (phase specimen).
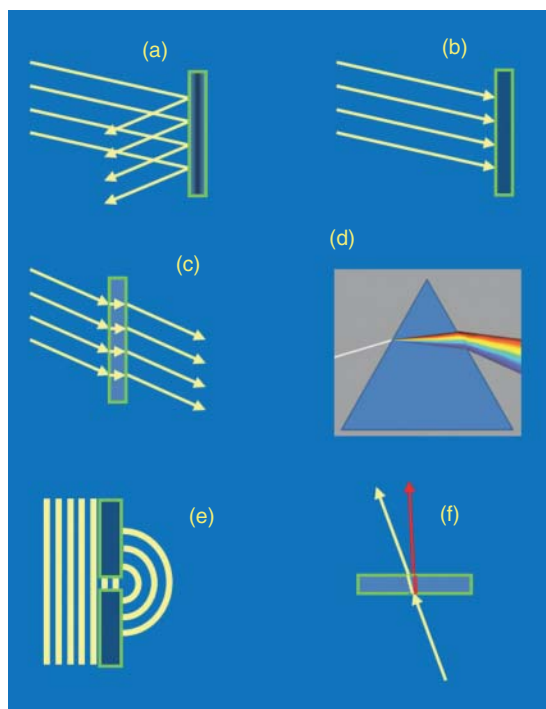
Figure 2   Basic principles in optics. (a) Reflection; (b) absorption; (c) refraction; (d) dispersion; (e) diffraction, and (f) double refraction.

reflection causes annoying aberrations in the microscopic light path.

- Qualitative and/or quantitative difference in *absorption* of light by the various structures of a specimen creates a contrast in terms of brightness or color in the *bright field mode* of a light microscope (see later), making individual details of the specimen visible. A specimen having such properties is called an *amplitude specimen* (Figure 1b). However, please note that each biological object represents both an *amplitude* and a *phase specimen* (Figure 1c) because light passing the object is refracted in various manner and therefore has different distances (refraction) and different matters to pass. For these reasons, a phase shifting occurs too. Normally, this phase shifting is imperceptible for our eyes but may be visualized by some techniques (e.g., *phase contrast* microscopy), and is used in cytology and cell biology for a so-called optical staining of fine structures (e.g., chromosomes).

- *Refraction* is the change in direction of a wave owing to a change of the medium, especially a
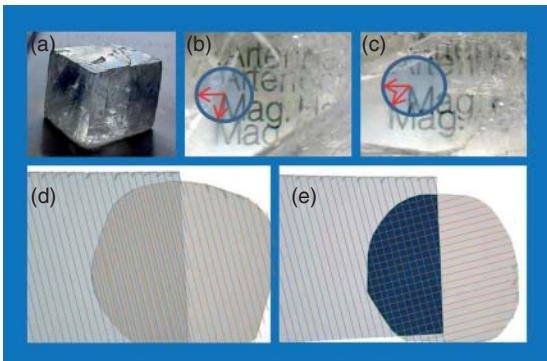
Figure 3 Birefringence. (a) Calc-spar crystal, (b) and (c) view through a calc-spar crystal with a double refracted printing, (b) and (c) after successive rotation; (d) and (e) polarizing filters, (d) filter in parallel orientation, and (e) filter in orthogonal or crossed orientation.

change of the velocity of the wave, when it is passing through. In optics, refraction is a phenomenon that occurs when waves travel from one medium to another with a different refraction index at an oblique angle. The *refraction index* of matter is a material property (n = *velocity of light in vacuum*/*velocity of light in the medium*). As known from the basics of physics, light is an electromagnetic wave. The white light (depending on the light source) is composed of different waves (red, blue, etc.). The refraction index of matter varies with the *wavelength* of the light passing through. Refraction is an important phenomenon for image formation by lenses. If we use a transparent birefringent material (e.g., gypsum) with a definitive thickness and therefore a definitive phase shift (e.g., green wavelength will be deleted by destructive interference [compare also the text later on]) in between the two polarizing filters, then the background of an image in the microscope appears in the so-called first-order purplish-red. Such a mounted gypsum crystal is often referred to as *first-order red plate* or *retardation plate*. The first-order retardation plate is standard equipment that is frequently utilized to detect birefringent structures, for example, fibers and trichomes in plant-derived specimens.

- *Dispersion* is the phenomenon in which the velocity of a light wave depends on its frequency. Dispersion causes the splitting of white light in different colors, for instance, in a full rainbow but leads to the annoying *chromatic aberration* (see later on) in lenses too. Blue light is refracted

to the greatest extent followed by green and red (Figure 2).

- *Interference* is a physical phenomenon in which two waves superimpose to form a resultant wave of heightened or lowered amplitude. This interaction of waves only occurs when the waves are correlated or coherent with each other, for example, light from the same source (see also linear polarized light).
- *Diffraction* occurs when light encounters an obstacle or while passing through small openings, for example, the specimen structures in the microscope. This phenomenon is described as the obvious bending of waves around small barriers and the spreading out of waves past small openings (Figures 2 and 4). Figure 4 shows a simple scheme of diffraction after passing an obstacle with two slits. Note: a real image shows much more complicated diffraction patterns, and for each wavelength, we observe different patterns. We can see circular wave crests and wave troughs appearing in the space downstream of the slits. These circular waves interfere with each other. Bright regions will appear as a result of constructive interference and dark regions by destructive interference. The two curves on the top shows the cross section of the light intensity of the 2D image (the so-called airy discs). The central peak symbolizes the undiffracted or zeroth-order light (= direct or undeviated light). In real images, this undiffracted light represents the background light. The peaks next to central peak are the diffracted parts of higher order. The smaller the wavelength is, the nearer the peaks are. This principle is the basic principle of image formation in a compound microscope too. The greater the number of diffracted parts gathered by the objective is, the better will be the information about the object and the resolution of the image!
- *Linear polarized* light and *birefringence* (also referred to as *double refraction*): Light emitted from a lamp vibrates in all direction orthogonal to the beam and is called *nonpolarized light*. If light shows vibration only in one single layer, it is called *polarized light*. Polarized light originates from reflection or from passing through birefringent materials such as crystals of calc-spar (Figure 3a) or a so-called polarizing filter (Figure 3d and e). If light of a normal lamp or sun light, both are nonpolarized light, passes through a birefringent material, then the light ray is divided into two linearly
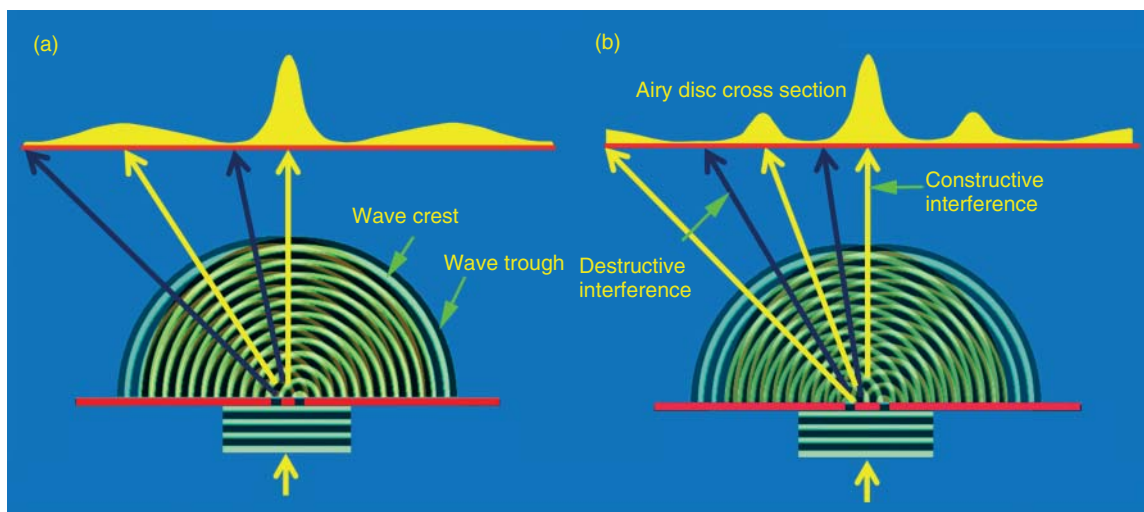
Figure 4 Light passing through a wall (in red) with two slits and gets scattered. The resulting waves interfere with each other giving a diffraction pattern. Constructive interference will appear where wave crests intersect and reinforce each other; destructive interference results where crests and troughs meet and neutralize each other. The two curves on the top shows the light intensity of the image (in 2D images referred to as airy discs). (a) and (d) shows different interference patterns because of different distance of the slits.

polarized light rays every single one with a vibration direction, that is, perpendicular to one another. From these two rays, the one that follows the law of refraction is called the *ordinary ray*, whereas the other is referred to as *extraordinary ray*. The speed and the index of refraction of the two rays differ in different birefringent materials. Therefore, both values serve as typical matter properties. Figure 3b and c demonstrates the effect of birefringence. The printing appears doubled. The ordinary ray creates an image of clear and dark appearance, and the extraordinary ray causes a second grayish image of shifted position. If one rotates the crystal, the position of the more indistinct letters rotates around the distinct letters. A material that refracts in such a way is termed anisotropic. Because of the different path and speed of the two rays, birefringence causes a phase shift (Figure 1c). The usage of such a phase shift in microscopy is shown later. The second way to produce polarized light is the usage of polarizing filters. Figure 3d and e shows pieces of polarizing film available in photo shops. These filters are often made from polyvinyl alcohol (PVA) plastic with an iodine doping. Mechanical elongation of the sheets during manufacture causes the PVA chains to align in one particular direction. Valence electrons from the iodine dopant are able

to move along the polymer chains, but not transverse to them. Because light is an electromagnetic wave with a magnetic and an electric component, only such waves with vibration directions perpendicularly to the chains are transmitted. This light has only one vibration direction and is called *linear polarized light*. In Figure 3d, the two polarizing filters are arranged in parallel (red and blue lines indicate the layer of the polarized light emitted from the filter) and the overlapping area appears only slightly darker. In contrast, Figure 3e shows the orthogonal position of the filters. In this case, the overlapping area appears dark – no light is passing through. This is the proper position of the polarizing units in a polarizing microscope (see also later on). If we use polarizing filters in this way, the background of an image in the microscope appears black. However, every birefringent structure in a specimen causes a rotation of light, and by vector addition, this structure becomes visible (more or less white on a black background). The pros and cons are discussed later. Now, let us come back to the earlier mentioned phase shifting.

Besides the physics of light, it is also necessary to know the basics about the function of lenses. Figure 5a shows the light path in a very thin lens. Image *magnification* and *position* depend on the *focal*
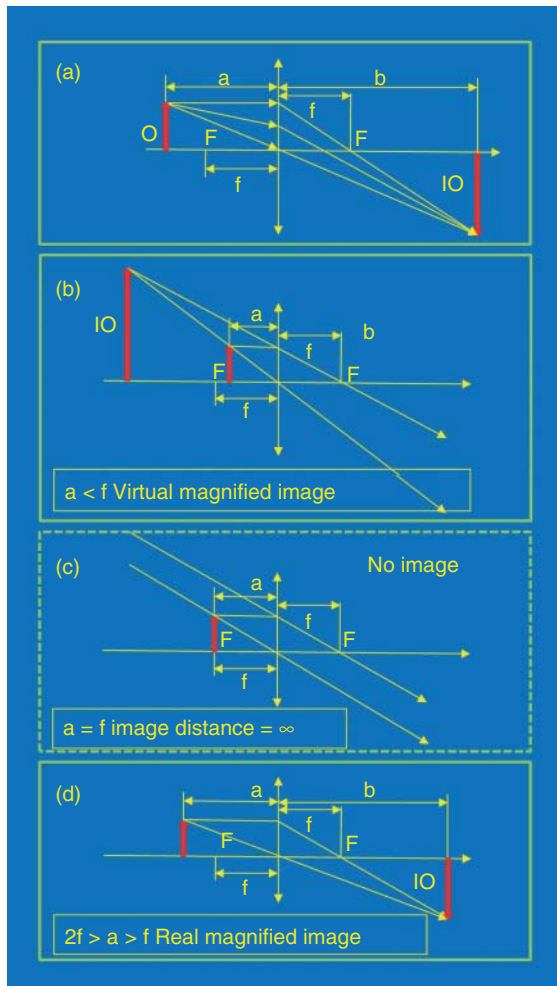
Figure 5 (a) Optical properties for a very thin lens (= vertical line in the center); a, distance between object and lens; b, distance between image and lens; if, focal length; F, focus; O, object; and IO, image of the object. Effects of different positions of an object in the periphery of a thin lens (b–d).

Figure 6 Major optical aberrations of simple lenses. (a) Spherical aberration; (b) chromatical aberration; and (c) field curvature.

length of lens and the distance between the lens and the object. Figure 5b–d shows three major cases of object positioning. First, the object is within the focal length (Figure 5b). This yields a *virtual magnified* upright image. In light microscopes, the eyepieces hold such a function. Second, the object is exactly in the focal length (Figure 5c). In this case, no image is obtained. Third, the object is out of focal length (Figure 5d). This results in a *reversed real magnified* image – as produced by the objectives of light microscopes. The image of an object visible in a compound microscope is bottom-up and inverted.

Real lenses unfortunately show the so-called (optical) aberrations (Figure 6). In the following, the most importing aberrations are mentioned:

- One tedious error is the so-called *spherical aberration* (Figure 6a). Light waves passing through the periphery of a lens are not brought into identical focus with those passing closer to the center. This constitutes one serious optical error because the image of an object is spread out longitudinally and transversely rather than being in sharp focus.
- The second aberration results from the compound nature of white light. As mentioned earlier, blue light is refracted to the greatest extent followed by green and red. This leads to fringes surrounding the image caused by wavelength-dependent foci dispersed along the axis, a phenomenon called *chromatic aberration* (Figure 6b).

- A third but (for many practical purposes) less important defect of lenses is the *field curvature* or *flatness of field* (Figure 6c). This defect leads to an image that is sharp either in the center or in the periphery. The image is only sharp if it is projected on a curved area.

To overcome the listed aberrations, the optical parts of a light microscope are corrected to a degree depending on the expense of the respective parts (see later on).

2.2 Compound Microscope and Stereo Microscope

The optical microscope, also often referred to as *compound microscope*, has to be distinguished from the so-called *stereo microscope*. The stereo microscope is an indispensable apparatus for a quick inspection of botanical materials and is sometimes referred to as *dissecting microscope*. It is designed for the visualization of material too small for the naked eye but too big and maybe too opaque for the study in a compound microscope. It uses two optical paths, each with one's own objective and eyepieces mounted on an appropriate stage to provide slightly different viewing angles to the left and right eyes. This way, it provides the operator with a three-dimensional upright visualization of the sample being examined.

In contrast, modern *compound microscopes* (= microscope in the context of this chapter) also have two oculars but only one single light path. The following sections give a short overview of the parts of an optical microscope, its functions, and usage.

Some short introductive statements in the beginning:

- Commonly visible light transports the object information from the specimen through the microscope to the eye of the observer.
- The researcher observes an image of the object but not the object itself!
- The properties of light and the properties of the microscope translate the object information into a microscopic image.
- Only the best management of the light path allows translation of object information as accurately as possible into a more or less aberration free image!

On the basis of this short introduction to optics, the following sections explain the basic technical principles of light microscopes in addition to the best practice for using this instrument.

Basic principles of a modern compound microscope: "Compound" commonly refers to the fact that in order to enlarge an image, light passes through a series of lenses in a line. Each group of lenses magnifies the image over the previous one (one light path, multiple lenses = compound microscope).

The core components are the *objective* (one or more mounted in the nosepiece) and the *eyepiece* (Figure 7). Objective and eyepiece are composed of multiple lenses. Figure 8 shows the old-fashioned construction with a so-called finite tube length (commonly in between 160 and 210 mm) and a modern one with an infinity optical system. The older type needs objectives, which are designed to project a diffraction-limited image at a fixed plane (the intermediate image plane), which is described by the microscope tube length and located at a prespecified distance from the rear focal plane of the objective (a). The infinity type of construction (b) makes it easier to put additional optical equipment (e.g., a drawing tube) in the light path by stretching out the parallel optical path. Objectives for this type of microscope emerge rays in parallel bundles and require a supplementary tube in the light path to bring the image into focus at the intermediate image plane. The names used for this lens are different in German and English literature and therefore ambiguous.

In both systems, specimen are pictured at a short distance beyond the front focal plane of the objective through a medium of a well-defined refractive index, usually air ($n = 1.003$) or immersion oil ($n = 1.52$). Different media require different types of objectives.

2.3 Parts of a Compound Microscope

- *Base and stand*: Both are necessary for a stable aligned light path.
- *Illumination system*: A typical microscope contains the following components:
  - *Illumination lamp*: *Tungsten-halogen* lamps are most important in optical microscopy and
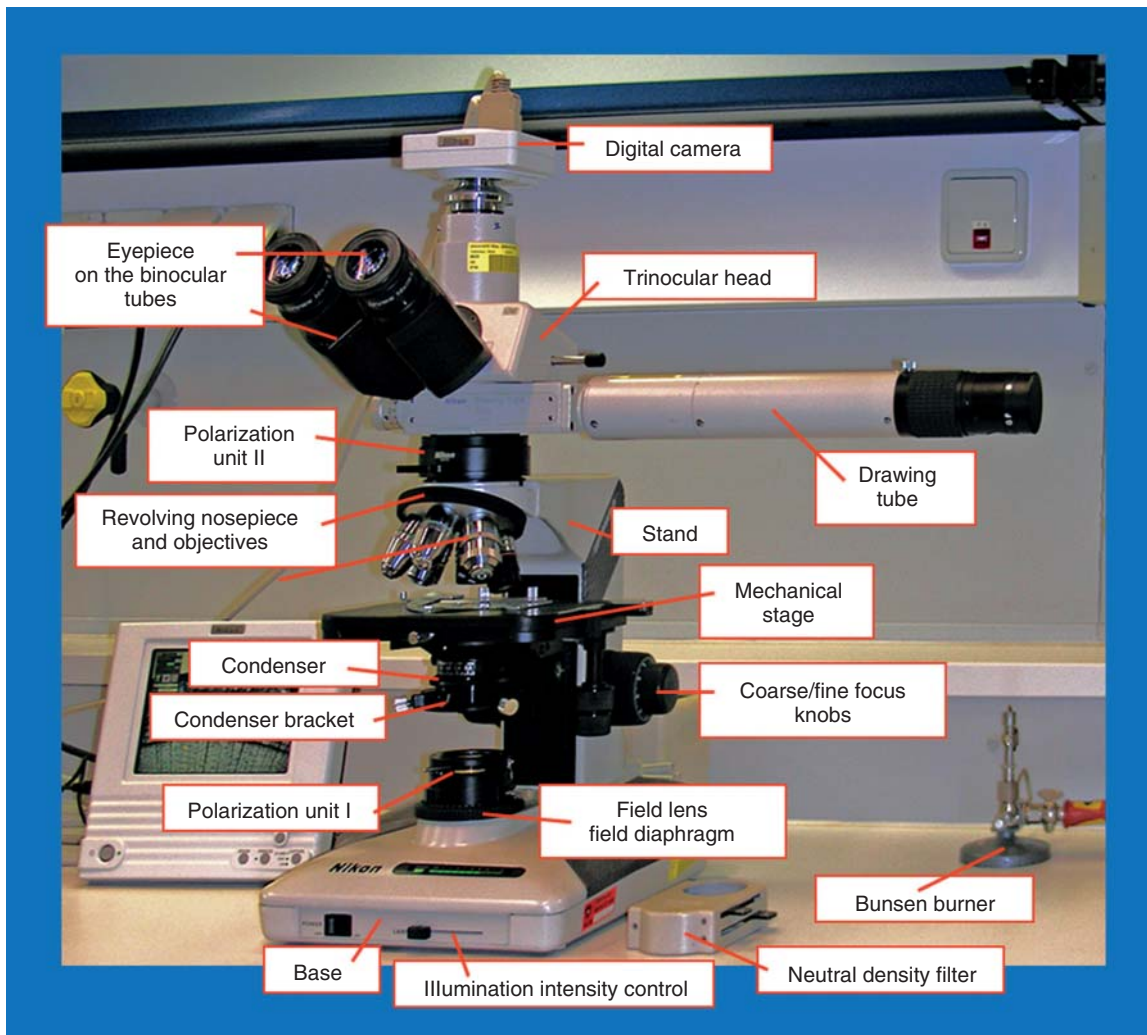
Figure 7 Main parts of a typical microscopic working place and of a compound light microscope.

continue to be the type of choice for pharmacognostic imaging. Because of generating a continuous distribution of light across the visible spectrum, they are ideal for microphotography too. (See also http://micro.magnet.fsu.edu/primer/anatomy/lightsourceshome.html)

- Eco-illumination (e.g., Nikon microscopes Ci-E/Ci-L): This is a newly developed high luminescent LED system with low power consumption. The new system produces evenly distributed illumination and reduces the cost of lamp replacement.

- *Field lens*: It is responsible to focus the image of the lamp filament at the plane of the condenser aperture diaphragm.
- *Field diaphragm opening*: It serves as a virtual light source for the microscope and its image is focused by the condenser after axial movement onto the specimen plane (see below Koehler illumination).
- *Condenser* (see below): This system facilitates the illumination of the specimen. The field diaphragm only controls the width of the bundle of light reaching the condenser. It does not
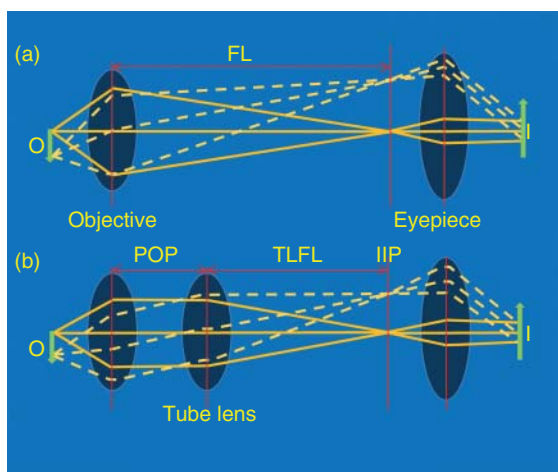
Figure 8 The light path in compound microscope with finite tube length (a) and with infinity tube length (b) [FL, focal length; I, image; O, object; POP, parallel optical path; TLFL, tube lens focal length; and IIP, intermediate image plane].

affect the optical resolution (NA see below) directly.

- *Mechanical stage* with *specimen holder* and an *X–Y translation mechanism*: It facilitates the exact translation of the specimen in *X* and *Y* directions. It is necessary for recording the X/Y position of interesting parts of the specimen at low magnification for a later observation at high magnification with an oil immersions objective (see below). For this purpose, a small drop of immersion oil is put on the cover glass and therefore the switch back to a dry objective is forbidden because of the improper refraction index for such an objective and the possibility to contaminate the frontal lens of the dry objective with oil.

- *Coarse* and *fine focus knobs*: They enable the operator to focus layers of interest in the object by moving the stage axially.

- *Condenser*: It is typically mounted directly beneath the stage in a bracket, which can be raised or lowered independently of the stage. This bracket can be slightly moved in *X* and *Y* directions and facilitate the alignment of the condenser in the correct position of the light beam (Figure 9).

The condenser gathers light from the light source and concentrates it into a light cone (Figure 10). After passing through the opening in the stage or the object, the light diverges to an inverted cone with the same angle. In the space behind the transparent object, an image of the object is formed by interference through

the combination of undiffracted and diffracted light of higher order (Figure 10). The higher the order of diffracted light gathered by the objective is, the better is the translation of object structures in the microscopic image (remember also Figure 4). The angle of the cone is the so-called aperture. The combination of the refractive index *n* from the medium between the specimen and the front lens of an objective (air, immersion oil, etc.) and the one-half angular
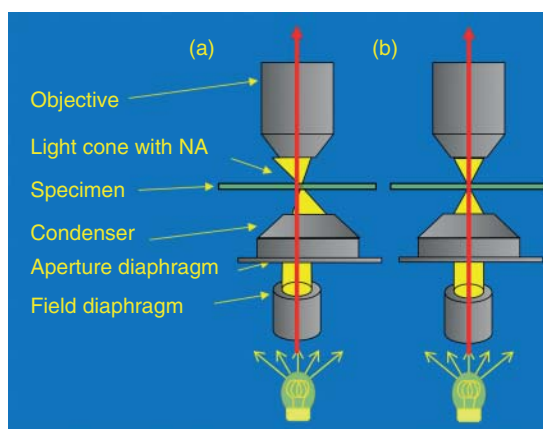
Figure 9 Alignment of the light cone produced by the condenser: (a) misalignment and (b) correct alignment.
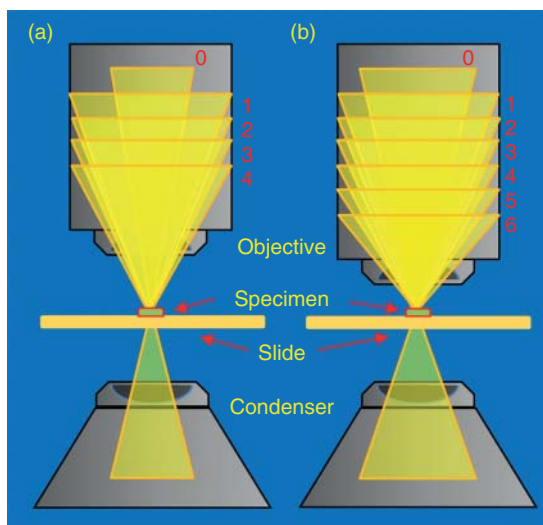
Figure 10 A simplified scheme of the light cone produced by the condenser and the specimen gathered by the objective of undiffracted ($0 =$ zeroth order) and the diffracted light cones of higher order (1–6): (a) numerical aperture smaller and (b) numerical aperture higher.

aperture α is referred to as numerical aperture (NA, $NA = n*\sin(\alpha)$). The diameter and numerical aperture of the light cone are regulated by adjusting the aperture iris diaphragm enabled by a lever or by rotating a collar on the condenser housing. It is very important that the numerical aperture of the light cone gathered by the condenser fits into the numerical aperture of the objective in use. The greater the difference is, the poorer will be the information about the specimen (compare Figure 10a and b).

Figure 10 shows the different amounts of diffracted light cones gathered by two objectives with different NA and the different angles of the light cone of zeroth order caused by the change of the opening of the aperture diaphragm. Every change of magnification (objective) needs a corresponding adjustment of the aperture diaphragm to provide the proper light cone for the numerical aperture of the objective chosen.

Cave: The most common errors are:

- the control of the illumination intensity by opening and closing the aperture diaphragm,
- the opening of the aperture diaphragm being too small (increasing contrast and decreasing resolution), or
- an unintended movement of the condenser vertically or axially after the correct adjusting.

These errors are the main sources of image degradation and therefore possible misinterpretation of structures in the specimen.

For different microscopic techniques such as *phase contrast* or dark *field imaging* (the light of zeroth order is stopped by a proper plate, and oblique light only causes the image formation), different types of condensers are designed.

- *Objectives*: Together with the condenser, objectives are the most important and expensive components of a microscope. Important parameters are inscribed on the barrel of the objective itself (Figure 11).

Figure 11 depicts the engraved information of a 60× *plan apochromat* objective (e.g., flat-field correction, optical aberration correction, lateral magnification, numerical aperture, and tube length) necessary to determine what the objective is designed for. Microscope manufacturers offer a wide range of objective designs to meet the performance needs of specialized imaging methods (e.g., *phase contrast*, *fluorescence*, and *polarizing* microscopies) or to modify the effective working distance of the objective (Figure 13). Finite-corrected and
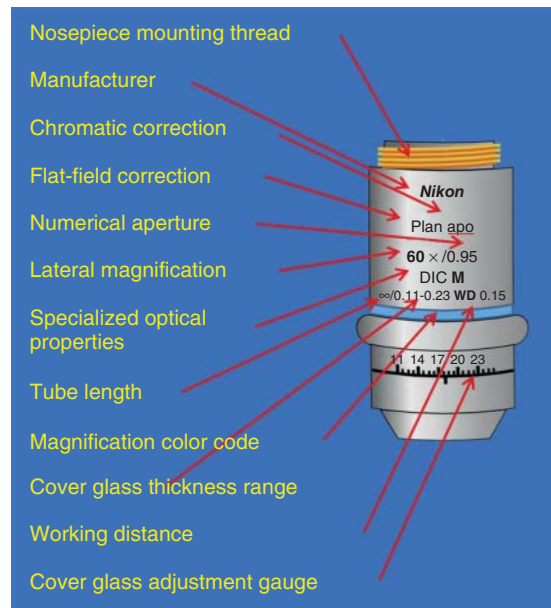
Figure 11 60× plan apochromat objective, including common engravings that contain the specifications.

infinity-corrected objectives are not interchangeable, that is, finite-corrected objectives do not fit into infinity-tube length microscopes and vice versa. To overcome the defects of lenses listed earlier, the manufacturers of microscopes developed different types of objectives:

- *achromatic* –corrected for axial chromatic aberrations in two wavelengths (red and blue) and for spherical aberration in green (compare Figure 6a and b);
- *fluorite* (or *semi-apochromates*; similar to the achromats, the fluorite objectives are corrected chromatically for red and blue light; in addition, they are corrected spherically for two colors); and
- *apochromatic* objectives. These are corrected chromatically for the colors red, green, and blue (and sometimes for dark blue too) and corrected spherically for two colors.

All listed objective types suffer from pronounced field curvature (Figure 6c). Most manufacturers produce flat-field corrected objectives. For many purposes, the field curvature is negligible; however, in photomicrography, it is an intolerable malfunction. That kind of objectives are called *plan achromates*, *plan fluorites*, or *plan apochromates*; although these corrections are expensive, the resulting objectives are state of the art in contemporary photomicrography.
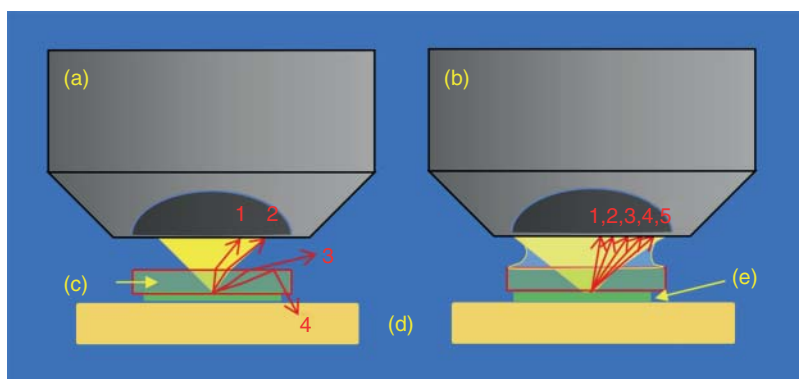
Figure 12 Influence of the medium between the specimen and the front lens of an objective: (a) dry objective, medium = air; (b) oil immersion objective, medium = oil; (c) cover glass; (d) slide; and (e) specimen.

Table 1 Resolution and numerical aperture by objective correction.

Magnification	Plan achromat		Plan fluorite		Plan apochromat	
	NA	Resolution (µm)	NA	Resolution (µm)	NA	Resolution (µm)
4× (air)	0.10	2.75	0.13	2.12	0.20	1.375
10× (air)	0.25	1.10	0.30	0.92	0.45	0.61
20× (air)	0.40	0.69	0.50	0.55	0.75	0.37
40× (air)	0.65	0.42	0.75	0.37	0.95	0.29
60× (air)	0.75	0.37	0.85	0.32	0.95	0.29
100× (oil)	1.25	0.22	1.30	0.21	1.40	0.20

Source: Reproduced with permission from M. W. Davidson. Microscopy Basics: Resolution. MicroscopyU http://www.microscopyu.com/articles/formulas/formulasresolution.html.

Another very important factor is the medium between the cover slip and the front lens of the objective. The higher the refractive index of this medium is, the higher is the angle of the light cone gathered by the objective front lens.

Figure 12a shows the light cone for a so-called dry objective and Figure 12b shows the light cone for an oil immersion objective. It is obvious that the usage of an oil immersion objective brings more diffracted light cones of higher order into the front lens of the objective, and therefore leads to a better resolution of the image by interference and to a higher NA of the oil immersion objective. Table 1 depicts the differences of the most used objective types, their possible numerical apertures, and the maximum of defining power. As given in the table, a better correction and a higher NA lead to a higher resolution.

Another, sometimes crucial, feature of objectives is the working distance. Figure 13 shows the working distances for objectives with different magnification. Usually, the smaller the magnification is, the greater will be the distance between the front lens of the objective and the cover slip of the specimen. However, this is not true for objectives with a magnification smaller than 4× and for the so-called long distance objectives (Figure 13d). However, note that two objectives with the same magnification but different working distances cannot have the same NA for geometrical reasons (compare Figure 13c and d).

- *Polarization* units I and II: two polarizing filters – the polarizer (unit I in Figure 7) and the analyzer (unit II in Figure 7) – are necessary to switch a normal bright field microscope into a polarizing microscope. The first filter is placed between the lamp and the condenser; the second is placed under or within the binocular tube. Polarized light microscopy is a contrast-enhancing technique that improves the information of the image obtained from birefringent materials in the specimen. This technique was originally developed for the investigation of minerals and crystals of chemicals. For this purpose, a very expensive apparatus is necessary. Polarized light microscopes have a high degree of sensitivity for birefringent structures and thus can be utilized for botanical drug identifications by means of structures with birefringent properties (starch, cell walls, calcium oxalate crystals, etc.). The demands for this application are less expensive and only two pieces of polarizing film are at least necessary. The polarizing unit shown in Figure 7 is a combination of
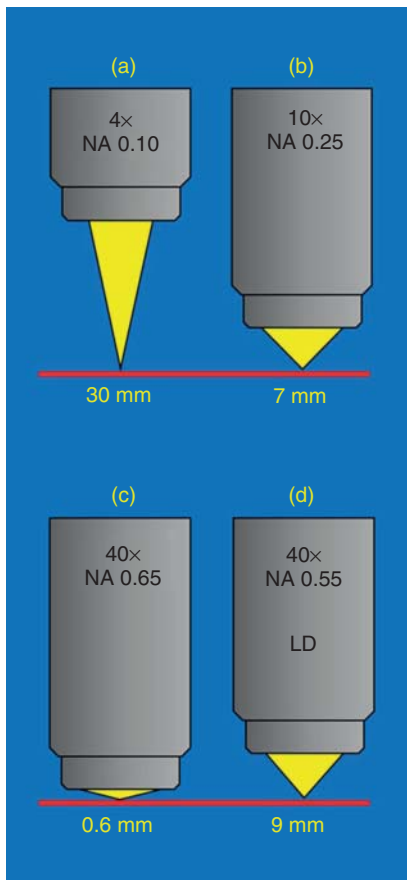
Figure 13 Working distances (length of yellow light cone) for objectives with increasing magnification (a–c) and the comparison of two objectives with the same magnification but different working distances and therefore different NA's too (c and d). LD, long working distance objective.

a polarizer and a compensator (see in the previous text). The compensator causes a phase shift in a predictable value, and after obstructive and destructive interference, the background of the specimen is purplish-red and not black as without the compensator all birefringent structures are yellow or blue colored (depending on the relative position against the polarizer). A tricky technique for the visualization of very thin cellular structures is the combination of a polarizing unit with a dark field condenser.

- *Drawing tube*: This equipment enables a more or less easy tracing of microscope images while looking into the eyepieces.

- *Trin- or binocular tube*: This part of the microscope holds the eyepieces. A trinocular tube provides the possibility to look into the eyepieces and take a photograph or transmit a video of the session concurrently.

2.4   The Proper Use of a Microscope

2.4.1   Köhler Illumination

The *Köhler illumination* is a technique that provides optimum illumination, resolution, and contrast by aligning and focusing the condenser and field diaphragm and setting the aperture diaphragm of the condenser to fit the numerical aperture of the objective lens. It is important to understand more than the procedural steps to gain a Köhler illumination. Being familiar with the concepts behind these steps will foster the knowledge necessary to improve image quality for challenging specimens. An understanding of the illumination and image-forming light path is important in order to realize the full imaging potential of microscopes.

Figure 14 depicts two sets of the so-called *conjugate planes*. Each plane within a set is said to be conjugate with the others in that set because they are simultaneously in focus. The left side (Figure 14a) shows the image-forming light path, where the focus is on the conjugate field plans. The right picture (Figure 14b) shows the illumination light path. Here, the focus is laid on the conjugate aperture plans. In normal observation mode, the conjugate set of field planes can all be simultaneously viewed when the specimen is focused, being referred to as the *orthoscopic mode* and providing an *orthoscopic* image. Viewing the other set of conjugate aperture planes requires an eyepiece telescope in place of an eyepiece or a built-in Bertrand lens. Both pieces of equipment give us the ability to focus on the rear aperture of the objective. This mode is termed as conoscopic, aperture, or diffraction mode, and the image observed at the objective rear aperture is known as the *conoscopic image*.

The importance of these different views was first pointed out by Köhler (1893): *When something is in focus in one set of conjugate planes*, *it is "maximally out-of-focus" in the other set of planes*!

To achieve the Köhler illumination, proper adjustment of the condenser and the field diaphragm is critical. The following adjustments must be made by the
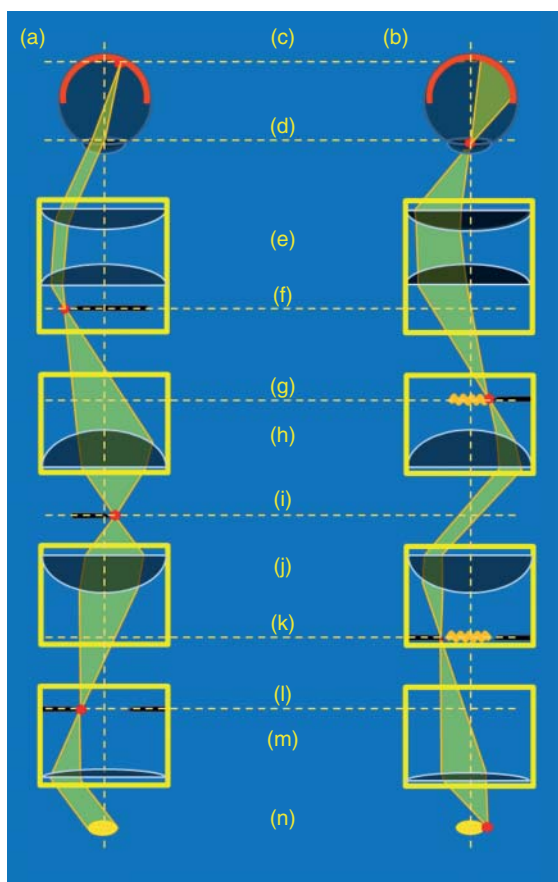
Figure 14 (a) Image-forming light path – conjugate field plans and (b) illumination light path – conjugate aperture planes. (c) Retina; (d) eyepoint – iris diaphragm of eye; (e) eyepiece; (f) intermediate image plane with fixed diaphragm; (g) objective back focal plane; (h) objective; (i) specimen; (j) condenser; (k) aperture diaphragm; (l) field diaphragm; (m) collector lens; and (n) lamp filament.

- Stop down the field diaphragm most of the way. While looking into the eyepieces, carefully raise or lower the condenser until the polygon-shaped edge of the field diaphragm is completely inside the field of view and is sharply focused. The edge may exhibit a slight blue tint. Using a highly corrected condenser, no colored tint is seen. Variation in color along the edge of the field diaphragm indicates that some components of the light path are not properly aligned (see Figure 9). For routine work, this may not matter, but a proper alignment is necessary to obtain optimal analog and digital pictures.
- If the image of the field diaphragm is not centered, align the illumination by centering the condenser or/and the field diaphragm. Then, open the field diaphragm until it just disappears from view.
- Adjust the condenser aperture diaphragm. The opening size of the aperture diaphragm is always a compromise between resolution (open diaphragm, Figure 15c) and contrast (more or less closed diaphragm, Figure 15a and b), and it depends on the characteristics of the specimen. When the diaphragm is opened too much, stray light generated by refraction can cause glare and lower the overall contrast. The opposite view is shown in Figure 15a, the aperture diaphragm is nearly closed. In this case, artificial refraction fringes occur in the image.

With a correct Köhler illumination, the image appears bright and evenly lighted providing ideal illumination for all purposes.

2.4.2 Eyepiece Adjustment

The eyepieces (oculars) provide an image to the user. The visual acuity and the distance between the eyes vary from person to person. This requires an individual adjustment. Neglecting this need may raise discomfort for the user. For some persons, it is not easy to make both images coincide. Modern microscopes have a diopter-adjustment ring on one side of the bino- or trinocular tube and hence, the possibility to change the distance between the eyepieces. The following steps are necessary to obtain an optimal adjustment:

- Place a specimen on the stage and select a low magnification objective.

microscopist each time the microscope is used and each time an objective is changed!

- Place a specimen on the stage and bring it to focus using the objective with the lowest magnification and therefore the longest working distance to avoid damaging the specimen or the front lens of the objective. It should also be in mind, as a rule, that the higher the magnification is, the lesser will be the working distance. Now, open the field diaphragm on the base and the aperture diaphragm beneath the condenser all the way. Adjust the light control, so that the light intensity is comfortable.

Figure 15 Diatom: *Stauroneis phoenicenteron* (Nitzsch) Ehrenberg, photographed with NIKON plan Apo objective 60× with NA 0.65. Effect of different aperture openings: (a) aperture nearly closed; (b) aperture optimally controlled; and (c) aperture maximally opened.

- If the diopter-adjustment ring is on the left side, close the left eye and look at the specimen with the right eye. Locate a well-defined target (e.g., the edge of structures) and bring it into focus.
- Open the left eye and rotate the diopter-adjustment ring alternately clockwise and counter-clockwise until the defined target is in focus too.
- Set the proper interpupillary distance for your eyes by moving the eyepiece tubes together or apart as necessary. Rarely, if the brain is not in the mood to align the two images, then it is helpful to have assistance. First, stop down the aperture diaphragm and turn the lamp to maximal power. An assistant person shall now look at your eyes where the bright spots of the microscope are also visible, and can adjust the distance between the eyepieces properly for you. After this, counter rotate the lamp control and open the aperture diaphragm to a proper position and look on a faraway object and then quickly into the eyepieces. Repeat the procedure as long as necessary.

2.5 Dirt in the Microscope

Clean optics and cleanliness while working with a microscope are preconditions for successful microscopy. Routine care of the optical surfaces as a part of regular use will help to preserve the image quality. However, even with utmost care, one cannot always avoid dirt and dust leaving marks in microscopic images. Whether dirt has strong or little effect on the visual or recorded image depends on its location within the light path of the microscope – closeness to the object or to a camera sensor results in a greater impact on the image (compare the chapter Koehler illumination).

For one thing, the surfaces (on top and below) of the anatomical specimen – the surfaces of the cover glass and the microscope slide – have to be spotless to guarantee an optimal image. For another thing, dirt can be located on one of the diverse optical lenses:

- the front lens of the objective,
- the ocular,
- the condenser front lens,
- if using a camera, the camera sensor and the optics of its adapter.

Further possibilities may be other glass surfaces in the light path such as the surfaces of lamps or diverse filters.

The following tips are required for routine use:

- Keep the microscope covered when not in use.
- Keep fingertips away from optical surfaces.
- After use, carefully remove contaminants from all outer optical surfaces using manufacturer's recommendations.

2.5.1 Locating the Dirt

Dirt in the optical path is always the source of annoyance. Fundamentally, the critical surface can be determined when that surface is moved and the dirt follows its movement.

- First of all, be sure that the anatomical specimen itself is clean. If the dirt moves while moving the sample using the XY stage, the dirt is located on the cover slip or slide.
- Check the objective: If the dirt is only visible when using one objective rather than another, clean the surface of the front lens of the respective objective.
- Check the ocular lenses by rotating the eyepieces.
- By moving the condenser up and down, dirt on the field diaphragm and the condenser can be detected.
- Carefully rotate the components of the microscope and the camera. Always remember that the dirt moves when the corresponding surface moves.

For a macroscopic control of optical surfaces such as the lenses of the ocular and objectives, an eyepiece holding the "wrong way" upside down provides useful service as magnifier.

For cleaning, use a dust blower (never blow with your mouth to avoid moisture!), distilled water and high purity cotton, soft facial tissues, or cotton buds. Never use alcohol or other organic solutions such as acetone for cleaning because of possible damage to objective front lenses!

3 PLANT TISSUES AND THEIR VALUE FOR PHARMACOGNOSY

For microscopy, one ought to engage oneself in studying the anatomy of the botanical raw material to be become familiar with the characteristics of the used drug, a prerequisite for detection of foreign elements. Thorough specific knowledge of histology, cell forms, tissue elements, and cell contents is essential to judge the identity and quality of crude drugs and identify possible adulterations.

Plant cells are organized into various tissues. Thus, the cells work together as a functional unit. A tissue is composed of either a single cell type (simple tissue – e.g., parenchyma or collenchyma) or different ones (complex tissue – e.g., xylem or phloem). Three tissue systems are distinguished: the *dermal tissue* system (the outer protective layer of the plant), the *vascular tissue* system (responsible for the transport and distribution of nutrients and water), and the *fundamental tissue* system (ground tissue system, storage, and support for the plant). Figure 16 gives an overview of the three tissue systems and their associated cell types. Secretory structures are not considered, as they occur throughout the plant body and do
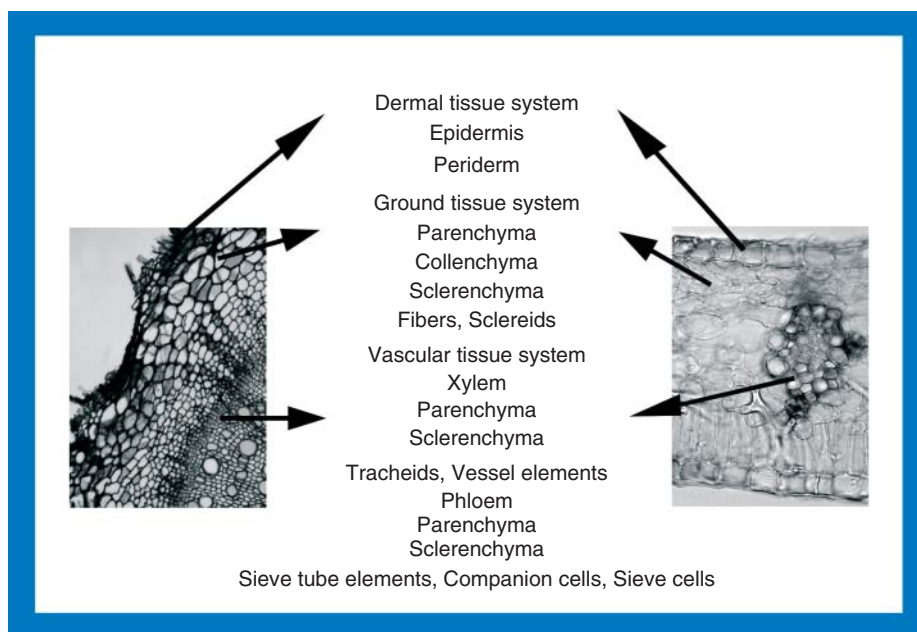
Figure 16   The three tissue systems and their corresponding tissues and cell types.
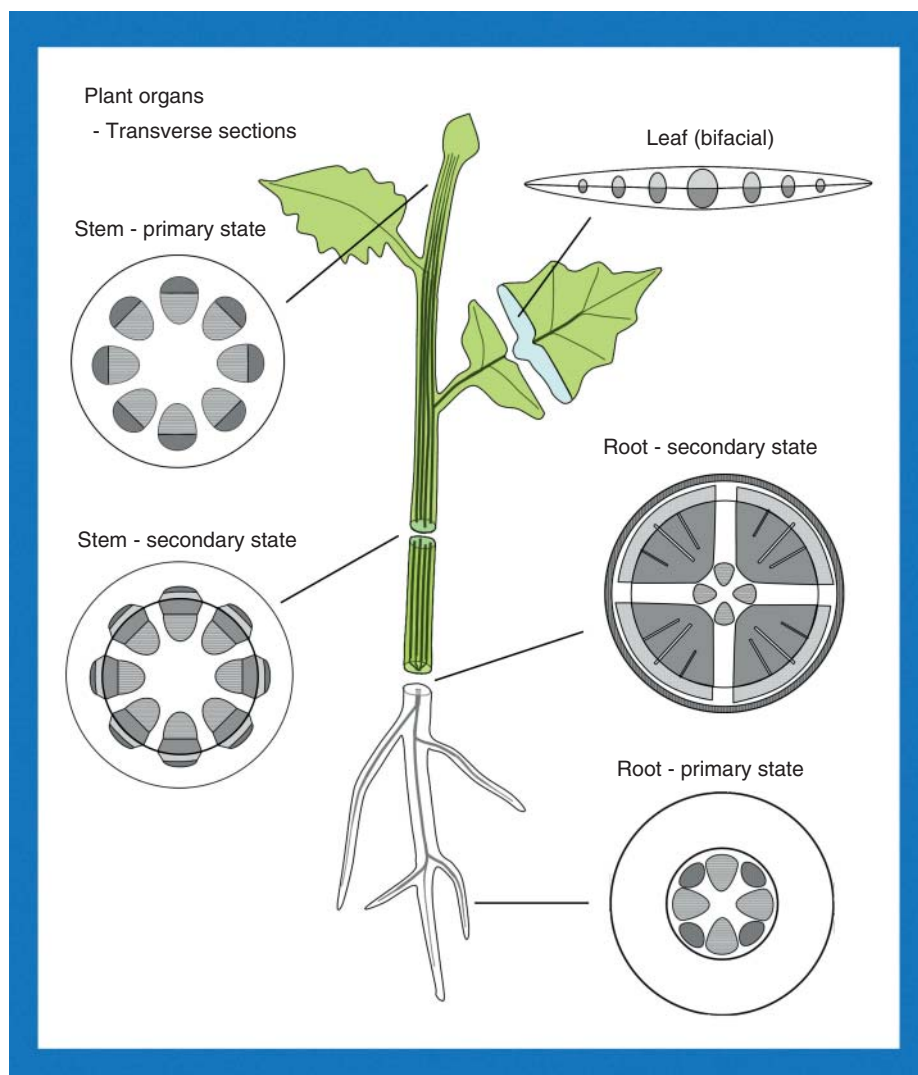
Figure 17 Dicotyledonous plant with all organs – leaf stem and root – in an early stage and short after the beginning of secondary growth.

not form definite tissues. Principal differences among the three plant organs – root, stem, and leaf – emerge from the relative position of the vascular and ground tissues, which are highly important characters for identifying the drug material in hand. Figure 17 shows the leaf, the stem, and the root from a young dicotyledonous plant. It is very important to internalize the typical anatomical structure of every organ.

The growth of the plant, that is, the development of all tissues depends on the activity of small regions of undifferentiated cells – the so-called meristems. There are two major types of meristems:

- *Apical meristems*, located at the tips of root and stem, allow for the primary growth (increase in length) of the plant. Tissues deriving from the apical meristems are also called *primary tissues*.
- *Lateral meristems* cause, by building up new cell material, the secondary plant body. The stems and roots of some plant groups (gymnosperms and dicotyledonous angiosperms) form wood and become thick. The lateral meristems comprise two types – the *vascular cambium* gives rise to the secondary growth (increase in girth) and the *cork cambium* yields the bark. Both build up

the so-called *secondary tissues*. According to the final position in the plant organ, a newly derived cell from a meristem undergoes differentiation whereby it achieves its distinctive characteristics in form, structure, physiology, and function. The extent of differentiation and specialization varies considerably. A broad range of helpful books in the field of plant anatomy are available (Cutler, Botha, and Stevenson, 2008; Evert, 2006; Foster, 1942; Kaussmann and Schiewer, 1989; Kraemer, 1910; Kremer, 2002; Metcalfe and Chalk, 1957).

3.1   Epidermis and Periderm

3.1.1   Epidermis

The *epidermis* is the outermost cell layer of a primary plant body, generally a single layer of cells lacking intercellular spaces. Sometimes additional cell layers can be found underneath the epidermis (i.e., the hypo-dermis). The anticlinal walls of the epidermal cells, seen on surface view, form characteristic patterns classified as straight-walled/polygonal, wavy-walled, or beaded and particularly useful for diagnostic purposes (Freire, Urtubey, and Giuliano, 2007; Saukel, 1982b, 1982c, 1983c, 1983d, 1984a; Soh and Parnell, 2011); for example, the epidermis cells of seeds and fruits vary considerably in size and form as well as thickness of the walls.

A waxy coating referred to as *cuticle* is present at the outside of the epidermis of all organs (stems, leaves, fruits, and seeds) except the youngest parts of the roots. The cuticle considerably varies in thickness and surface sculpture (e.g., striation) providing useful characters for discrimination of species.

Part of the epidermis cells differentiate into specialized cells early in their development – the *stomata* and the *trichomes*. Stomata (singular stoma), found on all aerial parts of the primary plant body except roots, are microscopic pores (or apertures) each formed by two guard cells, which regulate the gas exchange and water balance by controlling the size of the pores. Typically, an aerial cavity can be found beneath the stoma. The guard cells are surrounded by the subsidiary cells, both cell types forming together the stomatal complex. On the basis of the shape of the guard cells and the number, relative size, and arrangement of subsidiary cells, different types of stomatal complexes are

Anomocytic stomata: subsidiary cells indistinguishable in size, shape, and form from other epidermal cells

Anisocytic stomata: one subsidiary cell distinctly smaller than the others

Diacytic stomata: the common walls of the two subsidiary cells form a right angle to the long axis of the aperture

Paracytic stomata: two subsidiary cells parallel of the aperture

Graminaceous stomata: dumbbell-shaped ground cells with two subsidiary cells

Figure 18   Some important types of stomata used in most common literature and pharmacopoeia.

distinguished. The terminology of stomatal complexes (stoma + abutting cells) dates back to the work of (Prantl, 1872). Five types are commonly distinguished in present-day pharmacognostic literature: the *anomocytic* (= ranunculaceous) type, subsidiary cells are indistinguishable in size, shape, and form from other epidermal cells; the *anisocytic* (= cruciferous) type, one subsidiary cell distinctly smaller than the others; the *paracytic* (= rubiaceous) type, two subsidiary cells are arranged parallel to the aperture; the *diacytic* (= caryophyllaceous) type, the common walls of the two subsidiary cells are arranged at right angle to the longitudinal axis of the aperture, after (Metcalfe and Chalk, 1957); and additionally, the *graminaceous* type, with dumbbell-shaped guard cells and two subsidiary cells (Figure 18). The first four types are standard in the European Pharmacopoeia (7.0, Chapter 2.8.3. Stomata and stomatal index, p. 239).

This classification, although widely used in pharmacognostical literature, is unsatisfying in several respects. Prabhakar (Prabhakar, 2004; Prabhakar and Leelavathi, 1987) defined 11 types of stomatal complex (including the description of the old paracytic, diacytic, anisocytic types, the graminaceous type as tetracytic type, and a slightly changed anomocytic type), and presented a great number of illustrative drawings for all types in his work. Figure 19 illustrates these 11 types:

1. *Pericytic stomata*: a single subsidiary cell surrounds the two guard cells.
2. *Desmocytic stomata*: a single subsidiary cell surrounds the two guard cells with a conjoint wall variable in position.
3. *Paracytic stomata*: two subsidiary cells with their conjoint walls arranged parallel to the guard cells.
4. *Diacytic stomata*: two subsidiary cells with their conjoint walls arranged lateral to the guard cells.
5. *Anisocytic stomata*: three subsidiary cells, one distinctly small.
6. *Anisotricytic stomata*: three subsidiary cells, one distinctly large.
7. *Isotricytic stomata*: three subsidiary cells of more or less equal size.
8. *Tetracytic stomata*: four subsidiary cells, – two in polar, two in lateral position (graminaceous type).
9. *Staurocytic stomata*: four subsidiary cells, – two of them with their conjoint walls polar, whereas the other two are lateral to guard cells.
10. *Anomocytic stomata*: more than four subsidiary cells (other than tetracytic and staurocytic types).
11. *Cyclocytic stomata*: four or more subsidiary cells are arranged in a narrow ring.

Some of the stomata types are characteristic taxonomically even on the family level: for instance, the tetracytic type with dumbbell-shaped ground cells (former graminaceous type) are typical for Poaceae and other members of Poales; diacytic stomata are common in Caryophyllaceae, Lamiaceae, many members of Verbenaceae, and Polygonaceae; and Solanaceae and many members of Brassicaceae show anisocyctic stomata. More examples and details on the development of the stomatal complex can be found in a series of recommended publications (Carpenter, 2005; Croxdale, 2000; Fryns-Claessens and van Cotthem, 1973; Leelavathi, Ramayya, and Prabhakar, 1980; Ostroumova and Kljuykov, 2007; Tahir and Rajput, 2009).

Trichomes (or hairs) originate when epidermal cells enlarge or multiply. They represent epidermal appendages of varying complexity and size. Two principal types of trichomes are distinguished: *covering hairs* (Figure 20) and *glandular hairs* (Figure 21), the latter possessing secreting cells. Classification criteria are the number of cells (unicellular versus multicellular) and of longitudinal cell rows (uniseriate/biseriate/multiseriate) building a trichome, habit (single vs branched), and shape (also note the base of the hair! Figure 20). Additional characters include the consistency of the cell wall (thick- or thin-walled, possibly lignified, and the sculpture of its outer surface (smooth/warty/striated, compare (Saukel, 1982b)). Furthermore, inclusions such as calcium oxalate crystals (e.g., very small needles in many members of Lamiaceae) or cystoliths (e.g., *Cannabis sativa* L.) are also possible. While covering trichomes usually have pointed terminal cells, glandular hairs further possess an often rounded secreting part at the end of the hair (Figure 21), which number and their arrangement can be of high diagnostic value. Therefore, for the description of glandular hairs, the differentiation between stalk and glandular head is important to note. In Figure 21 (4 and 9, 10), one can see that the cuticle of the head is inflated forming a cavity filled with essential oil. The function of the secretory cells is discussed in some publications (Bagchi, 2000; Bruni and Modenesi, 1983; Glover, 2000; Krings, Taylor, and Kellogg, 2002; Muravnik and Shavarda, 2012; Peng and
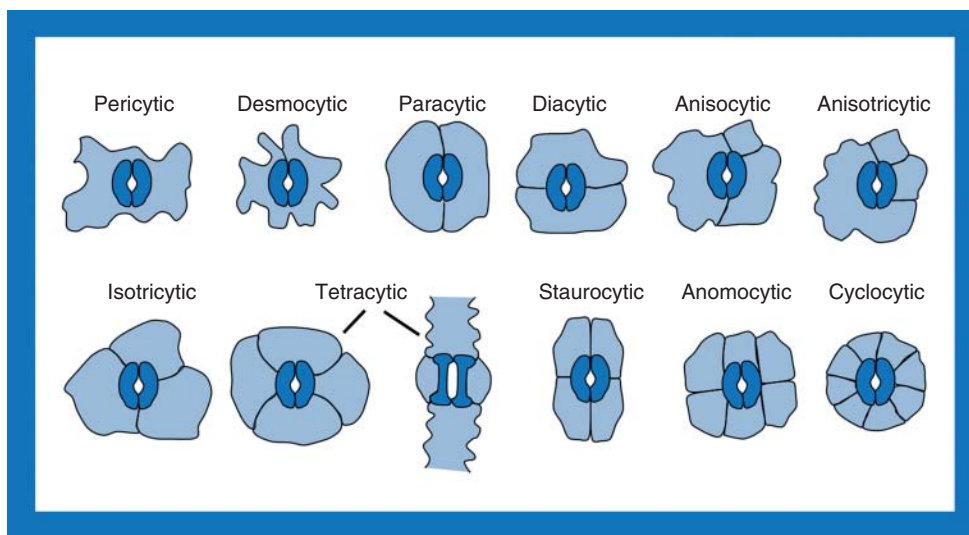
Figure 19 Eleven stomatal types. (*Source*: Adapted from M. Prabhakar (2004).)

Hu, 2007; Singh, Sharma, and Jain, 1974; Turner, Gershenzon, and Croteau, 2000; Wagner, 1991; Wagner, Wang, and Shepherd, 2004). Trichomes are widely used for species delimitations (Andrzejewskagolec, 1994; Eckhart, 1929; Hu, Balangcod, and Xiang, 2012; Prabhakar and Leelavathi, 1989; Saukel, 1984b; Shah and Kachroo, 1975; Singh and Jain, 1975; Singh, Sharma, and Jain, 1974; Xiang *et al.*, 2010), also as characters in phylogenetic research (Beilstein, Al-Shehbaz, and Kellogg, 2006; Cantino, 1990; Downing, Ladiges, and Duretto, 2008) and as a lead character for the identification of botanicals (compare Figure 21). Most trichomes or at least some of their cells are birefringent and can be detected very easily with the polarizing microscope.

In contrast to trichomes, *emergences* (e.g., *Urtica* sp.) can be formed by both epidermal and subepidermal tissues. Nonetheless, a differentiation between trichomes and emergences is not always clearly possible (compare (Corsi, 1992)). Both emergences and trichomes, however, serve as valuable anatomical differentiation characters. Typically structured emergences constitute good authentication characters for Urticae folium or the fruits of Agrimoniae herba.

In roots within primary state, the outermost layer is the so-called *rhizodermis*, differing to the epidermis of the aboveground organs insofar, as cuticle and stomata are lacking. The rhizodermis (always present in a narrow zone just above the root cap) carries the root hairs, which enable an extension of the root surface and thus, improve the uptake of water and nutrients from the soil. Commonly, the root hairs are of a short life span and die after a few days. Along with them, the rhizodermis cells themselves also perish and are replaced by the *exodermis* (outer layer of root cortex made of suberized cells, lying beneath the rhizodermis) as protecting tissue.

3.1.2 Periderm

Undergoing secondary growth of the plant, the rhizodermis and exodermis in roots and the epidermis in stems may be replaced by the formation of a *periderm* (compare Figure 22). A *cork cambium* (= *phellogen*) develops and produces the *phellem* (= *cork*) toward the outside and the *phelloderm* (not always present) toward the inside, all three together forming the periderm (Figure 22b). Commonly, the first phellogen arises in the subepidermal layer, thus replacing the epidermis. It may also occur deeper in any living tissue outside the vascular cylinder, but the occurrence is always genetically defined. Subsequently, continuous secondary growth and further growth in girth may give rise to new phellogens, each time, in deeper regions of the living tissues (secondary phloem). Doing so, the periderm separates inner and outer tissues, causing the tissues outside the innermost phellem to die, thus forming the
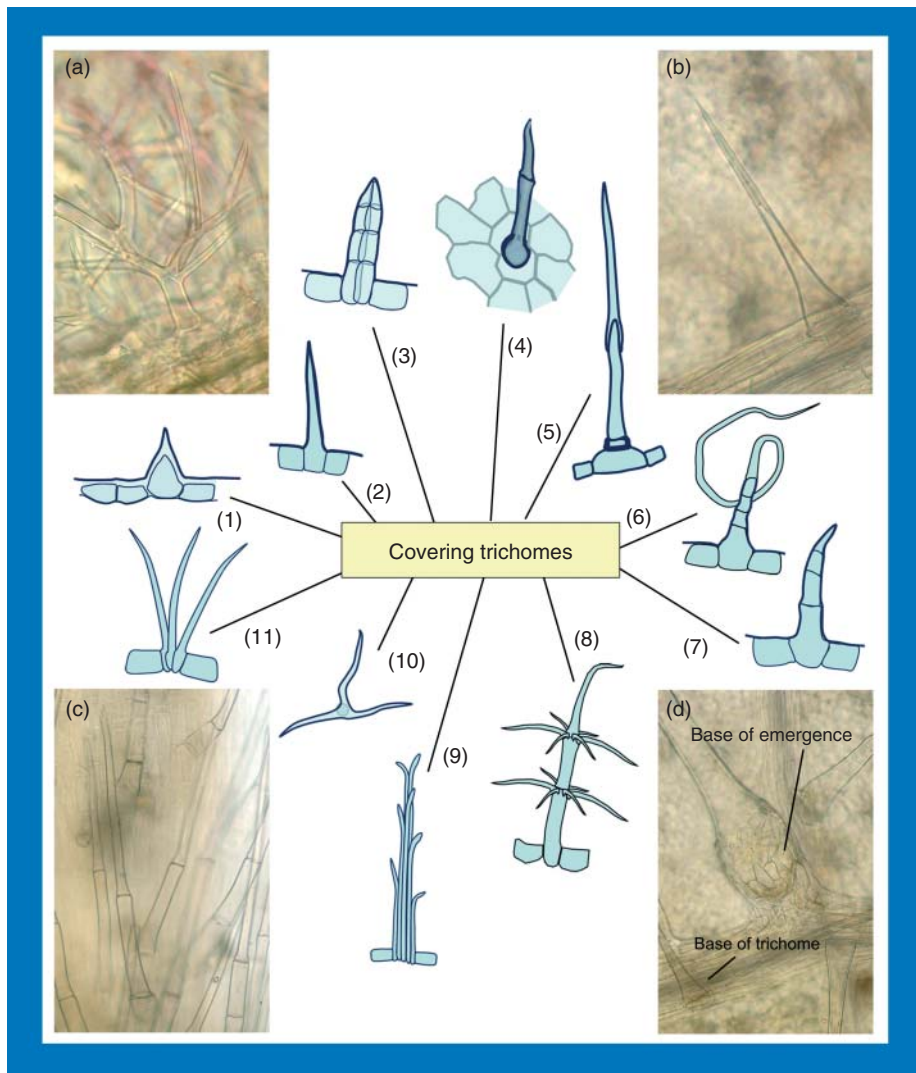
Figure 20 Covering trichomes (a–c and 1–11) and emergence (d): (a) dendritic hair of *Lavandula angustifolia* Mill.; (b) unicellular trichome; (c) multicellular uniseriate trichomes; (d) base of an emergence (*Urtica dioica* L. leaf); (1) unicellular papillae; (2) unicellular trichome; (3) multicellular, biseriate trichome (e.g., many members of Asteraceae); (4)–(7) multicellular uniseriate trichomes; (5) specialized uniseriate trichome from *Plantago lanceolata* L.; (6) multicellular uniseriate trichome with weak hair tip from Asteraceae; (8) dendrite hair from *Verbascum* sp.; (9) multicellular and multiseriate trichome (hair of pappus) from Asteraceae; (10) unicellular stellate hair of Brassicaceae; and (11) an aggregate of unicellular trichomes forming a false stellate hair typically for many members of the Malvales.

rhytidome or outer bark (i.e., the innermost periderm and all tissues outside it) and the inner bark (the tissues between vascular cambium and the active phellogen).

Cork cells are dead at maturity and may be filled with diverse substances (air, fluid, or solid contents). They may also be more or less pigmented (e.g., the white pigment of birch tree is betulin; tannins;

and the phelloderm often containing chloroplasts) or else less. In transverse section, they are arranged in regular radial rows and are often tangentially elongated (compare (Silva *et al.*, 2005)).

The number of layers of cork and phelloderm varies greatly among different species as does the development of the first and subsequent periderms.
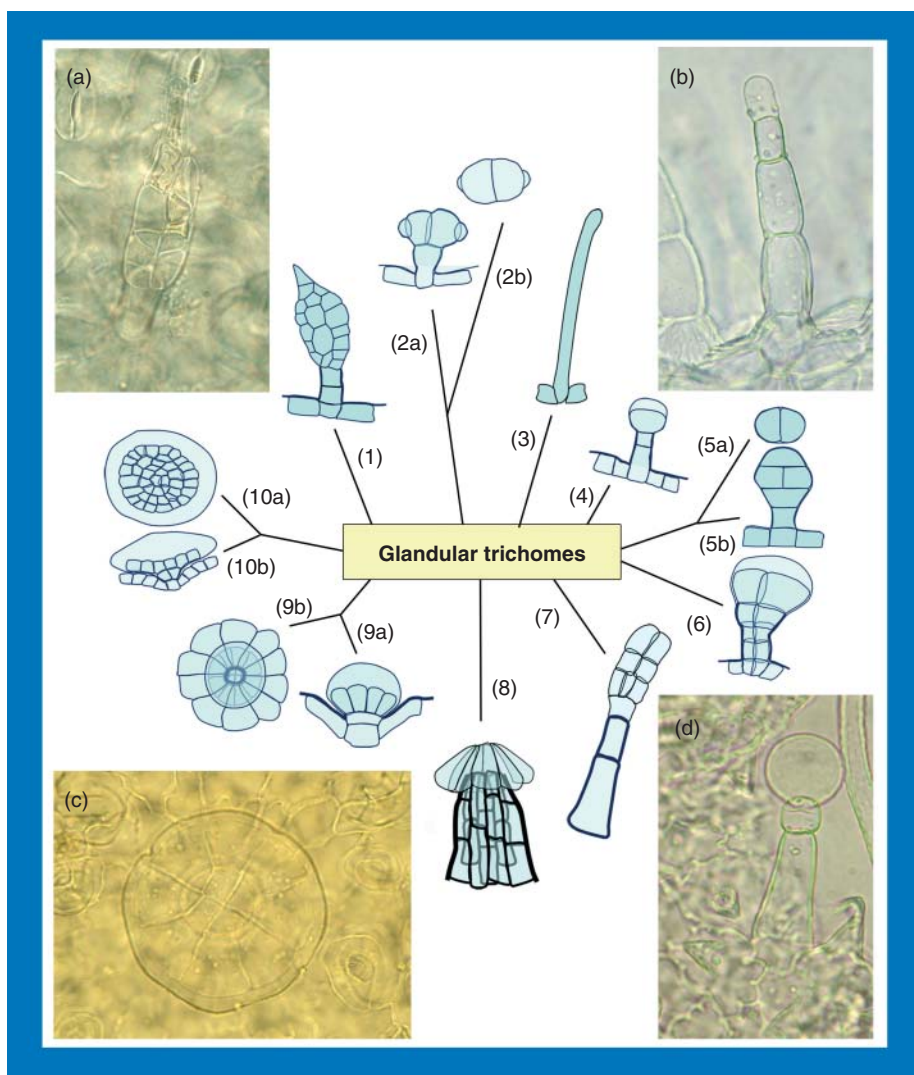
Figure 21   Glandular hairs: unicellular stalk with multicellular head typical for *Plantago* sp. (a); uniseriate glandular trichomes of *Primula* sp. (b); oil gland typical for some members of Lamiaceae, stalk uniseriate, head with a multiple of two cells (c); and uniseriate glandular trichome with enlarged tip (d). (1) Same type as (a); (2) lateral view (2a) and surface view (2b) of hair: unicellular stalk and four cells forming the head, typical for *Digitalis* sp.; (3) unicellular trichome discernable on the rounded tip, typical for *Epilobium* sp.; (4) uniseriate glandular trichome, abundant type; (5) glandular trichome characteristic for Malvaceae, the head is alternatively one and two celled; (6) biseriate glandular trichome from Asteraceae; (7) uniseriate stalk (often very similar to covering trichomes) and multicellular head (irregular divided) very characteristic for Solanaceae; (8) multiseriate stalk and peltate head from *Cannabis sativa* L.; (9) the same as (c); and (10) surface view (10a) and cross section (10b) from complex glandular trichome from *Betula* sp. with a multicellular peltate head.

3.2   Parenchyma, Collenchyma, and Sclerenchyma

The three types of ground tissues – the parenchyma, the collenchyma, and the sclerenchyma – vary greatly in strength and flexibility.

Parenchyma is the most abundant ground tissue. Its cells are commonly thin-walled (thick-walled in the pith of herbal stems, Figure 23c) with a polyhedral shape, variable in morphology and physiology, with intercellular spaces present (Figure 23b). Parenchyma cells may be specialized with reference to, for example, photosynthesis (chlorenchyma) or

SECONDARY ROOT

– Transverse section

(a)

(b)

Rhytidome (outer bark)

Inner bark

Vascular cambium

Phellem

Phellogen

Phelloderm

Periderm

Xylem

Periderm

Cortex

Sec. phloem

(c)

(d)

Vessels

Medullary ray

Fibers

Figure 22 Schema of a transverse section of a secondary root forming rhytidome; segments of (a) vascular cambium; (b) periderm; (c) secondary xylem; and (d) outer part of the root.

storage of diverse substances (storage parenchyma). According to the position of the parenchyma, it is also referred to as bark parenchyma, wood parenchyma, pith parenchyma, pericarp parenchyma, endosperm (triploid parenchyma of angiosperm seeds originated by double fertilization), and so on.

Collenchyma cells have mechanical support function in plants, and occur commonly in herbaceous stems, and also in leaves and flower parts, predominantly located in the area beneath the epidermis. It is composed of more or less elongated cells with primary cell walls unevenly thickened (clearly visible in cross section, Figure 23d), pliable and nonlignified. Depending on the localization of the wall thickening, different types can be distinguished: *lamellar collenchyma* (thickening localized on the tangential walls), *angular collenchyma* (wall thickening developed in the corners of the cells with no intercellular channels in the tissue), and *lacunar collenchyma* (wall thickening localized around intercellular spaces). As all types usually are found together, the differentiation between the types can be neglected in terms of their value for crude drug identifications.

Figure 23 (a)–(c) Different types of parenchyma (transverse sections): (a) aerenchyma – enhances the diffusion of air within an organ and contains numerous, large, regular intercellular spaces; (b) parenchyma of cortex with fibers in intercellular spaces (storage parenchyma – contains storage substances such as starch, oil, proteins); (c) parenchyma cells of pith with wall thickening; (d) collenchyma; (e) fibers (longitudinal) and sclereids; and (f) fibers accompanied by calcium oxalate crystals.

Sclerenchyma cells possess hard, rigid, and usually lignified secondary walls and have a supporting and protective function. There are two types of sclerenchyma cells, the *sclereids* and the *fibers* (Figures 23e and f and 24).

Sclereids (= stone cells) are typically isodiametric or short cells, but of highly variable form, frequently with branching pit channels (Figure 24). Depending on their shape, they are named *brachysclereids* (ovoid to somewhat irregular shape, may occur single or clustered, Figure 24e), *macrosclereids* (elongated, rod-shaped, occurring in seed coats forming the palisade-like external boundary tissue), *astrosclereids* (star-like cells, with long arms, e.g., *Camellia sinensis* (L.) Kuntze leaf, *Illicium verum* Hook.f. fruit, Figure 24f), *osteosclereids* (bone-shaped, *Pisum* seat coat, Figure 24d), and *trichosclereids* (hair-like, sparsely branched, occurring in roots, stems and leafs of aquatic plants, e.g., *Monstera* sp. leaf, Figure 24c). Trichosclereids cannot be easily

differentiated from fibers. One remarkable distinction is branching. Sclereids are widely distributed in the plant kingdom and may be particularly often found in barks, roots, and the pericarp (= the whole tissue outside the seed(s)) of fruits and in seeds.

Fibers are long, spindle-shaped cells with thick secondary walls, often lignified containing simple pits. They may occur isolated or in bundles, often accompanying vascular bundles forming sheaths, often in combination with calcium oxalate crystals (Figure 23f). Commonly they are named after the tissue in which they occur (e.g., phloem fibers, xylem fibers, and cortical fibers).

For quality control in pharmacognosy, the sclerenchyma, that is, the distribution, abundance, size, and shape of sclereids and fibers, represents an important feature for the identification of drugs and their differentiation to others. Most of these cells have a birefringent wall and can be detected very easily with the polarizing microscope.

3.3 Vascular Tissues

The *vascular tissue system* represents the conductive system of the plant. It includes the *xylem*
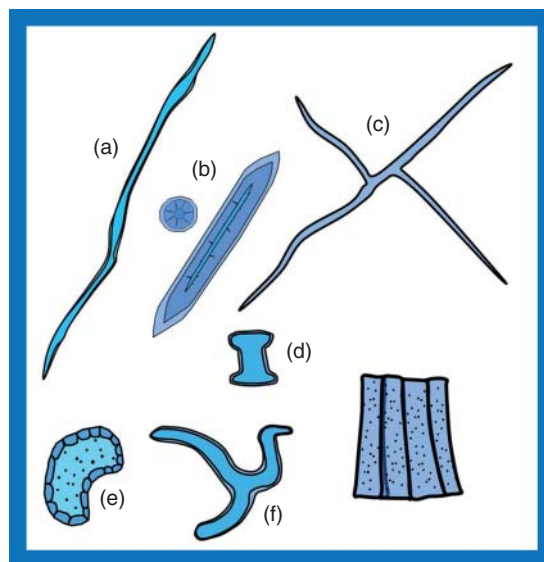
Figure 24 Sclereids and fibers: (a) fiber with uneven thick wall typically for *Althaeae officinalis* L. radix; (b) thick-walled fiber with distinct pits and laminated wall; (c) trichosclereid from *Monstera* sp.; (d) osteosclereid from the seed coat of Apiaceae; (e) Brachysclereid; (f) astrosclereid; and (g) macrosclereid from the seed coat of the members of Brassicaceae.

and the *phloem* – two complex tissues, each composed of more than one type of cells (Figure 25). The xylem carries water and dissolved minerals absorbed by the roots to the stem and leaves. The conducting cells are called *tracheary elements*, including *tracheids* and *vessel elements*. While tracheids – elongated, tapering cells with lignified cell walls and pit-pairs on their common walls that allow water to pass – occur within the whole plant kingdom, vessel elements cannot be found in the xylem of gymnosperms. In contrast to tracheids, vessel elements generally have a larger diameter, are shorter in length, and their common walls where they join end-on-end possess large perforations (perforation plate) or absent (Figure 26b). Several vessel elements together form a long tube called vessel. The differentiation between tracheids and vessels though is not always as clear as it is thought (see under: http://www.mobot.org/MOBOT/research/APweb/ and then choose CHARACTERS). In addition, the xylem contains parenchyma for food storage and sclerenchyma for supplementary mechanical support.

The phloem transports different types of oligosaccharides and nutrients produced during photosynthesis from the sources to the cells that consume or store them. The sieve elements are the conducting cells of the phloem. They include the *sieve-tube elements* with their *companion cells* lying next to them and deriving from the same immediate mother cell (angiosperms) or the *sieve cells* (gymnosperms and some primitive angiosperms). Sieve cells are associated with *Strasburger cells*, which are closely linked functionally with the sieve cells but are not formed from the same mother cell.

The sieve tube elements are living, tubular cells that are arranged end to end. These end walls, called *sieve plates*, contain large pores like a sieve, which enables the flow of solutes. Lacking a functioning nucleus, each sieve tube is accompanied by a companion cell that controls the activity of the sieve tube element.

Other cell types present in the phloem include parenchyma cells for storage and radial translocation of food substances and sclerenchyma fibers for mechanical support.

Xylem and phloem nearly always occur together, forming *vascular bundles*. According to the relative position of their constituent xylem and phloem, different types can be recognized (Figure 25).

- *Collateral vascular bundles*: The xylem and phloem tissues are opposed to each other in conjoint bundles with the xylem being the innermost and the phloem the outermost part in the stem or rhizome, both often accompanied by fibers. A fascicular cambium may be present between phloem and xylem, thus making the vascular bundle open (cambium present) or closed (cambium absent).
Collateral vascular bundles are the most frequent type in both gymnosperms and angiosperms.

- *Bicollateral vascular bundles* are also conjoint bundles similar to the collateral bundles but contain two strands of phloem on either sides of the xylem, being referred to as internal (toward the center) and external (toward the periphery of the central cylinder) phloem, each separated from the xylem by a cambium (sometimes the innermost cambium is missing).
Bicollateral vascular bundles are seen only in leaves of angiosperms with internal phloem in the stem, though they are not as common as the above-mentioned collateral vascular bundles (characteristic, e.g., Curcubitaceae, Gentianaceae, and Solanaceae). The term internal phloem in primary stem indicates the situation where phloem is found internally to the xylem but not included within the vascular bundle.

- *Concentric vascular bundles* are vascular bundles in which one vascular tissue completely encloses the other one. In *leptocentric* (= amphivasal) bundles, the xylem surrounds the phloem, common in rhizomes of monocot plants and a few dicotyledons. When the xylem is in the center, enclosed by the phloem, the vascular bundle is called *hadrocentric* (= amphicribal). This type occurs in pteridophytes (e.g., *Lycopodium*) and a few aquatic angiosperms.

- *Radial vascular bundles* are characteristic for roots. Xylem and phloem occur in separate groups (each group of xylem/phloem referred to as *arch*) alternating with each other.

For microscopic examinations, for one thing, the type of the vascular bundle(s) present and their arrangement are important for the definition of the plant organ. Furthermore, a first assignment to a group within the plant kingdom (such as monocots and dicots) may be possible:
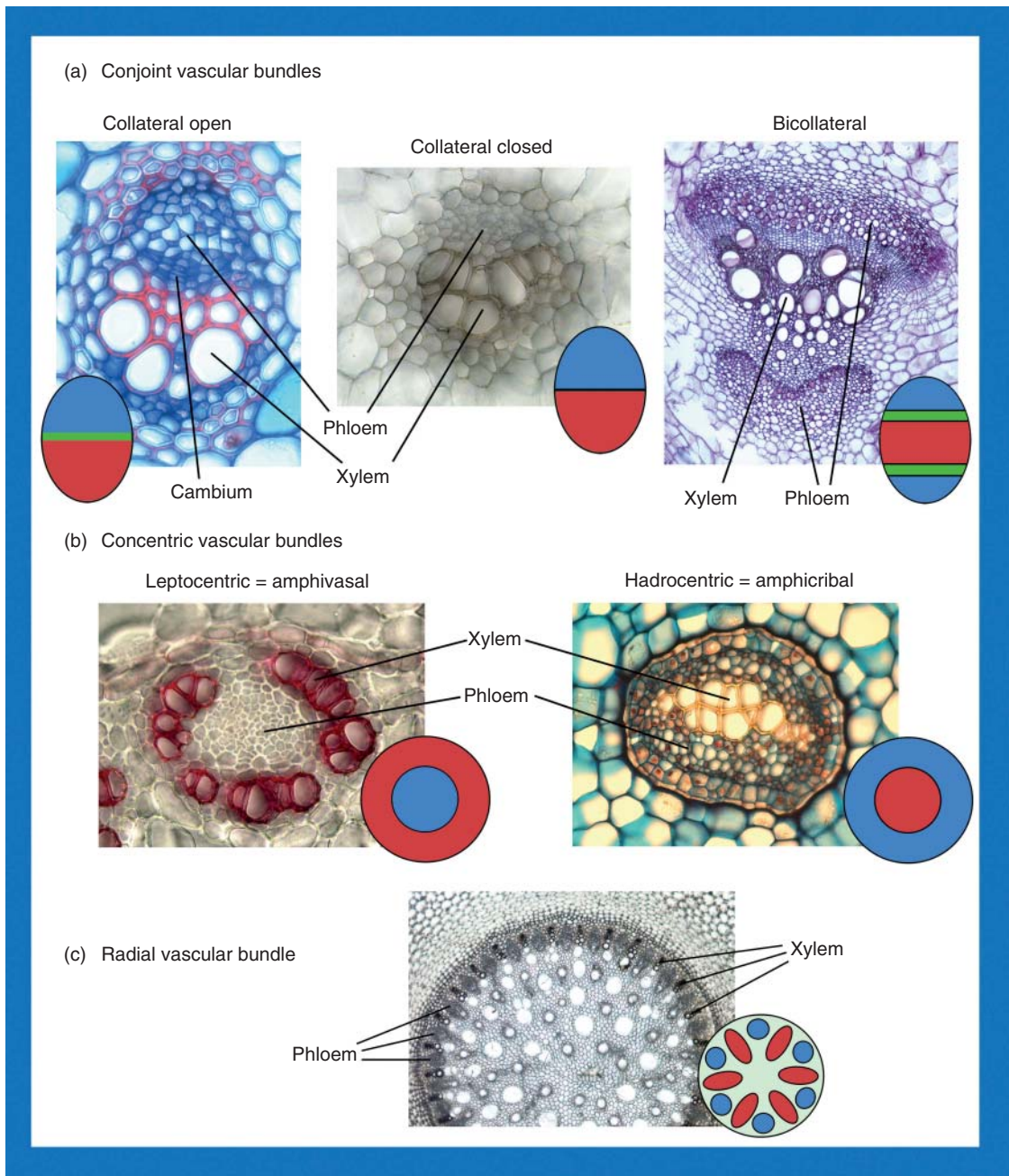
Figure 25 Different types of vascular bundles with picture and corresponding schemata (transverse sections); red = xylem, blue = phloem, and green = cambium.

- *Dicots*: Stem with one ring of open collateral bundles only. Root with two (= *diarch*) up to six (= *hexarch*) strands of xylem (the so-called protoxylem) alternating with phloem.

- *Monocots*: Aerial stem with many scattered closed collateral bundles; rhizomes with scattered leptocentric bundles in the center. Root commonly with more than six arches (*polyarch*).
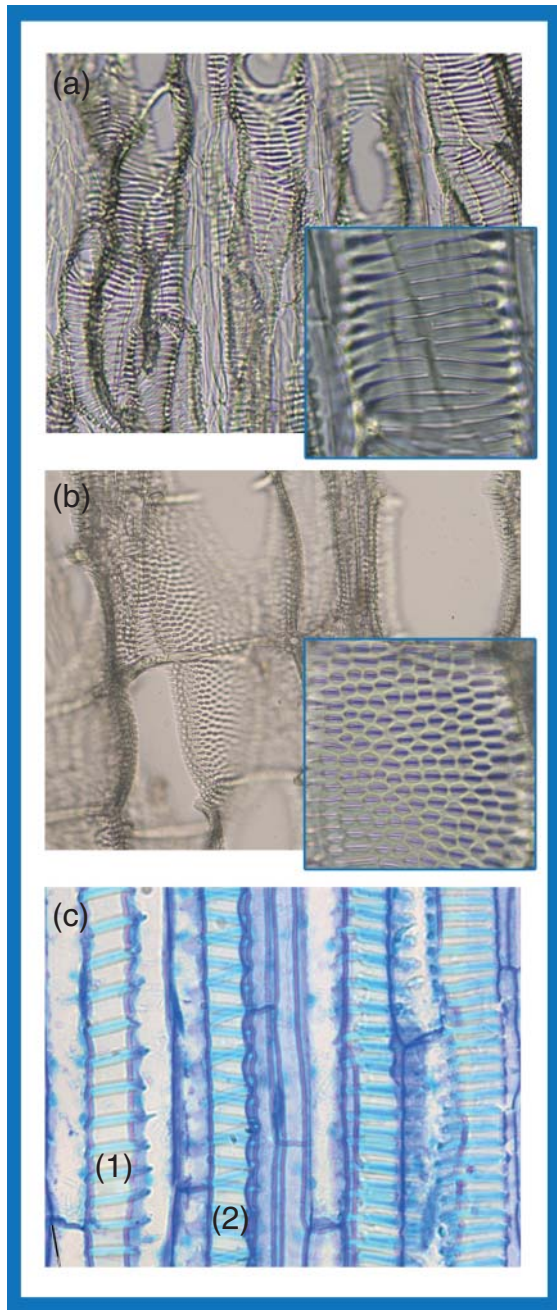
Figure 26 (a) Vessels with reticulate thickening; (b) vessels with bordered pits; and (c) vessels with annular (1) and helical (2) thickenings.

For another thing, the anatomical details of the single tissues provide useful information. The presence or absence of vessels usually indicates whether the botanical material in hand belongs to the gymnosperms or angiosperms.

The diverse types of secondary wall thickening of the vessel elements (Figure 26) – annular thickenings (secondary walls occur as rings), helical thickenings (spiral thickenings), scalariform thickenings (ladder-like, typically present in the fern group), and reticulate thickening (net-like, very common in dicotyledonous herbs) – have to be examined closely. In addition, two principal types of pits occur in cells with secondary walls and have to be considered: *simple pits* are canals in the secondary walls with straight sides, whereas with *bordered pits*, the secondary walls arches over the pit cavity and partially encloses the opening (pit *aperture*). The overhanging secondary wall constitutes the border. Together, the single features allow for further discrimination. Annular and helical thickenings are nonspecific – they may occur throughout the whole plant kingdom and are abundant in young organs. Some interesting works in this context are (Bliss, 1921; Carlquist, 2010; Jansen *et al*., 2004).

3.4 Secretory Structures

Secretion of diverse substances – either waste products or substances with special physiological functions – is a constant process in all plant cells. As a clear distinction between secretion and excretion cannot always be made, in this text, the term secretion refers to both of them. The visibly differentiated secretory structures, which are of diagnostical value in pharmacognosy, comprise a number of diverse cells and tissues, occurring within all tissue systems. Two basic works in this field are from (Tschirch, 1906) and (Fahn, 1988).

Particularly useful among the external secretory structures prove to be glandular trichomes varying greatly in their structure (mentioned earlier, compare glandular trichomes, Figure 21). Thus, the occurrence and the lack of a type of glands can provide evidence for the presence of a particular plant family or genus or species. Figure 21, for instance, shows a picture of a typical glandular hair of *Plantago* leaf (Figure 21a), a gland characteristic for Lamiaceae and some other families (Figures 11, 20, and 21c) and typical trichomes of Malvaceae (Figures 5a,b and 21).
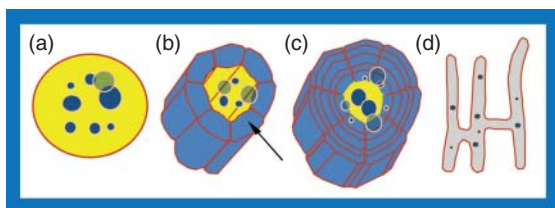
Figure 27 Important secretory structures: (a) oil cell with suberized cell wall; (b) schizogenous duct or cavity with the typical epithelial layer (arrow); (c) schizolysigenous cavity; and (d) simple laticifer.

The internal secretory structures range from *idioblasts* (single, isolated cells containing volatile oil, mucilage, calcium oxalate crystals, tannins, etc. (Foster, 1956), see Chapter 3.5 Cell Compounds) to large structures such as *secretory canals* (= ducts) and *cavities* or *laticifers*. The occurrence of oil cells (Figures 27a and 28c) is typical for families of Magnoliidae (basal group of dicots) and for some families of monocots (Acoraceae and Zingiberaceae). These cells comprise walls with suberin layers to avoid poisoning of the surrounding cells.

Secretory ducts and *cavities* closely resemble each other, only differing in the shape of their lumen (tubular intercellular spaces vs irregular spaces with no extensions). Usually, they are formed either schizogenously (Figures 27b and 28a and b) by the dissolution of the middle lamella and the enlargement of the resulting intercellular space because of separation of the adjacent cells (forming the so-called epithelial layer) or schizolysigenously through dissolution and autolysis of entire cells (Figure 27c). Thereby, the development is initially schizogenous and then lysigenous (compare (Curtis and Lersten, 1990; Fornasiero, Bianchi, and Pinetti, 1998)). For the identification of botanicals, the origin of the ducts is less important. What counts are the presence/absence of an epithelial layer and the appearance of the secretory channel. The anatomy and spatial position, the frequency of occurrence of the different types, along with the size of the secretory structures, and the size and shape of the epithelial cells provide valuable anatomical characters for differentiation of species (Fritz and Saukel, 2011a, 2011b).

Laticifers are specialized cells (single-celled laticifer = simple laticifer) or series of cells (compound laticifer) that contain a fluid with a milky appearance, which composition (and color) may vary widely in different species. On the basis of their structure and development, they are separated into two groups: *nonarticulated laticifers* are single, elongated cells with branching (e.g., *Euphorbia* sp., *Morus* sp., Figure 27d) or without branching (*Vinca* sp., *Urtica* sp., *Humulus* sp., and *Cannabis* sp.) but undergoing no fusions with other cells. *Articulated laticifers*, on the other hand, consist of many elongated cells, either anastomosing (e.g., tribe Cichorieae, *Carica papaya* L., and *Papaver* sp., Figure 28f) or nonanastomosing (e.g., *Chelidonium* sp. and *Allium* sp.). A good review on this topic can be found in Hagel, Yeung, and Facchini (2008) and Pickard (2008). Looking rather inconspicuous to an inexperienced observer, the presence and the arrangement of laticifers are quite important and therefore constitute a useful feature.

3.5   Cell Compounds

Some cell contents – referred to as *ergastic substances* (storage or waste products, products of metabolism) – play an important role in the identification and detection of adulterations in pharmacognosy, which, among others, include starch, proteins, fixed oil and fats, mucilages, resins and volatile oils, tannins, calcium oxalate, and calcium carbonate.

3.5.1   *Starch*

Starch (= amylum) is one principal storage polysaccharide in plants, thus providing energy for them. It consists of two fractions – amylose (long chains of glucose residues connected in α-1,4 linkage; unbranched; soluble in water; stains blue with iodine) and amylopectin (further α-1,6 linkages, one per 30 α-1,4 linkages; insoluble in water; stains brown with iodine). The ratio amylose : amylopectin varies according to the species with amylose usually accounting for 10–30% of the total structure of starch, whereas amylopectin constitutes the major compound with 70–90%. Starch occurs only in grains built up by the amyloplasts.

Starch grains are of varying shapes and sizes (Figure 29). A layering around the formation center (= hilum) as a result of the alternation of the two polysaccharides is often visible in light microscopy, faint or clear, in the form of a point or cleft, either concentric (e.g., graminaceous plants) or eccentric (e.g., potato) depending on the central or peripheral
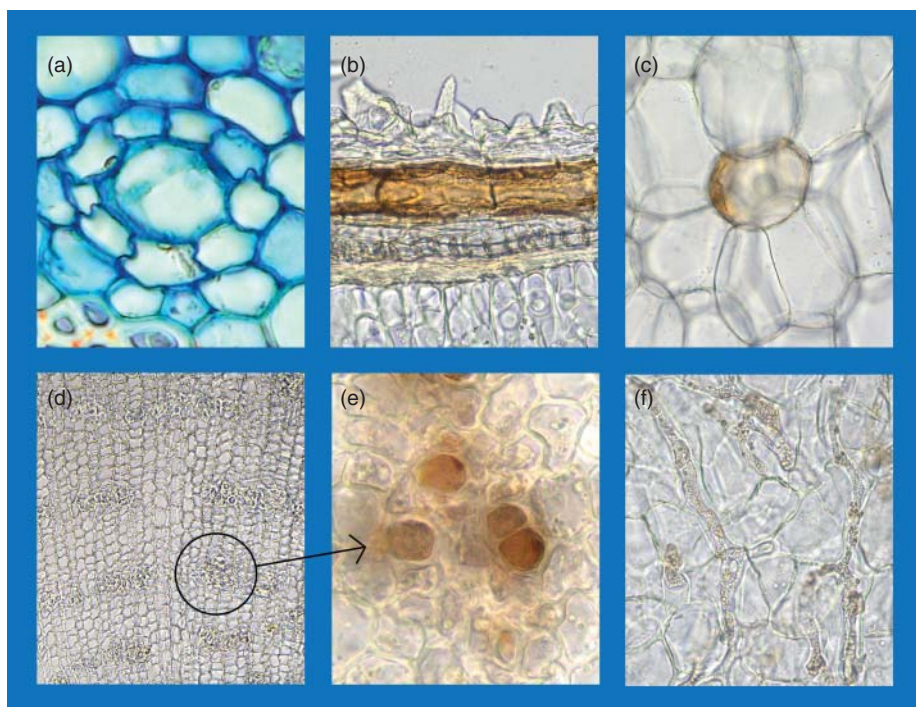
Figure 28   (a) Transverse section of secretory duct of *Anthriscus sylvestris* (L.) Hoffm. stem; (b) longitudinal section of secretory duct of anisum fruit; (c) oil cell of *Alpinia officinarum* Hance root; (d) overview of secondary phloem showing laticifers arranged in concentric circles from *Taraxacum sp*. root (transverse section); and (e) transverse and (f) longitudinal sections of laticifers typical for species of the tribe Cichorieae.

location of the formation center. Starch grains may occur single, compound (more than one formation center within one amyloplast), or in clusters.

For microscopic examination, starch grains first ought to be surveyed with bright field and under polarized light (Figure 29). Under polarized light (Figure 29e), the grains appear bright against the dark background and show a dark Maltese cross with its center at the hilum. Figure 29f demonstrates the use of a retardation plate, the grains are beautifully colored and the background is purplish-red. Furthermore, especially for identifying starch if one does not clearly see starch grains, the starch may be stained with iodine (the amylose : amylopectin – ratio being responsible for the resulting blue to bluish-brown, Figure 29c). Each contrasting method depicts other aspects of the observed structures.

Very similar to starch but with a quite different appearance is inulin, the second storage polysaccharide, used in some parts of the plant kingdom only. It is found extensively in the roots and rhizomes of Asteraceae. Inulin is insoluble in a mixture of alcohol and glycerol. If fresh material is placed in such a solution, inulin will precipitate in the form of sphaerocrystals within the cell.

3.5.2   Aleurone

Aleurone grains (Figure 30) are composed of protein-crystalloids, globoids, and a protein matrix, surrounded by a protoplasmic membrane. Calcium oxalate may occur as single crystals or as crystal rosettes. Aleurone grains are widespread in seeds (e.g., *Ricinus communis* L. and *Illicium verum*, Figure 30), particularly oily seeds (e.g., Umbelliferae). For example, seeds of Poaceae and Polygonaceae show aleurone cells (cells that contain aleurone grains) forming the so-called aleurone layer, a layer covering the surface of the endosperm. The form and size of the grains are often characteristic for a species. Aleurone grains stain yellowish-brown with iodine and thus can be easily distinguished from starch this way.
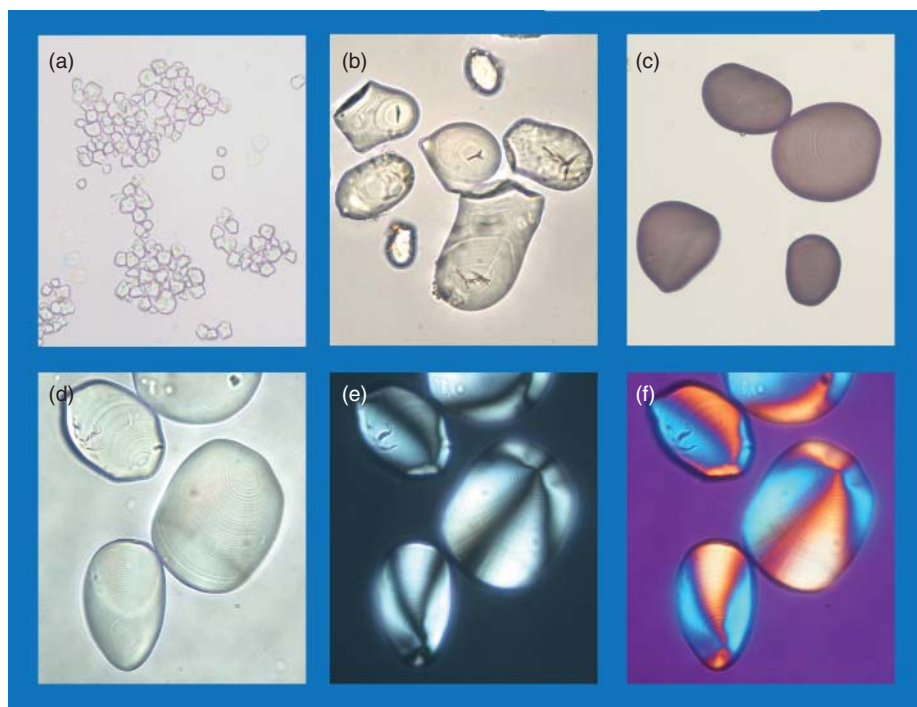
Figure 29   (a) Amylum Oryzae polyhedral grains with sharp angles, hilum as central point, no striations; (b) Amylum Sago: the hilum appears dark in bright field; (c)–(f) Amylum Canna: single grains with eccentric layering: (c) stained with iodine, (d) bright field, (e) polarized light, and (f) polarized light with compensator first order.
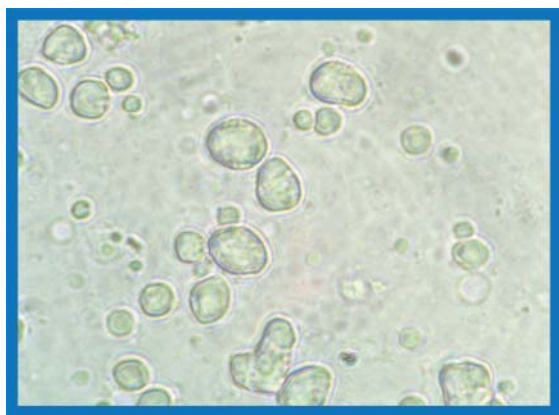
Figure 30   Aleuron grains of *Illicium verum* seed.

3.5.3   Crystals

● Calcium oxalate: Calcium oxalate crystals are mainly found not only in the cytoplasm but also within the cell wall or cuticle. Because of their abundance and importance in systematics, there are a great number of publications concerning the development, distribution, appearance, and function of calcium oxalate crystals (Franceschi and Horner, 1980; Franceschi and Nakata, 2005; Horner, Franceschi, and Hill, 1978; Prychid and Rudall, 1999; Sharawy, 2004). According to their shape, they are divided into different categories, the most common are mentioned in Figure 30. Calcium oxalate crystals are important features for identification. One first approach for a terminology can be found in (Haberlandt, 1914). One ought to note not only the type of crystals and the tissue they are within, but the exact shape, the size, the distribution, and the frequency. They may occur single or in rows, along the vascular bundles or exactly the opposite (e.g., along the nerves of leaves or within the intercostal region), or form a sheath around fibers (rectangular cells form elongated threads, each bearing a *prismatic crystal* or a *druse* (Netolitzky, 1908), Figures 23f and 31). While prismatic crystals are quite common in the plant kingdom, some kinds are specific for only a few plant families – *raphides*, for

instance, appear within Geraniaceae, Rutaceae, Oleaceae, Rubiaceae, Phytolaccaceae, Onagraceae, and some others (Netolitzky, 1905); *styloids* may be found in Iridaceae, Ruscaceae, and Melianthaceae; and *crystal sand*, jam-packed within parenchyma cells of leaves, is typical for, for example, Solanaceae (*Atropa belladonna* L.) and Caprifoliaceae (*Sambucus* sp.). Working with

Figure 31   The most common types of calcium oxalate crystals.
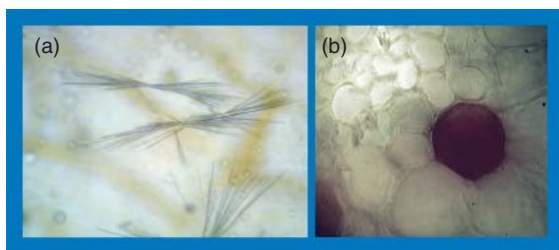
Figure 32 (a) Detection of calcium oxalate with sulfuric acid: calcium oxalate dissolves with formation of needles of gypsum and (b) tannin sac in *Sambucus* sp.

polarized light can assist in detecting crystals. The treatment of a specimen with sulfuric acid leads to the formation of gypsum needles, which may be used as proof for the calcium oxalate as the present kind of crystal (Figure 32).

• Hesperidin and diosmin: The flavonoid glycosides diosmin and hesperidin sometimes occur in crystalline form as sphaerocrystalline masses. In contrast to calcium oxalate crystals, they are not visible under polarized light and dissolve slowly with a solution of potassium hydroxide giving yellow color.

3.5.4 Cystoliths

Cystoliths are concretions of calcium carbonate together with the material of cell walls. Some epidermal cells (lithocysts) do not undergo periclinal divisions but exhibit local growth and expand to the inside, resulting into an irregular concretion composed mainly of calcium carbonate attached to a stalk. Cystoliths are confined to just a few families of the seed plants (Boraginaceae, Cannabaceae, Plantaginaceae, Solanaceae, Urticaceae, Moraceae, Acanthaceae, and Combretaceae), and therefore, can be quite helpful for microscopic identification. Figure 33a shows a typically cystolith of *Urtica urens* L. leaf. Cystoliths have no birefringent properties.

3.5.5 Tannins

Tannins are polyphenolic compounds that are widely distributed throughout the whole plant kingdom, usually located within the vacuoles of the cells or the cell wall matrix. Tannins are soluble in water;

however, they tend to oxidize and polymerize to brown and reddish-brown phlobaphenes, in which case, they get insoluble. Tannins give a positive reaction with a solution of $FeCl_3$ (stains bluish-black or greenish-black depending on concentration of $FeCl_3$ and composition of tannins). In some families, tannin sacs occur in stem or leaf (*Sambucus* sp., Figure 32b).

3.5.6 Volatile Oils

Volatile oils are mixtures of terpenes and terpenoids that can be found in special secretion cells, secretory canals, or cavities within the whole plant as small oil droplets (see Section 3.4). Frequently, volatile oils are associated with other substances such as resins, gums, benzoic, or cinnamic acids (oleoresins and gum resins). Volatile oils dissolve in alcohol and stain orange-red with Sudan.

3.5.7 Fixed Oils

Fixed oils, also called *fatty oils*, are esters of fatty acids and may appear as liquid droplets throughout all plant tissues. Most abundantly, they occur in seeds and fruits (e.g., endosperm of umbelliferous fruits) but are also found in leaves of many angiosperms as long forgotten features – the so-called oil *bodies* (Lersten *et al.*, 2006). Freehand cross sections of fresh leaves stained with Sudan IV can verify the presence of oil bodies. This for leaves newly rediscovered organelles are responsible for many droplets of fixed oils in chloral hydrate specimens.

3.5.8 Mucilage

Mucilage is a polysaccharide complex that is believed to help the plant to retain water (e.g., common in many succulent plant species, e.g., *Aloe* sp.) and swells or dissolves in water (Evert, 2006). Under the microscope, stratification is often visible. Frequently, mucilage occurs in cells containing raphides (e.g., *Epilobium* sp.), this being the case when staining with corallin. In other cases, mucilage may be verified by staining with a solution of *methylene* blue (Figure 33b) or using Chinese ink. Dry powder is mixed with the black-colored Chinese ink. After a short time, the particles of the Chinese ink will be displaced by the hydrated mucilage.
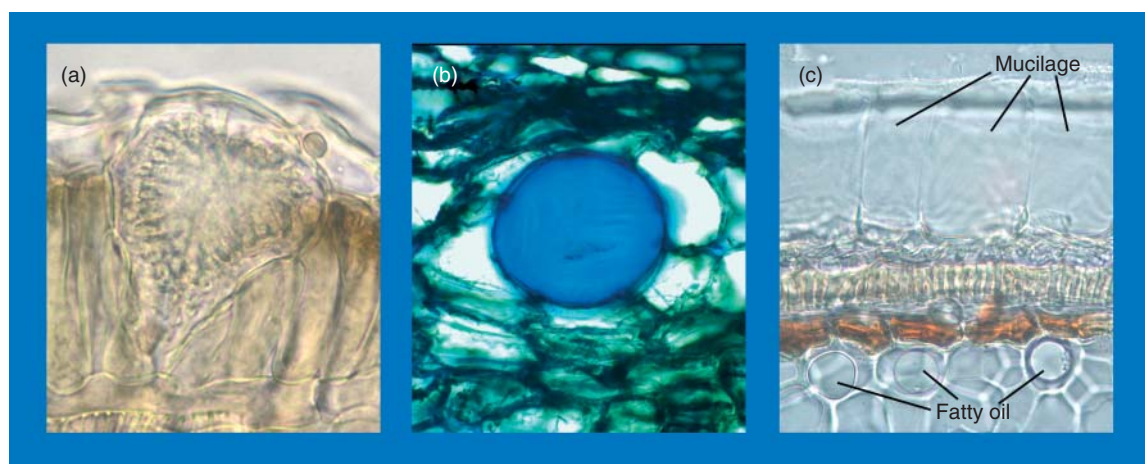
Figure 33 (a) Cystolith in the epidermis of *Urtica* leaf; (b) mucilage in *Hibiscus sabdariffa* L. stained with methylene blue; and (c) transverse section of *Linum* seed: mucilage in the epidermis and fatty oil droplets within the endosperm (disturbed oil bodies).

4 EXAMINATION OF CRUDE DRUG MATERIALS

The proof of the identity of plant species is different to the proof of a chemical substance. Chemicals have some very specific features such as spectral characteristics, melting point, and specific weight. In contrast, botanical material is harvested from living populations with sometimes very significant intraspecific variations depending on location, harvesting time and method, genetic diversification, and many other reasons. In addition, many of our herbals are contaminated by alien plant material; therefore, a great experience during the testing process is necessary. Only if a complete characterization of a drug is available, it is possible to detect adulterations at once and prevent poisoning or reduction in quality of the material in question. Guiding schemes for a complete description of botanical drugs can be found in an old but very exemplary work (Moll and Janssonius, 1923, pp. 25–45). Therein about 100 detailed, the so-called pen-portraits are available, which provide the possibility to compare the listed drugs in a perfect manner. For the beginner, it is very important to compare all drugs with authentic specimens in a living or preserved state of good quality. The comparison with fresh material from a botanical garden or with well-defined herbal specimens or retained samples is indispensable! A skilled investigator can correctly identify a large number of herbals by long experience and familiarity with their appearance and organoleptic properties. In our literature overview, the reader finds a notable number of books mostly available via the Web site *internet archive* (http://archive.org). Nearly all classical books on the field of microscopic analyses of botanicals (drugs as well as vegetable food, mushrooms, algae, and lichens) with often very helpful drawings or photographs were published between 1865 and 1945, most of them are available via internet (see Table 2 and cited literature).

Nevertheless, the necessary skills for a proper identification are difficult to learn solely based on written description: a personal teaching is always the best.

The standard criteria for the characterization of drugs are:

- Macroscopic appearance – should lead to a morphological interpretation of which parts of plants are present. Very helpful is the use of a stereo microscope (*see* optical introduction) to gain information on surface structures (trichomes and fibers), which go beyond the optical range of the naked eye.
- Organoleptic properties – these properties are very important for the first glance and a very quick associatively guided identification of drugs known to the investigator in addition to the detection of impurities.
- Anatomical (oxalate crystals, vessels, etc.) and the so-called fine morphological (surface of leaves with stomata and trichomes) characters visible by

Table 2 Important books in the field of pharmacognosy and microscopy available in the internet.

Author, title	Year	Url
Louis Braemer, Armand Suis, Atlas de photomicrographie des plantes médicinales, Paris: VIGOT FRERES, EDITEURS	1900	http://archive.org/details/atlasdephotomicr00brae
Friedrich Czapek, Biochemie der Pflanzen, zweiter Band, Verlag von Gustav Fischer in Jena	1905	http://archive.org/details/biochemiederp02czap
A. De Bary, Comparative Anatomy of the Vegetative Organs of the Phanerogams and Ferns. Oxford, The Clarendon Press	1884	http://archive.org/details/comparativeanato00baryuoft
Cornelius A. J. A. Oudemans, Handleiding tot de Pharmacognosie van het Planten-en Dierenrijk, tweede druk, Amsterdam, C.L. Brinkman	1880	http://archive.org/details/handleidingtotd00oudegoog
Friedrich A. Flückiger, Alexander Tschirch, Frederick B. Power, the Principles of Pharmacognosy, An Introduction to the Study of the Crude Substances of the Vegetable Kingdom. New York, William Wood & Company	1887	http://archive.org/details/mobot31753000258092
Friedrich August Flückiger, Alexander Tschirch, The Principles of Pharmacognosy (1887), William Wood & Company	1887	http://archive.org/details/principlespharm00tschgoog
Ernst Gilg, Lehrbuch der Pharmakognosie, Berlin, Verlag von Julius Springer	1905	http://archive.org/details/lehrbuchderphar00brangoog
K.Goebel, Organographie der Pflanze	1898	http://archive.org/details/organographieder02goebuoft
K. Goebel, Isaac Bayley Balfour, Organography of Plants Especially of the Archegoniatae and Spermophyta Part II Special Organography, Oxford, The Clarendon Press	1905	http://archive.org/details/organographyofpl02goeb
K. Goebel, Organographie der Pflanzen insbesondere der Archegoniaten und Samenpflanzen, erster Teil Allgemeine Organographie, Jena Verlag von Gustav Fischer	1913	http://archive.org/details/organographieder01goebuoft
George L. Goodale, Gray's botanical text-book (6th edn) VOL II Physiological Botany I Outlines of the Histology of Phaenogamous Plants, II Vegetable Physiology, New York Cincinnati Chicago American Book Company	1885	http://archive.org/details/graysbotanicalt02gray
T.F. Hanauseck, revised by Andrew L. Winton, Collaboration Kate G. Barber, 1st edn., New York, John Wiley & Sons	1907	http://archive.org/details/microscopytechn03hanagoog
Henry Kraemer, Scientific and Applied Pharmacognosy, New York, John Wiley & Sons	1920	http://archive.org/details/scientificapplie00krae
John Uri Lloyd, History of the Vegetable Drugs of the Pharmacopeia of the United States, in Bulletin of the Lloyd Library of Botany, Pharmacy and Materia Medica, Bulletin No. 18	1911	http://archive.org/details/mobot31753000007200
John Uri Lloyd, Origin and History of All the Pharmacopeial Vegetable Drugs, Chemicals and Preparations with Bibliography, Vol. I, Vegetable Drugs, Cincinnati, The Caxton Press	1921	http://archive.org/details/originhistoryofa01ameriala
William Mansfield, Histology of Medicinal Plants, New York, John Wiley & Sons, inc.; (etc.)	1916	http://archive.org/details/histologymedici00mansgoog
Hugo von Mohl, Mikrographie oder Anleitung zur Kenntniss und zum Gebrauche des Mikroskopes, Tübingen Verlag und Druck von L.G.Fues	1846	http://archive.org/details/mikrographieode00mohlgoog
Hans Molisch, Grundriss einer Histochemie der pflanzlichen genussmittel, Jena Verlag von Gustav Fischer	1891	http://archive.org/details/grundrisseinerh00moligoog
Hans Molisch, Anatomie der Pflanze, Jena Verlag von Gustav Fischer	1920	http://archive.org/details/anatomiedespfla00moligoog
H.H. Janssonius, Mikrographie des Holzes der auf Java vorkommenden Baumarten, Bd. I, Leiden, E.J. Brill	1906	http://archive.org/details/mikrographiedesh01jans
H.H. Janssonius, Mikrographie des Holzes der auf Java vorkommenden Baumarten, Bd. III, Leiden, E.J. Brill, 1914	1914	http://archive.org/details/mikrographiedesh03jans
H.H. Janssonius, Mikrographie des Holzes der auf Java vorkommenden Baumarten, Bd. IV, Leiden, E.J. Brill	1920	http://archive.org/details/mikrographiedesh04jans

Reference	Year	URL
H.H. Janssonius, Mikrographie des Holzes der auf Java vorkommenden Baumarten, Bd. VI, Leiden, E.J. Brill	1936	http://archive.org/details/mikrographiedesh06jans
H.H. Janssonius, Mikrographie des Holzes der auf Java vorkommenden Baumarten, Bd. VII, Gesamtregister zu Bd. I bis VI, Leiden E.J.Brill	1936	http://archive.org/details/mikrographiedesh07jans
Lucius E. Sayre, A Manual of Organic Materia Medica and Pharmacognosy, 4th edn, Philadelphia, P Blakiston's Son & Co	1917	http://archive.org/details/manualoforganicm00sayruoft
Smith Ely Jelliffe, An Introduction to Pharmacognosy. Philadelphia, New York, London, W.B. Saunders and Company,	1904	http://archive.org/details/introductiontoph00jelluoft
William Chase Stevens, Plant Anatomy from the Standpoint of the Development and Function of the Tissues and Handbook of Micro-Technic, Philadelphia, P. Blakiston's Son & Co.	1916	http://archive.org/details/plantanatomyfro00stevgoog
Eduard Strasburger, A.B. Hervey, A Manual of the Microscope in Vegetable Histology, Boston Samuel E. Cassino	1887	http://archive.org/details/microscopicbota00stragoog
Eduard Strasburger, Die Angiospermen und die Gymnospermen, Jena Verlag von Gustav Fischer	1879	http://archive.org/details/disangiospermenu00strauoft
Hermann Thoms, Handbuch der praktischen und wissenschaftlichen Pharmazie, Bd. V erste Hälfte, Botanik und Drogenkunde, Urban und Schwarzenberg Berlin Wien	1929	http://digital.ub.uni-duesseldorf.de/vester/content/titleinfo/1720739
Hermann Thoms, Handbuch der praktischen und wissenschaftlichen Pharmazie, Bd. V zweite Hälfte, Botanik und Drogenkunde, Urban und Schwarzenberg Berlin Wien	1931	http://digital.ub.uni-duesseldorf.de/vester/content/titleinfo/1720740
Alexander Tschirch, Die Harze und die Harzbehälter mit Einschluss der Milchsäfte, zweite stark erweiterte Auflage, Zweiter Band, Leipzig Verlag von Gebrüder Bornträger	1906	http://archive.org/details/dieharzeunddieh00tschgoog
Alexander Tschirch, Handbuch der Pharmakognosie, Erster Band, Leipzig, Verlag von Chr.Herm.Tauchnitz	1909	http://archive.org/details/handbuchderpharm1remy
Alexander Tschirch, Handbuch der Pharmakognosie, Zweiter Band, Leipzig, Verlag von Chr.Herm.Tauchnitz	1912	http://archive.org/details/handbuchderpharm21remy
Julien Vesque, Memoire sur lánatomie comparee de Lécorce, Theses, Paris G. Masson, Editeur, Libraire de LÁcademie de Medecine	1876	http://archive.org/details/mmoiresurlanato00vesqgoog
Julien Vesque, Epharmosis, sive, Materiae ad instruendam anatomiam systematis naturalis, Vincennes	1889	http://archive.org/details/cbarchive_52245_epharmosis1889
Julien Vesque, Epharmosis, sive, Materiae ad instruendam anatomiam systematis naturalis	1894	http://www.archive.org/details/mobot31753002555628
Otto August Wall, Notes on Pharmacognosy, St. Louis, Mo., Aug. Gast Bank Note and Litho. Co.	1902	http://archive.org/details/notesonpharmacog00wall
Alex. J. Wedderburn, Report on the Extent and Character of Food and Drug Adulteration, Published by Order of Congress	1894	http://archive.org/details/reportonextenta00weddgoog

microscopy only. Immediately visible in preparations of powders or after sectioning of drug material in question.

- Microchemical proof for some characteristic parts/substances (e.g., iodine solution and starch detection, etc.).
- Phytochemical characterization (this field is not discussed further in this chapter).

Besides some technical details, it should be pointed out that there is a necessity for more than one identification criterion anyway! Commonly, a collection of criteria with a lot of well-defined sub-criteria is used (*also see* Plant Tissues), and the sample in question must comply with all of them definitively. Therefore, the observation that the sample in question contains fibers or oxalate crystals or unicellular trichomes is not precise enough. Many questions should arise by the examination: How is the appearance of the fibers in the cross section? Are they equally thick walled, is the wall lamellate or not? What observations are possible in the longitudinal view? What value has the ratio between length and width, are the fibers pitted or not, is the wall equally thickened on the whole length?

One useful approach for collecting experience is the comparison of the images from the stereo microscope with the respective images from the same part of the investigated drug sample gained from the optical microscope. This approach ensures a meaningful use of the stereo microscope and will fasten the identification process. It is also strongly recommended to apply the nearly same procedure with different magnifications in normal microscopic work, first carefully inspecting at a high magnification and after this going back to a low magnification with the question in mind – what structures are visible now? After
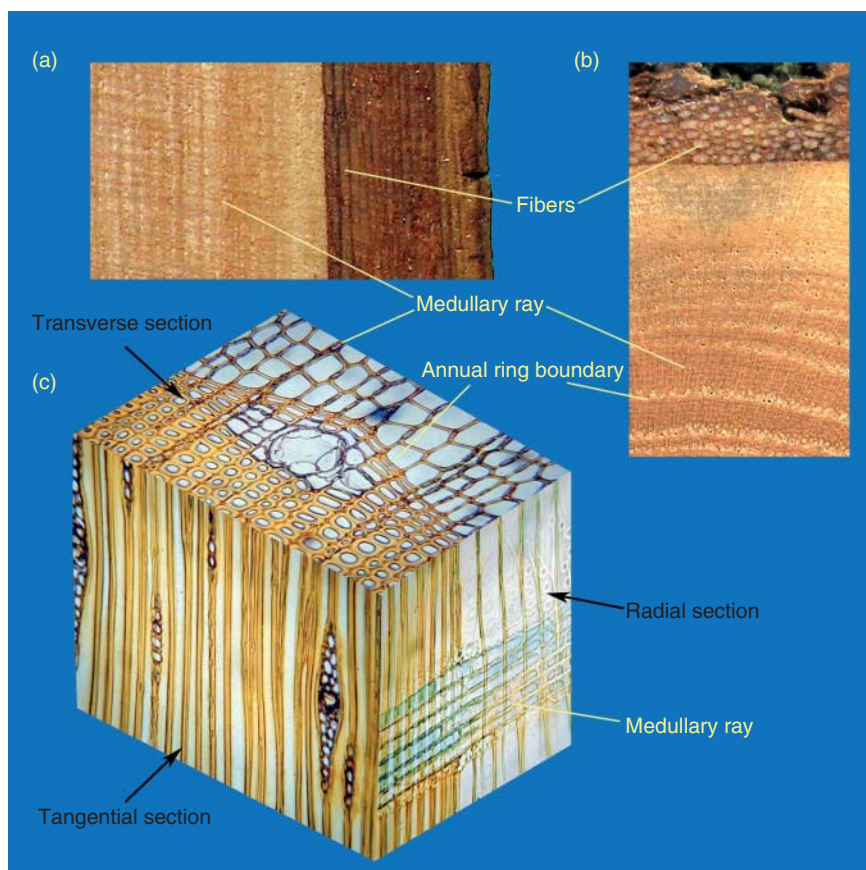
Figure 34   (a) Longitudinal view; (b) cross section of a wood; and (c) three-dimensional appearance of pinewood.

some repetitions, surprisingly, one can see at a lower magnification fine morphological and/or anatomical features that are actually too small for the resolution of the used objective (normal or stereo microscope) but our brain has learned to recognize those structures at low magnification too! A further point should be kept in mind – microscopical imaging of any kind is a two-dimensional imaging! Every three-dimensional imagination of the object in question must be carried out by our brain. The brain needs to superimpose many two-dimensional cross-sectional images to produce a three-dimensional structure of the specimen. We can provide an endorsement for our imagination by comparing cross sections and tangential or radial section of the same object (Figure 34). Note: It takes a long time to gain the necessary experience for proper microscopic drug differentiations.

Only if all aspects of the identification of a pure crude drug are internalized, the detection of adulterations will be meaningful and possible. Adulterations are all parts in a sample not included in the definition of a drug (e.g., the monograph if existing), in other words, all particles from the intended species not included in the monograph or description. *Intended adulterations* often occur when good quality material is mixed with or substituted by products of bad quality. *Unintended adulterations* comprise all materials stemming from the harvesting processes such as mould, dirt, stones, or an infestation by unwanted fungi (Figure 35), insects, mites, or other pests. The contamination limits need to be declared for every drug (*see* European Pharmacopoeia).

Figure 35 Unwanted fungi (arrow) in a preparation of *Plantago* leaf.

For the debutant, the realization of the correct type of plant organ is the base for any comparison with literature of the examined drug in question. This is a very crucial point! Hereafter, the reader can find some bundles of features to employ on this recognition.

4.1 Leaves

Leaves are mainly flat and green organs (but keep in mind leaves from the genus *Allium*), with a great variety of shapes and dissection of the margin (margin entire, toothed, or lobed; leaves palmate or pinnate; etc.). Only specialized leaves have a differenced outline, for example, rosemary or thyme leaves with revolute margins that create their needle-like shape or the so-called *unifacial leaves* with either the originally upper or the lower face as new outer surface. Unifacial leaves are common in the families of Pinaceae, Alliaceae, Acoraceae, Iridiaceae, and so on (Figure 36). In contrast to the normally constructed leaves referred to as *bifacial leaves*, unifacial leaves show a circle of vascular bundles in the case of round-shaped leaves or two rows of vascular bundles in the case of flat unifacial leaves (e.g., *Iris* sp. and *Acorus* sp.). Moreover, every leaf with a distinct petiole also shows, within the petiole, a more or less unifacial construction (Figure 36, right bottom corner). This may lead to misinterpretations (leaf vs primary stem). The epidermal cells are sometimes conspicuously striated (*Atropa belladonna* L.), with stomata, adaxial, or/and abaxial distributed, and often with trichomes of different but mostly characteristic outline (Figures 18 and 19). The *mesophyll* (ground tissue in between the epidermis) is mostly differentiated in *spongy parenchyma* and *palisade parenchyma* (Figure 37c). Sometimes the *spongy parenchyma* is developed as *aerenchyma* (*Menyanthes* sp. and *Tussilago* sp., Figure 37d). The position of the mesophyll together with the arrangement of the vascular bundles serve as diagnostic leaf characters, visible in cross sections only. The cross section of the petiole and the major veins can also be very diagnostic (Haron and Moore, 1996; Salmaki *et al.*, 2011). One often crucial question for the beginner is the determination of the abaxial face. Only after this decision is the information about the distribution of stomata meaningful. Most common leaves have a groove on the adaxial face at the position of the main venation (Figures 36 and 38) but a rip on the abaxial face.
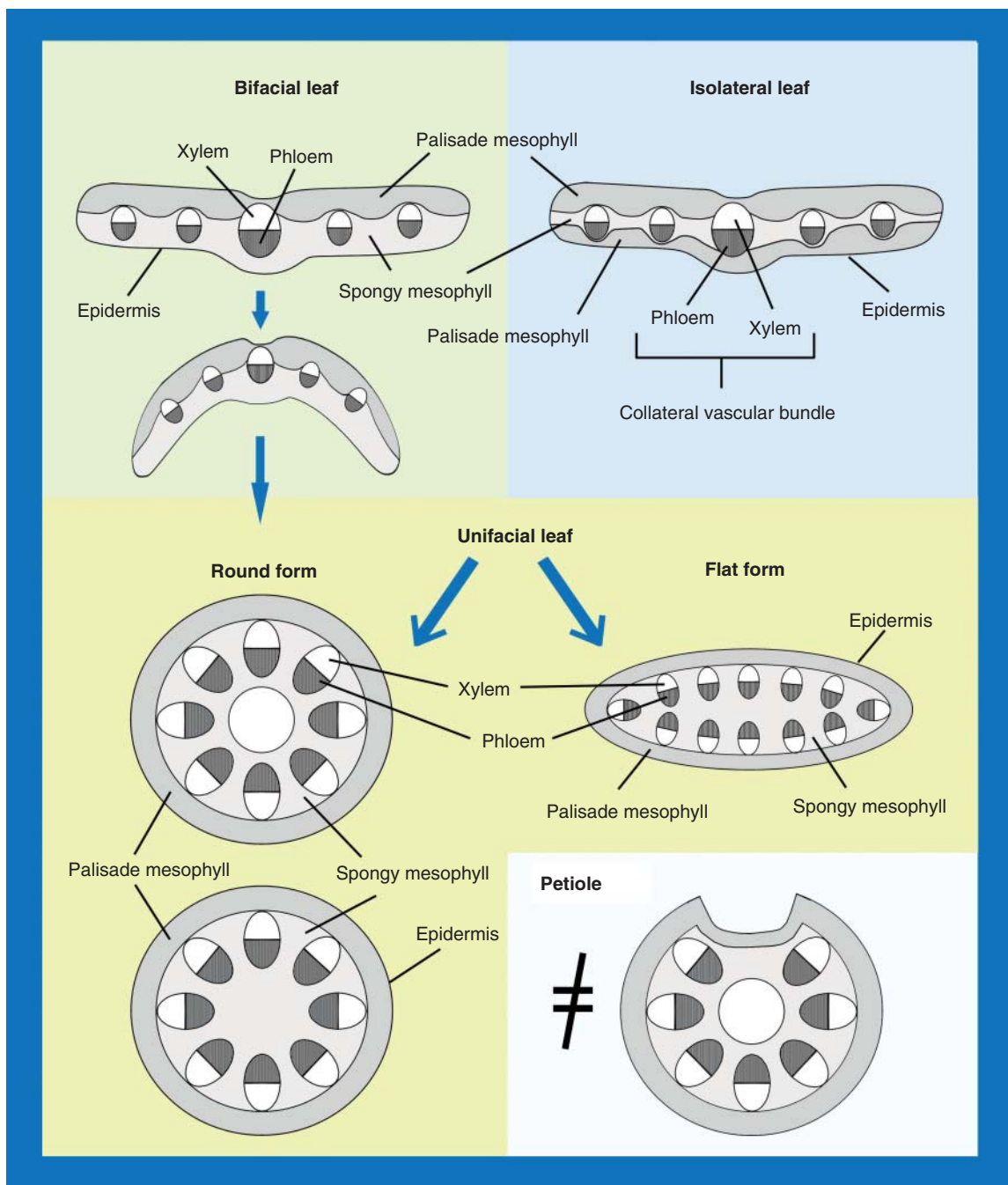
Figure 36   Cross sections of different leaf types.
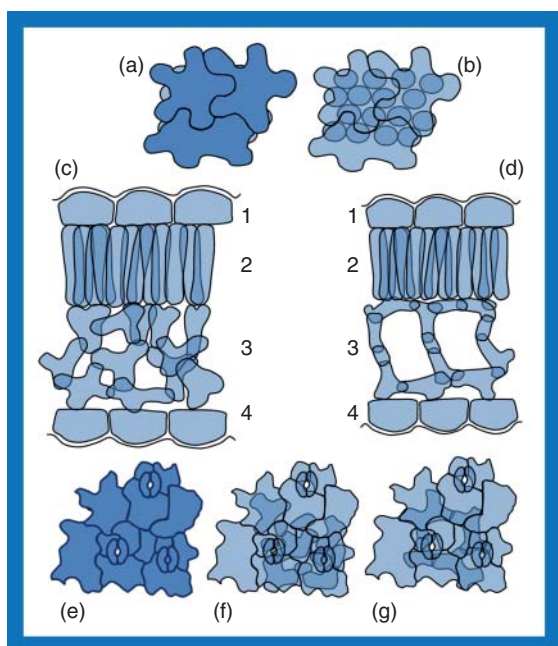
Figure 37 Cross sections and surface views of a bifacial leaf: (a) adaxial epidermis in focus; (b) palisade parenchyma below the epidermis in focus; (c) leaf with normal spongy parenchyma; (d) leaf with an aerenchyma; 1 adaxially epidermis; 2 palisade parenchyma; 3 spongy parenchyma (c) aerenchyma (d); 4 abaxial epidermis; (e) surface view of the abaxial surface with stomata; (f) spongy parenchyma in focus; and (g) aerenchyma in focus.

Therefore, the freshman has to look which part of the leaf surface is in focus first if the stage is moved upward, the intercostal region (Figure 38b) or the epidermis above the major veins (Figure 38a). Furthermore, the microscopist has to look at the cell layer beneath the epidermis. In the case of a bifacial leaf and the adaxial side, one should have the palisade parenchyma in focus (Figure 37b). Otherwise, the spongy parenchyma should come in focus (Figure 37f).

As annotated before, the type of stomatal complex, the overall number of stomata, and the distribution between the adaxial and abaxial sides can be of interest – *amphistomatic* (stomata on both sides), *hypostomatic* (stomata on abaxial side only), and *epistomatic* (stomata on adaxial side only) leaves.

Besides the stomata type, the overall number (absolute and relative) of stomata on the leaf surfaces can be of interest too. The *stomatal index* (SI) is one possibility to obtain comparable relative values.

$$SI = 100 * Stnr/(Stnr * Ecnr)$$

Stnr = number of stomata in a given area of leaf

Ecnr = number of epidermal cells including trichomes.

4.2 Flowers

Flowers are very typical structures of the angiosperm plant group. Common flowers consist of four kinds of leaf derivates attached to the tip of a short stalk and are sometimes arranged in flower heads or in capitula (Figure 39). Each of these kinds of parts is mostly arranged in a whorl on the receptacle. The main whorls are the *perianth* often differentiated in *calyx* (= *sepals*) and *corolla* (= *petals*) or the leafs are only from one type (= *perigon* with *tepals*), the *androecium* and the *gynoecium*. Depending on the absence or presence of one or both of the last two parts, the flowers may be monoecious or dioecious. The morphological and anatomical properties of botanicals derived from flowers vary largely. Having a look at Figure 39, one can easily imagine a lot of different structures visible in a microscopical slide (different leaf types, the stylus, the stigma, and so on). The epidermis of all leaf derivates concerned often with stomata (calyx and sepals), often with trichomes (calyx and corolla, and petals and tepals). In calyx and corolla, the mesophyll is often notably reduced, containing chloroplasts (calyx only) and sometimes colored vacuoles or chromoplasts (in the case of corolla). The chromoplasts often contain crystals of carotene (e.g., *Calendula* sp.). The anatomical characters of drugs derived from flowers and flowering tops depend above all on the number of different parts (single flowers versus capitula (Asteraceae) or whole inflorescences (e.g., *Tilia* sp.) that are present in the drug (compare (Saukel, 1984c)). If the drug consists simply of petals or/and sepals, the chief diagnostic characters will be found in the shape of the stomata, and in the papillae, on the surface (petals). Some petals are furnished with simple and glandular hairs in which case these may be utilized in establishing the identity of the powder (e.g., *Althaea* sp. and *Arnica* sp. (Saukel, 1984c)). Some species show multicellular papillae on the sepals (e.g., *Epilobium* sp. (Saukel, 1983a, 1983b)), petals or tepals (e.g., *Convallaria* sp.). Of main importance are the pollen grains. The size and structure of pollen grains are highly conserved and, therefore, they often
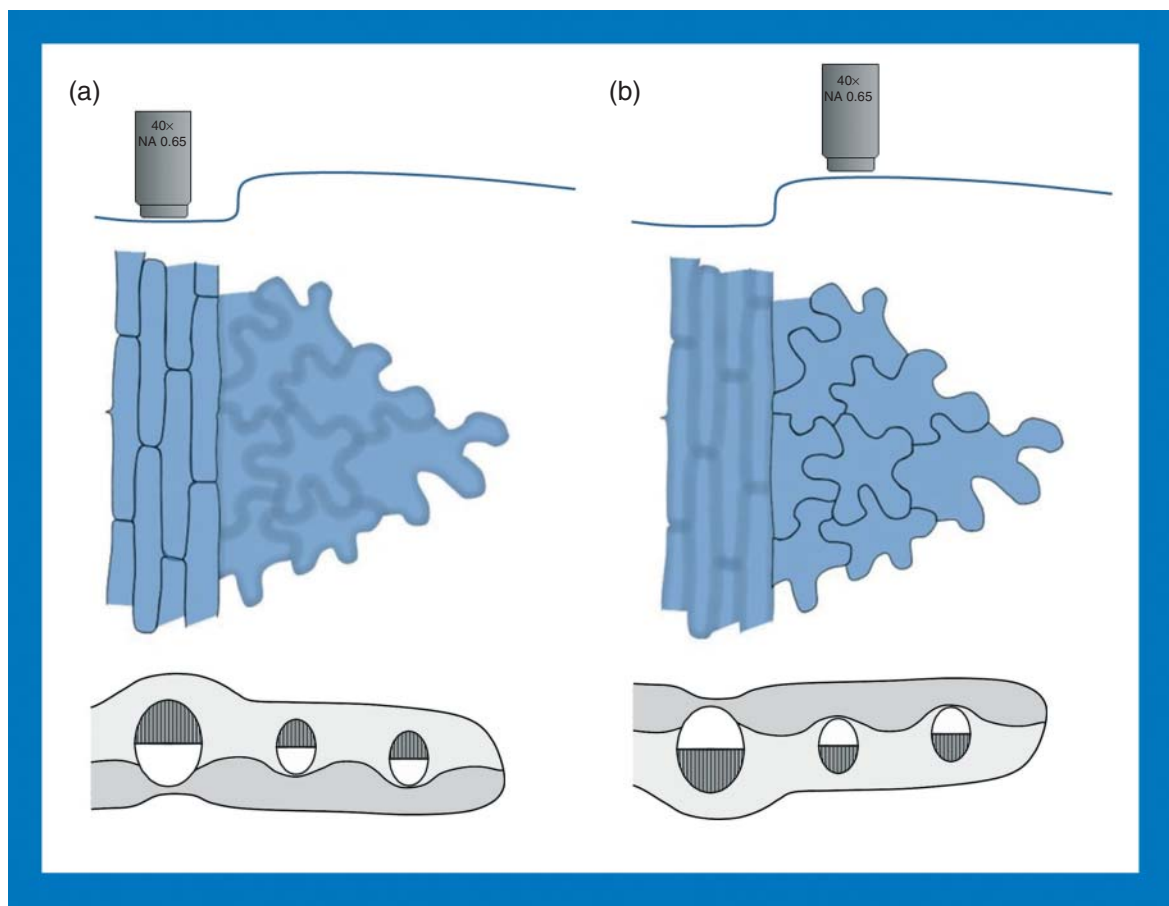
Figure 38 Shows a simple test for the decission which side of a leaf is visible in the specimen (a) Epidermis above the major vein is in focus – abaxial side and (b) epidermis of the intercostal region is in focus – adaxial side.

provide a significant feature for the identification of the species (Figure 40, (Saukel and Laenger, 1992)). The presence of secretory ducts may also sometimes be useful for determining adulterations (e.g., *Crocus sativus* L. without, *Carthamus tinctorius* L. with secretory ducts).

4.3 Seeds and Fruits

Both plant structures are the most ambiguous plant products concerning the identification in botanicals, which is evident in many cases of indehiscent fruits. The *seed coat* (= *testa*) is composed of several layers of cells, often exhibiting characteristic structures (not only for one species but also for one genus or family, e.g., willow herb and mustard (Bobrov *et al.*, 2004; Morozowska *et al.*, 2011; Netolitzky, 1926)). The seed coat shows commonly no stomata and chloroplasts but sometimes trichomes (cotton, willow herb, and willow). Parts of the testa are often modified to mucilage cell layers, or developed as sclerenchyma (e.g., Solanaceae and Brassicaceae), or as pigment layer. The seeds of many members of Solanaceae possess considerably thickened side and inner walls of the testa. Many seeds have a well-developed endosperm or/and perisperm filled with starch, oil, protein granules, or protein crystals, often concentrated in an aleuron layer (e.g., *Areca* sp., *Daturea* sp., and all caryopsis of Poaceae and Cyperaceae). The cells vary in thickness of the walls, also in porosity and the character of localized thickenings. Endosperm cells may show calcium

Figure 39 The inside view of the capitula of *Cosmea* sp. (Asteraceae).
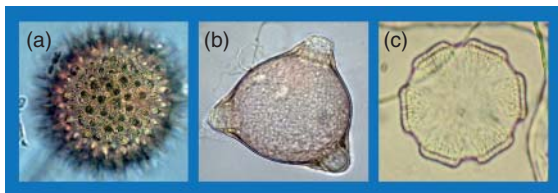
Figure 40 Pollen grains. (a) Malvaceae; (b) *Epilobium* sp.; (c) *Rosmarinus* sp.

oxalate crystals, in the case of cremocarps (Apiaceae) with very small crystals between drops of fixed oil. Some seeds have a conspicuous appendage referred to as *elaiosome*. Elaiosomes are important for the dispersal of seeds by ants and other insects.

The wall of *fruits* is parted in *exocarp*, *mesocarp*, and *endocarp*. The exocarp often anatomically resembles the structures of leaves or stems, which should be kept in mind; sometimes with characteristic epidermal cells (Figure 41, *Vaccinium myrtillus* L.), compare also (Base, 1897; Vogl, 1872, 1899; Winton and Moeller, 1906); and often with stomata and trichomes; if abundant, usually nonglandular and diagnostic (e.g., *Prunus* sp.). In the case of fleshy or dry and indehiscent fruits, the diaspore may include a combination of fruit, seed, and other plant parts. Figure 42 shows the powder fragments of *Rosae pseudofructus*, with unicellular hairs from the inside of the fleshy hypanthium. Sometimes, for example, in caryopsis, achene, and samara, a differentiation between seed and fruit is very difficult – especially

in the case of unknown objects. Compare also http://waynesword.palomar.edu/termfr1.htm. Cave: clear features to discriminate against seeds and of course vice versa are missing.

4.4 Stems

4.4.1 Aerial Stems

Stems at *primary state*: The epidermis develops normally stomata and often trichomes. The cortical parenchyma contains mostly chloroplasts. Vascular bundles are often enclosed in the endodermis, open and in most cases ordered in a circle (gymnosperms and dicots (Figure 43)), or closed collateral and scattered over the stem cross section (in monocots). Sometimes the endodermis remains as a bundle sheet containing starch. The subepidermal cells of the cortex often develop a collenchyma, sometimes located at edges or ridges (e.g., *Urtica* sp., *Hypericum* sp., and *Epilobium* sp. (Saukel, 1982a, 1982b). A pith is nearly always present, but very distinct only in gymnosperms and dicots with often thick-walled and pitted cells. In monographs of herbaceous perennial drugs, a clear description of respective features is often missing even though stems are quantitatively important parts of some drugs (e.g., *Achillea* sp., *Artemisia* sp., and *Agrimonia* sp.).

Stems after *secondary growth*: Only gymnosperms and dicots are qualified to develop a secondary growth (Figures 34 and 46) showing secondary xylem and phloem (secondary cortex) with medullary rays and cork with fiber bundles (Figure 34a and b) in transverse and radial sections. Figure 34c shows a piece of pinewood. Note that each cutting direction reveals a very different view of the wood. Real secondary wood is only rarely used as drug; however, in an intermediate form, we can see such structures in many herbs.

4.4.2 Rhizomes

Mostly we can see an obvious difference to aerial stems as rhizomes possess a greater part of parenchyma cells (Figure 44) and therefore often a lot of starch is present. If one puts a drop of iodine solution on a piece of a rhizome, the presence of starch is
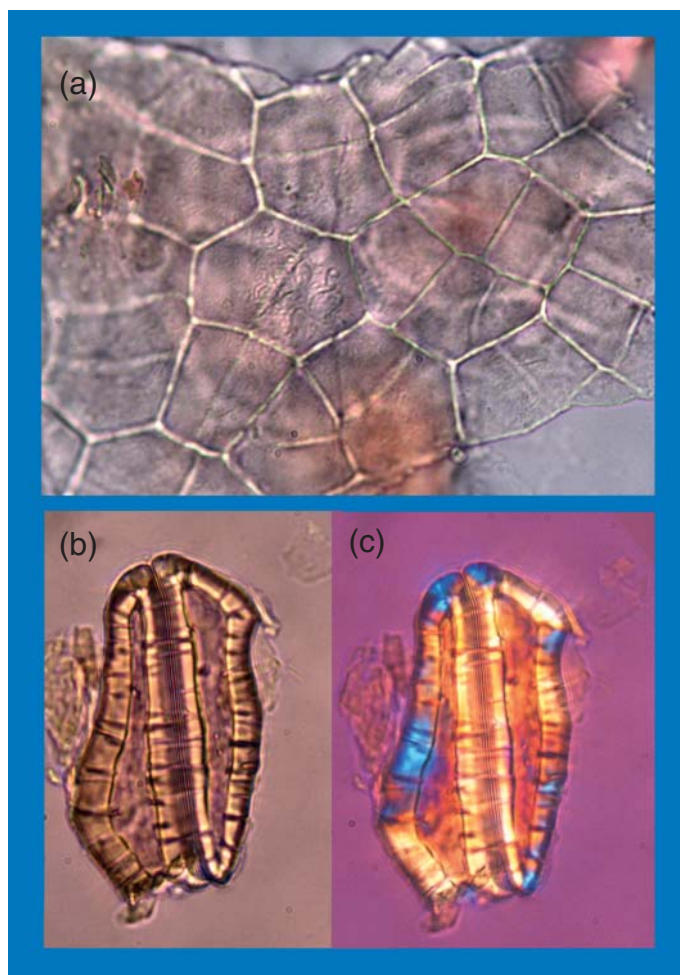
Figure 41 Powdered Myrtilli fructus. (a) Epidermal cells with a typical pattern; (b) sclereids of the endocarp in bright field; and (c) the same in polarizing light.

immediately visible. Sometimes the epidermis showing stomata but usually the epidermis is transformed to few external cork cell layers. In gymnosperms and dicots, the vascular bundles are enclosed within the endodermis, and developed as open conjoint bundles, in most cases arranged circularly (Figure 44b). The monocots develop leptocentric bundles in the center, and closed collateral bundles near the endodermis and within the cortex (Figure 44a), all being scattered over the stem cross section. The endodermis is often multiplied (e.g., Asteraceae) and in case of monocots forming a tertiary endodermis with thickened cell walls (Figure 45). Starch is abundant in the primary and secondary cortices, except for species unable to produce this polysaccharide (e.g.,

Gentianaceae and Asteraceae). Chloroplasts are rare in rhizomes. In numerous taxa, we find a combination of a short orthotropic rhizome combined with a taproot referred to as *beet* (e.g., *Gentiana* sp., *Taraxacum* sp., and many members of Apiaceae). In this case, the microscopic examination of cross sections should be conducted very carefully as the anatomical structures of the involved parts (root and stem) are quite different.

4.5 Barks

Barks are protecting tissues developed by two different meristems, the cambium and the phellogen
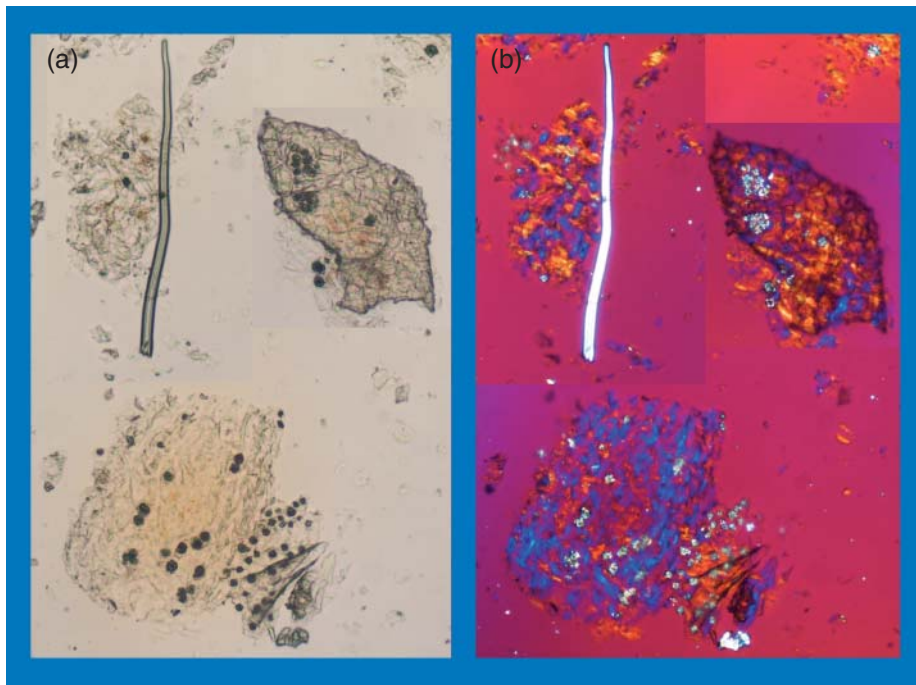
Figure 42 Powdered Rosae pseudofructus: (a) bright field and (b) polarizing light.

(Figure 46). The cells and tissues contained in barks are more or less uniform in many different drugs. Therefore, it is not so easy to distinguish between different barks. Often it is necessary to prepare cross sections for the detection of the presence of a primary cortex as in *Quercus* cortex. This detection is necessary to assure that the bark is harvested from younger trees. Important characters are the abundance of starch granules, the presence of stone cells, and/or fibers, often associated with oxalate crystals, the occurrence of oil cells, or secretory ducts. The structure of fibers is also very diagnostic if one looks at the cross section. Important findings are the shape (round and polygonal), evenly or unevenly thick walled, laminated or not, the visibility, and size of pit cavities. As an example may serve the discrimination between Frangulae cortex (*Frangula alnus* Mill.) with sclereids absent and Rhamni purshianae cortex (*Rhamnus purshiana* DC.) with sclereids present.

4.6 Roots

Roots at *primary state* – Primary roots itself are not used as drugs but may be present in some cases together with rhizomes (e.g., *Valeriana* sp.). The primary epidermis is labeled as rhizodermis (Figure 43) and always without stomata and cuticle but with root hairs, and after a short time mostly replaced by the hypodermis labeled as exodermis. One radial vascular bundle enclosed within the endodermis, with more numerous phloem/xylem parts in monocots (commonly >6, often numerous) than in dicots (commonly <6). Starch is abundant (primary and secondary cortices), except for the species that are unable to produce this polysaccharide (e.g., *Gentiana* sp. and many members of Asteraceae). Chloroplasts and pith wanting.

Roots after *secondary growth* – Similar to the stem of the same stage. The main difference is a small group of protoxylem cells at the beginning of the medullary rays (Figure 43).

Important are the verification of the starch granules (iodine staining directly on parts of the drug), the presence of fibers, the occurrence of tracheids and vessels, the type of vessels, and the amount of secondary xylem in relation to parenchyma.

For instance, the lack of fibers distinguishes Taraxaci radix (*Taraxacum* sp.) from many other roots of near relatives within Asteraceae (Fritz and
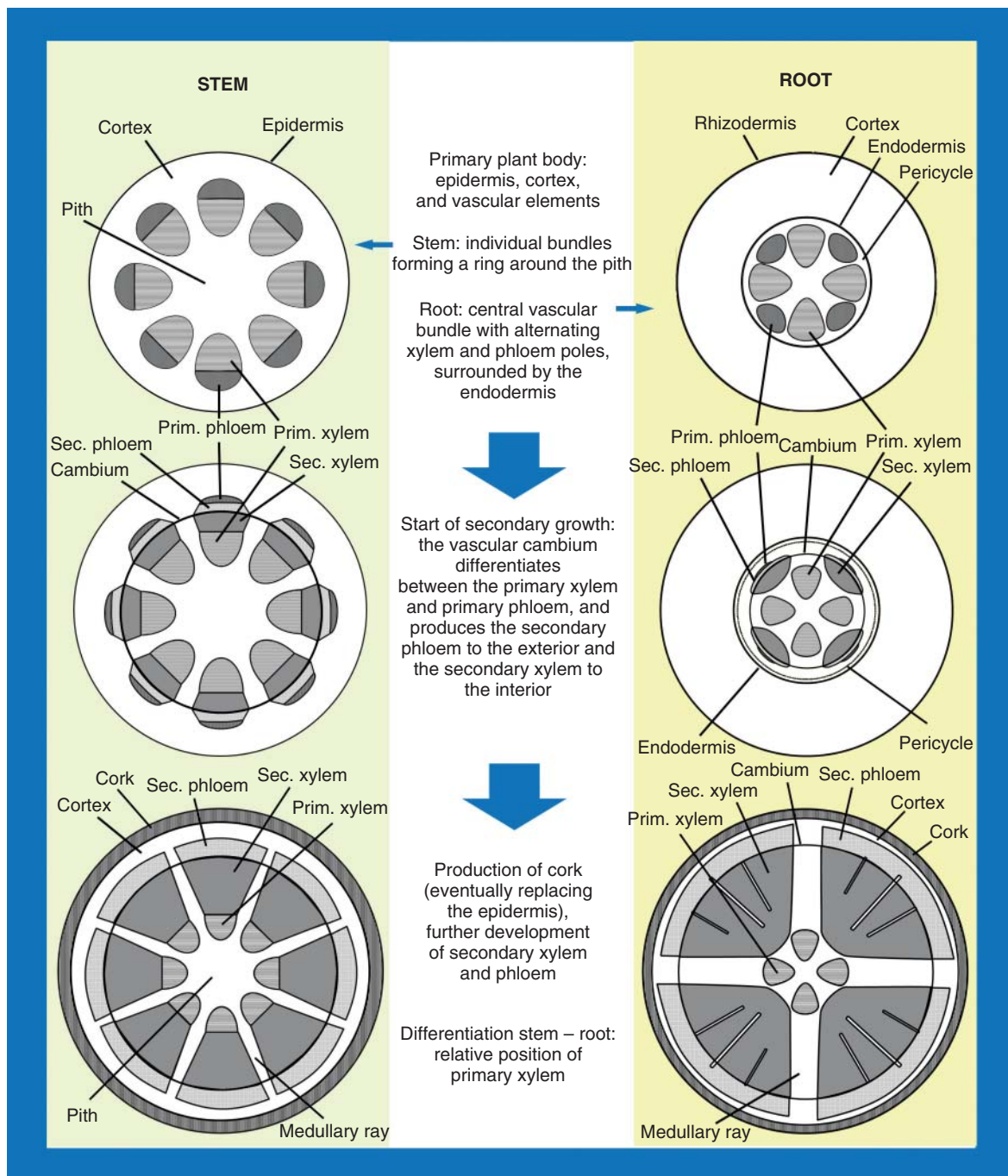
Figure 43   Development of secondary structures in stem and root.

Figure 44 (a) Monocotyledonous rhizome (*Convallaria majalis* L.), with leptocentric bundles in the center and collateral closed bundles in the region of the endodermis and (b) dicotyledonous rhizome (*Tussilago farfara* L.), vascular bundles arranged in a circle.

Figure 45 Cross section of a monocotyledonous rhizome showing the radial vascular bundle (*Allium* sp.): (a) bright field and (b) under polarizing light: red arrow – tertiary endodermis; green arrow – phloem; and blue arrow – protoxylem.

Saukel, 2011c). Again, Taraxaci radix can serve as example for a root lacking bordered pits, containing only vessels with reticulate thickening.

5 PREPARATION OF CRUDE DRUGS FOR MICROSCOPIC ANALYSES AND GENERAL USE OF REAGENTS

5.1 Microscopic Equipment

For microscopic analyses, some equipment besides the technical devices already explained earlier (see Section 2) is necessary for the preparation of adequate samples (for details, see Figure 47): microscope slides (typically $75\,mm \times 25\,mm \times 1\,mm$, on which the object is mounted on for the examination under the microscope); cover slips (= cover glasses; to be placed over the object on the slide); two types of razor blades, either the normal type for obtaining thin slices from softened material or single-edge razor blades for slivering hard and mostly unprepared plant material; a pointed forceps; a dissecting needle; a paintbrush (size 1 or 2); and absorbing paper/filter paper.

Furthermore, a great number of chemicals are employed in microscopic work but just a small number is actually required for the drug identification (Table 3). In addition, compare old and new literature (Schacht, 1862; Upton *et al.*, 2011).

Most important is the so-called *mounting medium*, in which the object being examined has to be embedded before examination. This mounting medium may either serve for temporary purposes only or can provide permanent slides for later evidence and archiving. Important for the quality of a mounting material is the chemical nonreactivity with the material being examined and its refractive index that ought to be similar to glass (n = 1.523 [cover slips]).

Prominent temporary mounting materials in pharmacy include water (1.33), glycerin (1.47), and an aqueous solution of chloral hydrate (ca. 1.5). Except for the examination of certain cell compounds such as starch, chloral hydrate is particularly useful, as the reagent causes to vanish not only starch and proteins but also chlorophyll and therefore serves as a useful and very fast bleaching/clearing agent. A proper way to get nice specimens of fresh material (especially small pieces of leaves) is to put small pieces of the material in the tube of a syringe together with water, then hold the syringe upright and push the plunger until all air is gone. After this, put one finger on the tip of the tube and drag the plunger back. Immediately one will see air bubbles leaving the plant material.
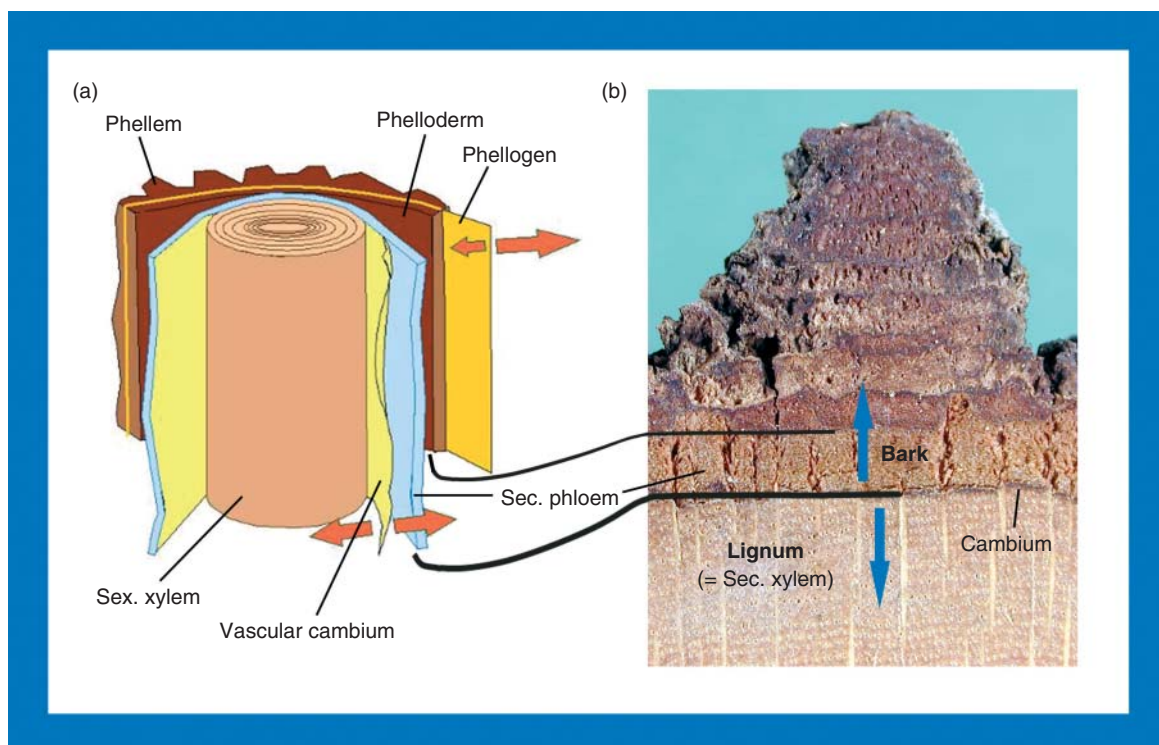
Figure 46 (a) Scheme of secondary growth and (b) photography of a woody stem of *Quercus* sp.

In this way, one can get more or less air-bubble-free specimens.

For permanent mounting, the used media ought to have stability over a long time without reacting with the embedded objects, without crystallizing or darkening. Canada Balsam and glycerin gelatin are well-known mounting materials. Canada Balsam is widely used in biology because of its high optical quality, its chemical inertness, and long preservability; however, the sections have to be dehydrated in a time-consuming procedure. Glycerin gelatin, on the contrary, is hydrophilic and therefore much simpler in use than the former mentioned and, thus, can be recommended for quality control in pharmacognosy.

Staining techniques can be employed to allow a better visualization of cell structures and can provide evidence for the chemical composition of certain cell components (e.g., starch, cutin, and lignin). Several types of staining media exist, each of them useful for a different purpose depending on the type of specimen and the cell structure being examined. However, for routine quality control, a few reagents

are well established and frequently used, which are listed in Table 3. For further information, one can find specialized old literature (Molisch, 1913, 1921; Poulsen and Trelease, 1886) and also new books in this field (Cutler *et al.*, 2008; Kremer, 2002).

5.2 Scanning Directions

For quality control of pharmaceutically used botanical raw materials, most often, dried and cut, or even powdered samples have to be analyzed. Thus, in order to determine the size, shape, and relative positions of the diverse cells and tissues, owing to the great diversity of the samples (e.g., roots, leaves, and fruits), one has to adapt the sampling techniques according to the present drug material and has to possess a good understanding of plant anatomy.

Depending on the species, the plant part and the anatomical structures that are supposed to be examined, different views of the object are requested in order to achieve a thorough picture. The object has

Table 3 Important chemicals for the use in microscopic analyses.

Distilled water	For temporary mounting – especially for the examination of starch and proteins

Solution of chloral hydrate
  60% chloral hydrate in water
	For temporary mounting;
	  Swells and finally dissolves starch, proteins, resins, volatile oils, and chlorophyll (bleaching agent)

Ether–ethanol
  Ether 1
  Ethanol (96%) 1
	For removal of fixed oils, fats, resins, volatile oils, tannins, and chlorophyll

Glycerin (pure)	Prevents dissolving of sugars; for temporary mounting
Glycerin gelatin 30 g gelatina alba solved in 120 mL water, 150 g 85% glycerol	For permanent mounting

Iodine
  Potassium iodide 2
  Iodine 1
  Water 100
	Indicator for starch: amylose turns blue and amylopectin stains yellowish brown

KOH 5% Potassium hydroxide in water	1,8-Dihydroxyanthrachinones stain red (Bornträgerreaktion)
$FeCl_3$ Stock solution: 10% $FeCl_3$ in water before use dilute with water 1 : 10	Stains tannins blackish-green (condensed tannins) and blackish-blue (hydrolyzable tannins)
Phloroglucinol–HCl 10% Phloroglucinol in ethanol 96% – before examining put HCl conc. directly on the object slide	Stains lignified cell walls red
Astra blue and Safranin 0.5% Astra blue in 0.5% acetic acid; 0.5% safranin in water; ratio 5–50 : 1	Stains nonlignified cell walls blue and lignified ones red
Methylene blue	Stains mucilage blue; staining of living cells
Tusche (Chinese ink) Diluted with water as needed	Detection of mucilage – mucilage stays clear in contrast to the dark preparation

CP-reagents
  60% Chloral hydrate in water 2
  85% Phosphoric acid 1
	Detection of proazulene

Sulfuric acid (H_2SO_4) conc.	Detection of calcium oxalate – calcium oxalate dissolves with formation of needles of gypsum

to be prepared in such a way that the material being analyzed allows the light of the microscope to pass through sufficiently for an adequate view. At the most, slices of the sample have to be obtained as thin as possible, preferably about 20–30 μm, to get a clear picture of all anatomical details and to avoid disturbing interferences. Figure 48 shows an improper specimen; the slice is too thick (recognizable on the sight into the secretory ducts) and full of air bubbles.

- Surface view: In order to see anatomical structures such as stomata or glandular hairs on the surface of plant organs, a surface view is required. In the case of translucent material (e.g., thin leaves or flowers may be transparent enough after treating with chloral hydrate – see Section 5.3), sectioning is unnecessary, and the sample may be placed directly on the slide. For thicker objects, a *paradermal section* – sectioning parallel to the surface – has to be performed.
- View of inner tissues: Usually, three directions of cutting are performed to provide a thorough picture of the distribution and structure of the inner tissues of a plant organ (Figure 49):
  - *Transverse section* (*cross section*) → The plant organ is sectioned at right angles to its main axis.
  - *Longitudinal sections*
    - *Radial section* → The plant organ is sectioned parallel to its main axis and directly through its center.
    - *Tangential section* → The plant organ is sectioned parallel to its main axis and not through its center.
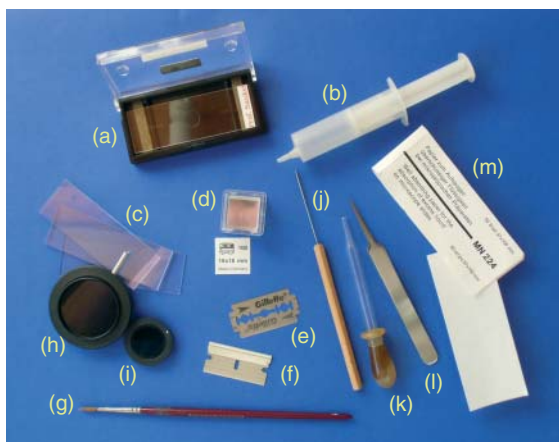
Figure 47 Necessary equipment for microscopic analyses: (a) objective (=stage) micrometer; (b) syringe; (c) microscope slides; (d) cover glasses (cover slips); (e) razor; (f) single-edge razor; (g) paint brush; (h) polarizer; (i) analyzer, this type is placed beneath the binocular tubus; (j) dissecting needle; (k) small pipette; (l) sharp pointed forceps; (m) absorbing paper.
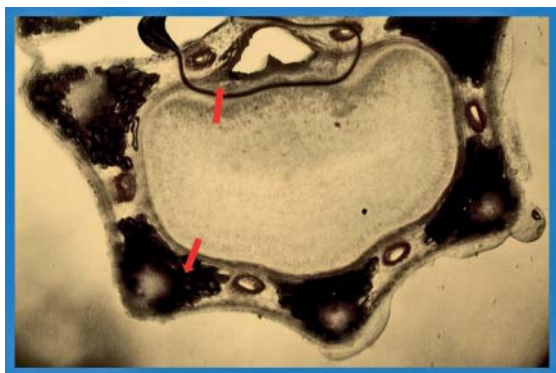
Figure 48 Improper cross section of an Apiaceae fruit showing a lot of air bubbles (red arrows).

The differentiation between the last two cutting directions is especially important for the analyses of the vascular cylinder of roots, rhizomes, and stems as they are giving two different views of the axial system, especially concerning the rays (Figure 43). With the radial section, the section is parallel to a radius – showing the rays lying across the axial system, whereas with the tangential section, the section is perpendicular to a radius – revealing the height and width of the rays (see Figure 34).

The correct cutting direction is most important to achieve a well-interpretable picture of the tissues.

A slightly thicker slice can be more adequate if the cells are properly cut at a right angle in contrast to a skewed one.

An interesting question resulting out may be raised in this context – is it easier to analyze powdered drugs or is the usage of slices the proper method? If we look in the early literature in the nineteenth century and later on, some authors preferred the inspection of powdered drugs only (Greenish and Collin, 1904; Jackson and Snowdon, 1968; Jelliffe and Rusby, 1895; Koch, 1906; Mitlacher, 1904; Schneider, 1921; Schürhoff, 1906; Vogl, 1872; Zörnig, 1912), whereas other authors used both preparations (Berg, 1865; Clayton and Hassall, 1909; Hassel, 1861; Melchior and Kastner, 1974; Moeller, 1886, 1889, 1892, 1901; Vogl, 1899; see also Table 2 for more interesting books in this field). Figure 50 shows the master pencil drawing from Chinese cinnamon preserved in the Library of Pharmacy at the University of Vienna. In the important European Pharmacopoeia, the pharmacognosist is instructed to powder the drug in question before the examination! In our opinion, both methods are useful depending on the examined drug. Note: for the newcomers, it is indispensable to prepare slices and preparations of the powder from the same material and make a thorough comparison. In some cases, for example, *Quercus* bark or *Frangula* bark, a complete cross section is unalterably necessary to decide if the drug complies with the monograph.

5.3 Preparing Sections

The majority of the raw materials used for analyses in quality control are received in dry condition and often require some preliminary treatment before sectioning to enable better results by softening the material. This may include soaking in water for some time, boiling for a few minutes, or just wetting the razor blade and/or the cut surface for a few seconds. Embedding the object in alcohol (60–90%) after softening in water may help to achieve good results with the following sectioning by hardening the material again.

If softening is necessary at all and in which way has to be decided according to the sample. Possible thermolability of cell contents (e.g., starch and mucilage) has to be kept in mind (no boiling!), and certain cell contents (e.g., mucilage) may demand dry sectioning.
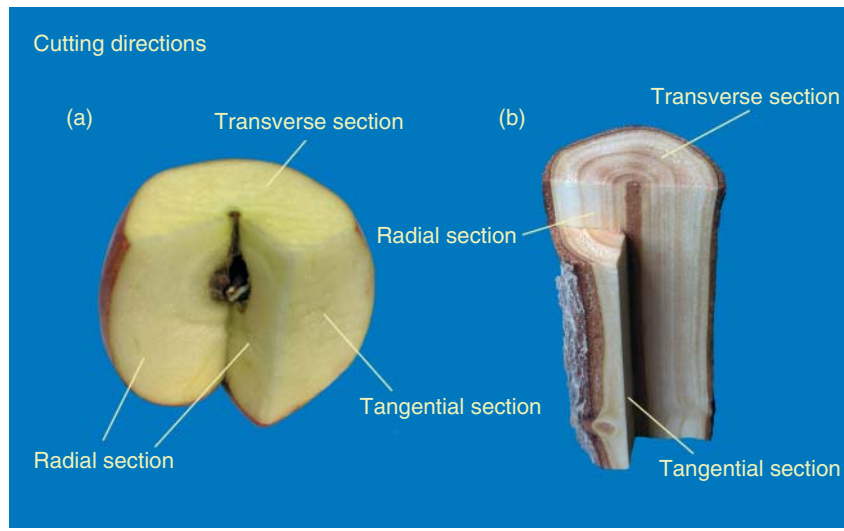
Figure 49 The three principle cutting directions shown on (a) an apple and (b) on a piece of a fir branch.

There are two different ways of sectioning: (1) free-hand sectioning and (2) sectioning using a micro-tome. For quality control, freehand sectioning is usually sufficient and needs less time. According to the requested kind of section, there are different ways to accomplish the cutting.

The easiest way of sectioning is to hold the sample material in one hand between thumb and forefinger while cutting with the razor blade holding in the other hand (Figure 51b). This, however, requires a certain firmness and size of the object (e.g., roots, stems, barks, seeds, and many fruits).

Thin and soft objects such as leaves, and material that is too small to hold in the hand, may need to be sectioned by embedding them in a block of polystyrene to provide stability (compare Figure 51). A block of polystyrene is cut into two halves with the object clamped in between (Figure 51c). Thereby, one has to pay attention that the object is correctly adjusted, so that the cutting direction is perpendicular to the edge of the polystyrene. After removing the supernatant material (Figure 51d), thin slices are obtained with a diagonal movement of a razor blade (Figure 51e). Very important is to employ the correct cutting angle (Figure 51f).

Alternatively, thin objects may be cut directly without embedding in polystyrene lying on an even polystyrene block, fixated with a finger or (in case of flat objects) a slide on top. One should use the fingernail and the edge of the fixating slide for guiding the razor blade.

For paradermal sections of leaves, the leaf is wrapped around the forefinger of one hand to give stability to the material while cutting with the blade in the other hand.

Preferably, all sections ought to be performed with the help of a stereomicroscope with enhanced magnification to achieve better results.

The resulting slice can be transferred onto a fresh slide, already prepared with a drop of clearing reagent (usually chloral hydrate solution or water) with the help of either a preparation needle or a paintbrush (Figure 51g–i). If thin and therefore weak cross sections are at hand, it is very helpful to place small splinters of a cover glass on the four edges of the planned position of the cover slip (Figure 51j). This procedure is important to avoid crushing the specimen. Afterward, a cover glass is placed on the slide on top of the slice (one should place the cover slip at a 45°, with one edge touching the mounting media, and slowly let go, Figure 51j). It is important to use enough reagent to cover the slice. Thereby, one has to avoid contamination of the top of the cover glass to protect the objective of the microscope from possibly aggressive substances and to guarantee a spotless view under the microscope later on. Apart from heat-sensitive objects, the slices are now heated carefully and as short as possible above a Bunsen burner to eliminate air and to bleach the object. The
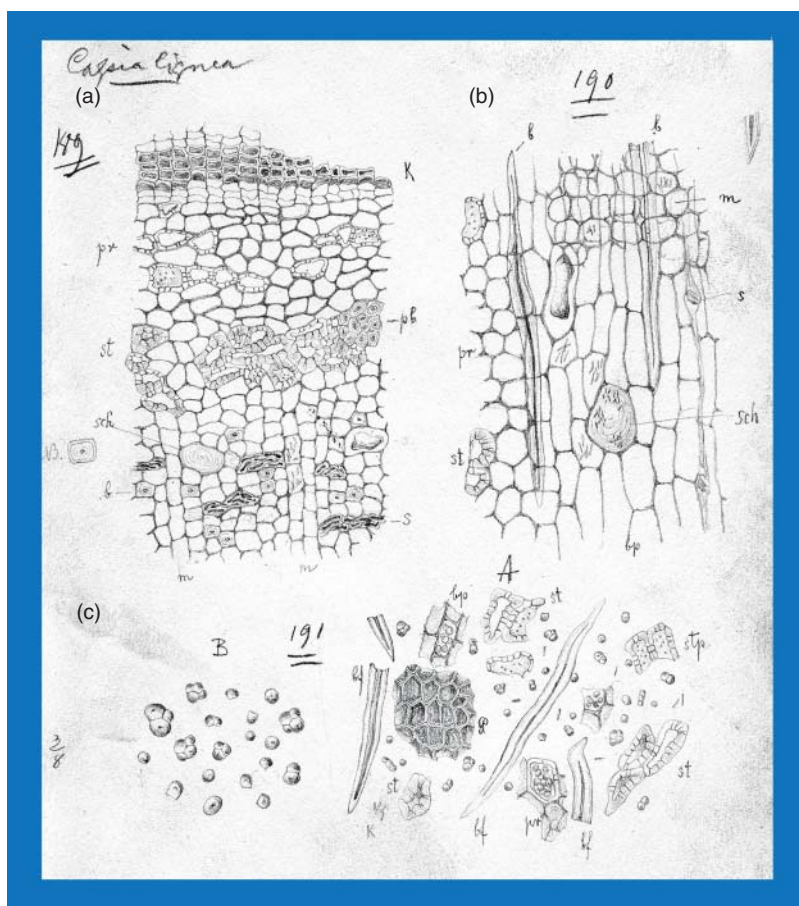
Figure 50 Chinese cinnamon preparations, the master pencil drawing from J. Moeller (the mid-nineteenth century) preserved in the Library of Pharmacy at the University of Vienna. (a) Cross section, (b) longitudinal section, and (c) the powdered drug.

best way is to position the slide above the flame, so that the flame does not reach the area of the cover glass (Figure 51l). Otherwise, boiling delay leads to little explosions caused by violent air bubble production. This may cause a dangerous situation for the investigator and has to be avoided. Vaporized solution always has to be added afresh in a way that the slices never desiccate. This process has to be repeated until no air bubbles disturb the examination of the sample and/or the material is sheer enough, so that all details can be clearly viewed under the microscope. One has to be aware that the heating process has an influence on the specimen and may cause a destruction of cell walls. Therefore, the quality of the slices ought to be checked regularly under the microscope all the while.

5.4 Permanent Slides

After preparing interesting objects in the way mentioned earlier, a small amount of glycerin gelatin, prepared on a separate slide, is melted carefully (without boiling to avoid bubbles!). Potential bubbles may be removed with a forceps before the sections are gently transferred from the first slide into the melted gelatin so that the plant material is covered completely by the gelatin. Thereby, attention has to be paid not to transfer a lot of the first mounting medium. A first transfer of the slices into pure glycerin before the final transfer into the gelatin may serve as intermediate step. Under the stereomicroscope, the correct position of the section can be controlled and, if necessary, corrected (e.g., a rolled up leaf may be stretched again).
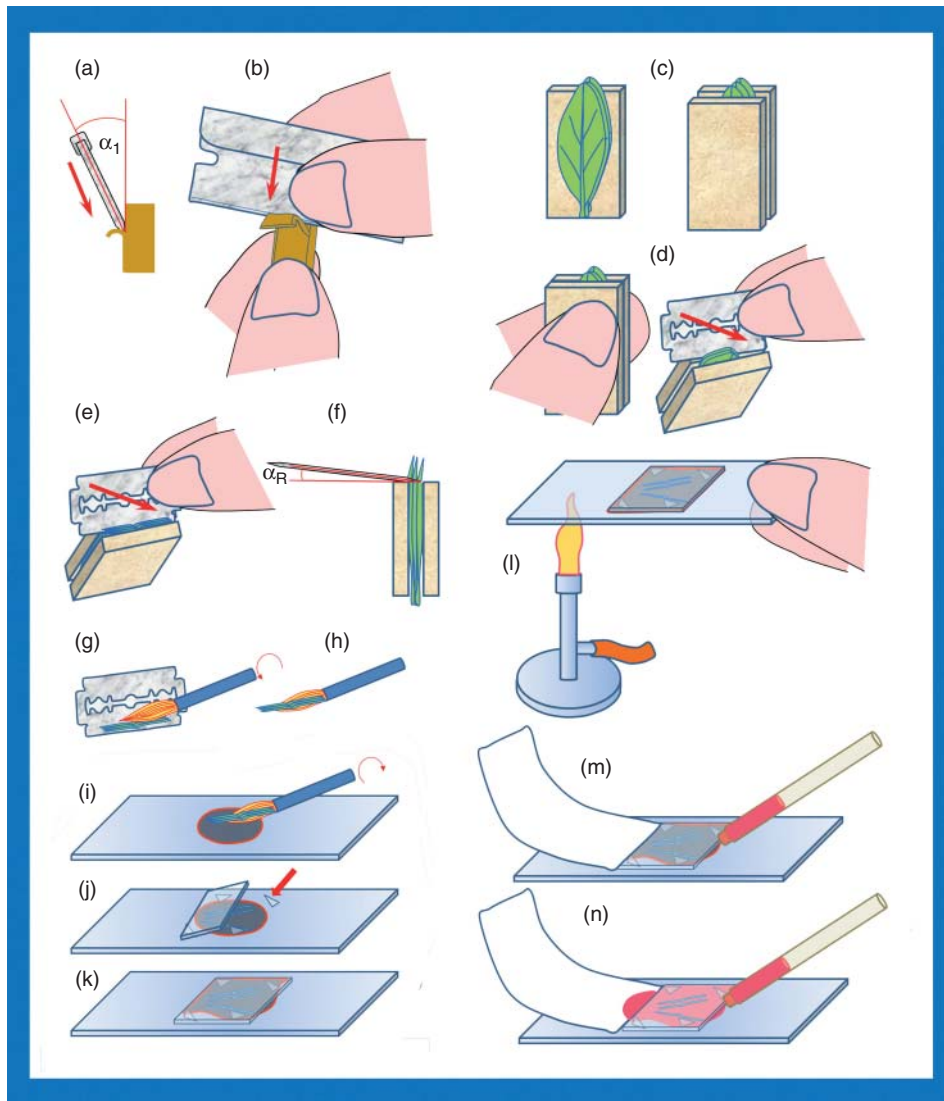
Figure 51 (a) Right cutting angle; (b) position for the work with a single-edge razor; (c) embedding of soft material in between two little blocks of polystyrene; (d) removing the supernatant material with the razor; (e) thin slices are obtained with a diagonal movement of the razor; (f) the correct cutting angle; (g) and (h) transferring the slice from the razor onto a fresh slide with a paintbrush, already prepared with a drop of clearing reagent (i); (j) placement of small splinters of a cover glass on the four edges of the planned position of the cover slip; (j) and (k) lowering of the cover slip onto the slide; (l) positioning the slide above the flame; and (m) and (n) staining of the specimen with a pipette on the one side and a strip of filter paper on the other side.

Now, a cover slip is placed onto the sample. The whole space under the cover glass should be filled with gelatin.

Especially beginners often have problems with the thickness of the produced slices, which causes the slices to fall over when the glycerin gelatin is hardening. Embedding small splinters of a cover

glass beside the object under the cover slip is an easy trick to prevent this happening (Figure 51j).

Permanent samples usually last for many years, though one has to be aware that delicate structures such as glandular hairs or other fine tissues may be destroyed. Thus, for leaves and flowers, the preparation of new slides is recommended.

5.5   Documentation

For documentary evidence, permanent slides and the preparation of both drawings and photographs are valuable methods. While photomicrography often is considered superior and more advanced than drawings, the latter also possesses some advantages and the two methods must complement each other.

Photography is not as time-consuming as drawings and the resulting pictures represent the microscopic objects exactly the way they appear: There are no subjective influences as may be possible in drawings, and the colors of the object can be shown as well. With drawings on the other hand, multiple optical layers can be united in one illustration (Figure 52), whereas in photography often not all structures can be brought into focus, important structures may be emphasized and disturbing contaminations excluded in the drawing.

It highly depends on the purpose of the pictures (analyzing, studying, presentation, etc.) and the examined objects themselves, if photography or drawing or both are useful.
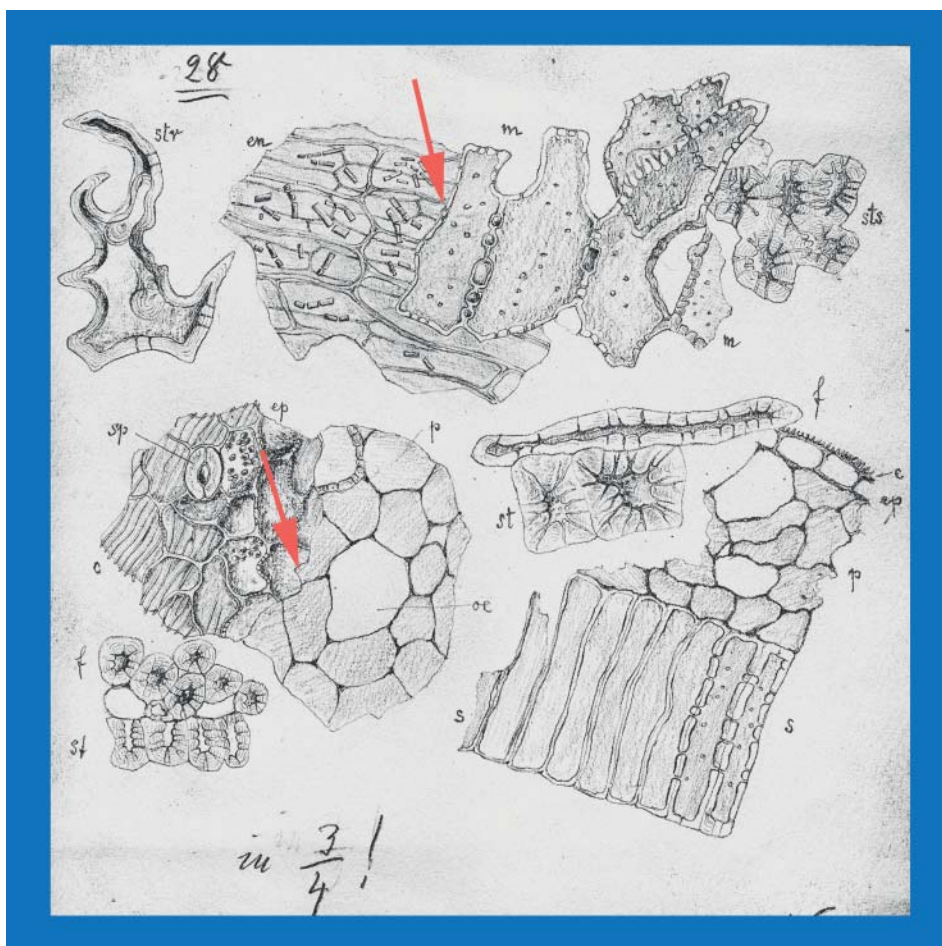
Figure 52   Tissue fragments from Anisi stellati fructus, the master pencil drawing from J. Moeller (the mid-nineteenth century) preserved in the Library of Pharmacy at the University of Vienna. (en) Testa; (m) sclerotic parenchyma; (sts) sclereids; (f) fibers cross section (left) longitudinal section (right); (s) macrosclereids; (ep) epidermis with stomata (sp); (p) parenchyma from the carpells with oil cells (oe); and (str) asterosclereid. The arrows show multilayer drawings!
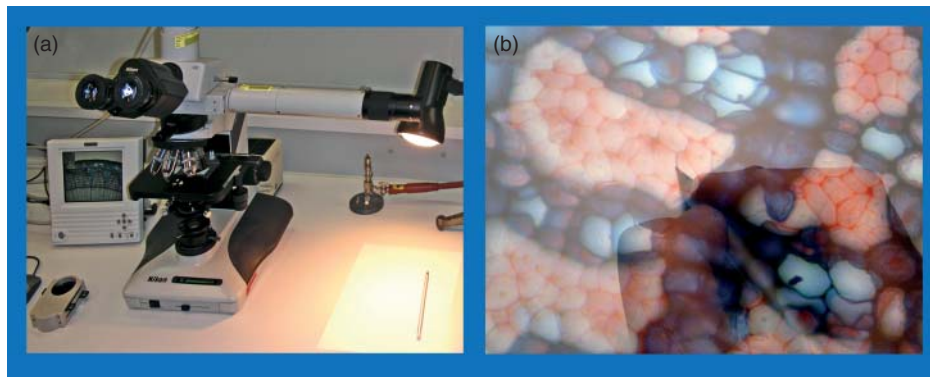
Figure 53 (a) Microscope with drawing tube attached and (b) digital image of a specimen and the pencil, captured through a trinocular tube with a digital camera.
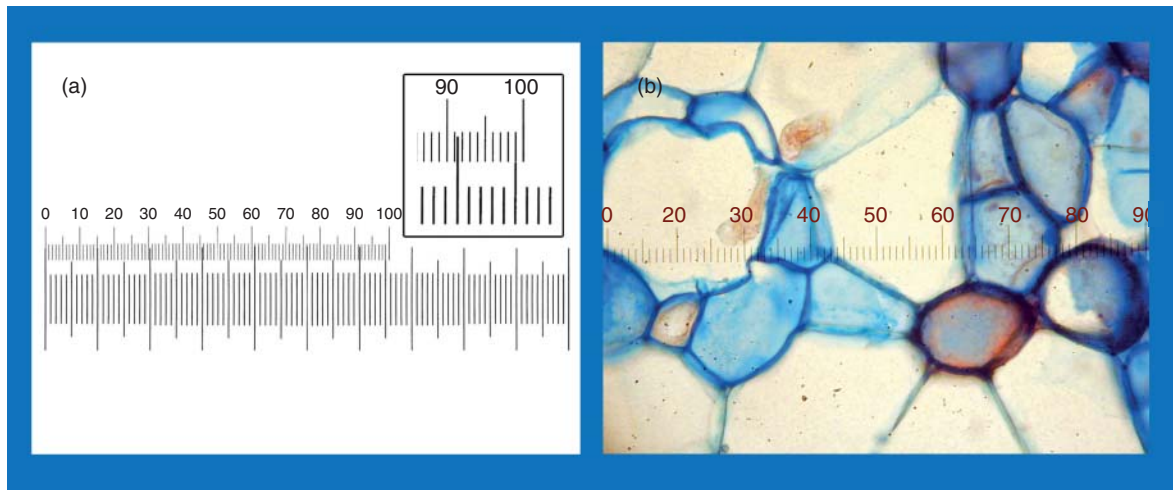
Figure 54 (a) Calibration of the ocular micrometer: two scales adjusted and (b) measurement of cells with the ocular micrometer.

5.5.1 Drawings

For scientific purposes, a drawing tube is attached to the microscope, extending off the right (or left) side (Figure 53a). Below the end of the drawing tube, the drawing paper is placed. Prisms inside the drawing tube allow seeing the paper, the pencil and the microscopic image at the same time while looking into the ocular of the microscope and thus enable accurate tracing of the images (Figure 53b). Important thereby is the correct adjustment of the light. One has to experiment with the lighting of the microscope. A low intensity of light in the microscope improves the image of the pencil and the drawing lines but the specimen might not be clearly seen anymore. A stronger intensity in contrast provides better view of the specimen but a worse of the drawing. The right balance between them has to be found. In addition, a light source next to the drawing paper may be helpful to get a stronger image of the drawing and the pencil.

5.5.2 Photomicrography

Two principal techniques for photomicrography are the usage of analog film material and the digital photography. The work with the old-fashioned analog cameras is not easy as one cannot see the result immediately and the great time span between taking

the picture and getting the finished photograph is often a problem. Furthermore, it becomes more and more difficult to find labs for the development of pictures or slide films. Nowadays, the development of digital photography is remarkable in resolution and charge/price of the cameras. The great advantages are the immediately visible images and the possibility of subsequent correction on a computer. The most important drawback is the smaller dynamic range that is the difference between the darkest and the brightest areas of an image, of the used photo sensors. Most digital cameras have a dynamic range of $1:65$, whereas the value for a slide film can be in the range of $1:100$ until $1:300$. This means that the choice of the exposure time is crucial. The necessity of avoiding faulty exposures in digital photography leads especially in photomicrography often to a grayish colored background instead of a white one.

It is possible to take digital photos directly through the ocular with an ordinary compact camera. Naturally, to receive pictures of better quality for scientific working and documentation, special microscope cameras usually are connected with the microscope in a steady way with the help of an adaptor. One low price option may be an adaptor fitting into the eye tube of the microscope by replacing an eyepiece. Modern research microscopes most often possess a trinocular tubus that enables a stable hookup for the camera over the base. The camera is typically attached to a computer, which provides the control functions for the camera and allows further image processing too. Compare also the important book in this field (Wu, Merchant, and Castleman, 2008).

5.6   Microscopical Measurements

Regarding the identification and purity control of drugs, the size of anatomical structures can be of important value. Microscopic measurements can be taken using a special ocular including a scale in relative units (*ocular micrometer/eyepiece micrometer*). For accurate measurement, any ocular micrometer has to be calibrated with the help of a *stage micrometer* – a microscope slide with a finely divided, precise scale engraved on the surface.

For calibration, with the ocular micrometer in use, the stage micrometer is placed directly on the stage and focused in the ordinary way. The two scales

should now be adjusted in a way that they appear parallel to each other and the first lines of the two scales are aligned (Figure 54a). Now the most distant unit where the two scales coincide again has to be determined. In the example (Figure 54a), 65 stage micrometer divisions cover the same distance as 99 ocular divisions. Thus, if one division of the stage micrometer is equivalent to $1\,\mu m$, one unit of the ocular micrometer (the smallest increment on the scale) equals $99/65 = 1.5\,\mu m$. This calibration has to be performed for every optical combination.

With the now known value of one eyepiece division, an accurate size of an anatomical structure can be calculated by multiplication (Figure 54b).

REFERENCES

Andrzejewskagolec, E. (1994) *Acta Societatis Botanicorum Poloniae*, **63**(2), 199–204.

Bagchi, G. D. (2000) *J. Med. Arom. Plant Sci.*, **22**(1B), 605–615.

Base, D. (1897) *Elements of Vegetable Microscopy*, Baltimore MD.

Beilstein, M. A., Al-Shehbaz, I. A., and Kellogg, E. A. (2006) *Am. J. Bot.*, **93**(4), 607–619.

Berg, O. C. (1865) *Anatomischer Atlas zur pharmazeutischen Warenkunde*, Verlag von Rudolph Gaertner, Berlin, p. 103.

Bliss, M. C. (1921) *Bot. Gaz.*, **71**(4), 13.

Bobrov, A., Melikian, A. P., Romanov, M. S., *et al.* (2004) *Bot. J. Linnean Soc.*, **145**(4), 437–443.

Bruni, A., and Modenesi, P. (1983) *Nordic J. Bot.*, **3**(2), 245–251.

Cantino, P. D. (1990) *J. Arnold Arbor.*, **71**(3), 323–370.

Carlquist, S. (2010) *Bot. J. Linnean Soc.*, **164**(4), 342–393.

Carpenter, K. J. (2005) *Am. J. Bot.*, **92**(10), 1595–1615.

Clayton, E. G. and Hassall, A. H. (1909) *A Compendium of Food-Microscopy with Sections on Drugs, Water, and Tobacco*, Baillière, Tindall and Cox, London.

Corsi, G. (1992) *Phyton-Annales Rei Botanicae*, **32**(2), 247.

Croxdale, J. L. (2000) *Am. J. Bot.*, **87**(8), 1069–1080.

Curtis, J. D., and Lersten, N. R. (1990) *New Phytologist*, **114**(4), 571–580.

Cutler, D. F., Botha, C. E. C. E. J., and Stevenson, D. W. (2008) *Plant Anatomy: An Applied Approach*, Wiley.

Downing, T. L., Ladiges, P. Y., and Duretto, M. F. (2008) *Plant Systemat. Evolut.*, **271**(3–4), 199–221.

Eckhart, W. (1929) *Osterreich Bot Zeitschr*, **78**(2), 129–156.

Evert, R. F. (2006) *Esau's Plant Anatomy, Meristems, Cells, and Tissues of the Plant Body: their Structure, Function, and Development*, John Wiley & Sons, Inc., New Jersey, p. 601.

Fahn, A. (1988) *New Phytologist*, **108**(3), 229–257.

Flück, H., Schlumpf, E., and Siegried, K. (1935) *Pharmakognostischer Atlas zur Pharmacopoea Helvetica*, Kommissionsverlag Wepf & Co., Basel.

Flückiger, F. A. (1867) *Lehrbuch der Pharmakognosie des Pflanzenreiches*, Rudolph Gaertner, Berlin.

Fornasiero, R. B., Bianchi, A., and Pinetti, A. (1998) *J. Herbs, Spices Med. Plants*, **5**(4), 12.

Foster, A. S. (1956) *Protoplasma*, **46**, 10.

Foster, A. S. (1942) *Practical Plant Anatomy*, D. Van Nostrand Company, Inc., New York.

Franceschi, V. R., and Horner, H. T. (1980) *Bot. Rev.*, **46**(4), 361–427.

Franceschi, V. R., and Nakata, P. A. (2005) *Annu. Rev. Plant Biol.*, **56**, 41–71.

Freire, S. E., Urtubey, E., and Giuliano, D. A. (2007) *Caldasia*, **29**(1), 23–38.

Fritz, E., and Saukel, J. (2011a) *Sci. Pharm.*, **79**(1), 157–74.

Fritz, E., and Saukel, J. (2011b) *Acta Biol. Cracov. Bot.*, **53**(1), 63–73.

Fritz, E., and Saukel, J. (2011c) *Pharm. Biol.*, **49**(8), 789–795.

Fryns-Claessens, E., and van Cotthem, W. (1973) *Bot. Rev.*, **39**(1), 67.

Gabel, C. E. (1912) *Microscopy and the Microscopical Examination of Drugs*, The Canyon Co., De Moines.

Glover, B. J. (2000) *J. Exp. Bot.*, **51**(344), 497–505.

GOVERNMENT_OF_INDIA. unknown. The Ayurvedic Pharmacopoeia of India, Part I. GOVERNMENT OF INDIA.

Greenish, H. G., and Collin, E. (1904) *An Anatomical Atlas of Vegetable Powders*, Churchill, London.

Haberlandt, G. (1914) *Plant Physiology*, MacMillan, London.

Hagel, J. M., Yeung, E. C., and Facchini, P. J. (2008) *Trends Plant Sci.*, **13**(12), 631–639.

Haron, N. W., and Moore, D. M. (1996) *Bot. J. Linnean Soc.*, **120**(3), 265–277.

Hassel, A. H. (1861) *Adulterations Detected; or, Plain Instructions for the Discovery of Frauds in Food in Medicine*, 2nd edn, Longman, Green, Longman, and Roberts.

Horner, H. T., Franceschi, V. R., and Hill, E. L. (1978) *J. Cell Biol.*, **79**(2), A38–A38.

Hu, G.-X., Balangcod, T. D., and Xiang, C.-L. (2012) *Biologia*, **67**(5), 867–874.

Jackson, B. P., and Snowdon, D. W. (1968) *Powdered Vegetables Drugs*, J. & A. Churchill Ltd, London, p. 203.

Jansen, S., Baas, P., Gasson, P., *et al*. (2004) *Proc. Natl. Acad. Sci. U. S. A.*, **101**(23), 8833–8837.

Jelliffe, S. E., and Rusby, H. H. (1895) *Essentials of Vegetable Pharmacognosy – A Treatise on Structural Botany – Designed Especially for Pharmaceutical and Medical Students, Pharmacists and Physicians*, D.O. Haynes & Co, New York.

Kaussmann, B., and Schiewer, U. (1989) *Funktionelle Morphologie und Anatomie der Pflanzen*, Gustav Fischer, Jena, p. 465.

Koch, L. (1906) *Einführung in die mikroskopische Analyse der Drogenpulver*, Gebrüder Bornträger, Berlin.

Köhler, A. (1893) *Zeitschrift für wissenschaftliche Mikroskopie und für mikroskopische Techni*, **10**(4), 8.

Kraemer, H. (1910) *A Text-Book of Botany and Pharmacognosy, Intended for the Use of Students of Pharmacy, as a Reference Book for Pharmacists, and as a Handbook for Food and Drug Analysts*, J.B. Lippincott, Philadelphia, London.

Kremer, B. P. (2002) *Das grosse Kosmos-Buch der Mikroskopie*, Kosmos.

Krings, M., Taylor, T. N., and Kellogg, D. W. (2002) *Evolut. Ecol. Res.*, **4**(5), 779–786.

Leelavathi, P., Ramayya, N., and Prabhakar, M. (1980) *Phytomorphology*, **30**(2–3), 195–204.

Lersten, N. R., Czlapinski, A. R., Curtis, J. D., *et al*. (2006) *Am. J. Bot.*, **93**(12), 1731–1739.

Melchior, H., and Kastner, H. (1974) *Gewürze botanische und chemische Untersuchung*, Paul Parey, Berlin und Hamburg, p. 290.

Metcalfe, C. R., and Chalk, L. (1957) *Anatomy of the Dicotyledons*, At The Clarendon Press, Oxford.

Mitlacher, W. (1904) *Toxikologisch oder forensisch wichtige Pflanzen und vegetabilische Drogen mit besonderer Berücksichtigung ihrer mikroskopischen Verhältnisse*, Urban & Schwarzenberg, Berlin Wien, p. 200.

Moeller, J. (1886) *Mikroskopie der Nahrungs- und Genussmittel aus dem Pflanzenreiche*, J. Springer, Berlin.

Moeller, J. (1889) *Lehrbuch der Pharmakognosie*, A. Hölder, Wien, p. 450.

Moeller, J. (1892) *Pharmakognostischer Atlas; mikroskopische Darstellung und Beschreibung der in Pulverform gebräulichen Drogen*, Springer, Wien, p. 443.

Moeller, J. (1901) *Leitfaden zu mikroskopisch-pharmakognostischen Übungen für Studierende und zum Selbstunterricht*, Alfred Hölder, Wien, p. 336.

Mohl, H. (1846) *Mikrographie, oder Anleitung zur Kenntnis und zum Gebrauche des Mikroskops*, L.F. Fues, Tübingen.

Molisch, H. (1913) *Mikrochemie der Pflanze*, G. Fischer, Jena.

Molisch, H. (1921) *Mikrochemie der Pflanze*, G. Fischer, Jena.

Moll, J. W., and Janssonius, H. H. (1923) *Botanical Pen-Portraits*, Martinus Nijhoff, Hague.

Morozowska, M., Czarna, A., Kujawa, M., *et al*. (2011) *Plant Systemat. Evolut.*, **291**(3–4), 159–172.

Muravnik, L. E., and Shavarda, A. L. (2012) *Nordic J. Bot.*, **30**(4), 470–481.

Netolitzky, F., Tischler, G., and Linsbauer, K. 1926. Anatomie der Angiospermen-Samen, Borntraeger, Berlin.

Netolitzky, F. (1908) *Bestimmungschlüssel und Anatomie der einheimischen Dikotyledonenblätter. Kennzeichen der Gruppe II: Drusenkristalle*, Wien, Moritz Perles.

Netolitzky, F. (1905) *Bestimmungsschlüssel und mikroskopische Beschreibung der einheimischen Dikotyledonenblätter. Kennzeichen der Gruppe: Raphidenkristalle*, Wien, Moritz Perles.

Ostroumova, T. A., and Kljuykov, E. V. (2007) *Feddes Repertorium*, **118**(3–4), 84–102.

Peng, L., and Hu, Z. H. (2007) *Fen zi xi bao sheng wu xue bao = Journal of Molecular Cell Biology/Zhongguo xi bao sheng wu xue xue hui zhu ban*, **40**(6), 395–402.

Pickard, W. F. (2008) *New Phytologist*, **177**(4), 877–887.

Poulsen, V. A., and Trelease, W. (1886) *Botanical micro-chemistry; an introduction to the study of vegetable histology*, S.E. Cassino, Boston.

Prabhakar, M., and Leelavathi, P. (1989) *Asian J. Plant Sci.*, **1**(1), 49–66.

Prabhakar, M. (2004) *Acta Bot. Sin.*, **46**(2), 242–252.

Prabhakar, M., and Leelavathi, P. (1987) *Int. Bot. Congress Abstr.*, **17**, 246–246.

Prantl, K. (1872) *Flora*, **55**305–312, 321–328, 337–346, 369–382.

Prychid, C. J., and Rudall, P. J. (1999) *Ann. Bot.*, **84**(6), 725–739.

Salmaki, Y., Zarre, S., Lindqvist, C., *et al*. (2011) *Plant Systemat. Evolut.*, **294**(1–2), 109–125.

Saukel, J. (1984a) *Sci. Pharm.*, **52**(3), 196–220.

Saukel, J. (1982a) *Sci. Pharm.*, **50**(3), 179–200.

Saukel, J. (1984b) *Sci. Pharm.*, **52**(3), 196–220.

Saukel, J., and Laenger, R. (1992) *Phyton-Annales Rei Botanicae*, **32**(1), 47–78.

Saukel, J. (1983a) *Sci. Pharm.*, **51**(2), 132–158.

Saukel, J. (1983b) *Sci. Pharm.*, **51**(2), 115–132.

Saukel, J. (1984c) *Sci. Pharm.*, **52**(1), 12.

Saukel, J. (1983c) *Sci. Pharm.*, **51**(2), 132–158.

Saukel, J. (1983d) *Sci. Pharm.*, **51**(2), 115–132.

Saukel, J. (1982b) *Sci. Pharm.*, **50**(3), 179–200.

Saukel, J. (1982c) *Sci. Pharm.*, **50**(1), 37–64.

Schacht, H. (1862) *Das Mikroskop und seine Anwendung, insbesondere für Pflanzen-Anatomie*, G.W.F. Müller, Berlin.

Schneider, A. (1921) *The Microanalysis of Powdered Vegetable Drugs*, Blakiston, Philadelphia.

Schroff, C. D. (1853) *Lehrbuch der Pharmacognosie*, W. Braumüller, Wien.

Schürhoff, P. (1906) *Qualitative botanische Analyse der Drogenpulver*, Julius Springer, Berlin.

Shah, A. M., and Kachroo, P. (1975) *J. Indian Bot. Soc.*, **54**(3–4), 138–153.

Sharawy, S. M. (2004) *Feddes Repertorium*, **115**(5–6), 441–452.

Silva, S. P., Sabino, M. A., Fernandes, E. M., *et al.* (2005) *Int. Mater. Rev.*, **50**(6), 345–365.

Singh, V. and Jain, D. K. (1975) *J. Indian Bot. Soc.*, **54**(1–2), 116–127.

Singh, V., Sharma, M., and Jain, D. K. (1974) *Bull. Bot. Survey India*, **16**(1–4), 27–34.

Soh, W. K., and Parnell, J. (2011) *Plant Systemat. Evolut.*, **297**(1–2), 1–32.

Solereder, H. (1908a) *Systematic Anatomy of the Dicotyledons*, Clarendon Press, Oxford.

Solereder, H. (1908b) *Systematic Anatomy of the Dicotyledons*, Clarendon Press, Oxford.

Stöger, E. A. (2005) *Arzneibuch der chinesischen Medizin*, Deutscher Apotheker Verlag.

Tahir, S. S. and Rajput, M. T. M. (2009) *Pak. J. Bot.*, **41**(5), 2137–2143.

Tschirch, A. and Oesterle, O. (1900) *Anatomischer Atlas der Pharmakognosie und Nahrungsmittelkunde*, Herm. Tauchnitz, Leipzig.

Tschirch, A. (1906) Die Harze und die Harzbehälter – mit Einschluss der Milchsäfte, 2. Auflage, Band II, Leipzig, Verlag von Gebrüder Bornträger.

Turner, G. W., Gershenzon, J., and Croteau, R. B. (2000) *Plant Physiol.*, **124**(2), 655–663.

Upton, R., Graff, A., Jolliffe, G., *et al.* (2011) *American Herbal Pharmacopoeia – Botanical Pharmacognosy – Microscopic Characterization of Botanical Medicines*, CRC Press, Taylor & Francis Group, Boca Raton, London, New York.

Vogl, A. E. (1872) *Nahrungs- und Genussmittel aus dem Pflanzenreiche: Anleitung zum richtigen Erkennen und Prüfen der wichtigsten im Handel vorkommenden Nahrungsmittel, Genussmittel und Gewürze mit Hilfe des Mikroskops: zum allgemeinen sowie zum speciellen Gebrauche für Apotheker, Droguisten, Sanitätsbeamte*, Industrielle etc. Verlag der G.J. Manz'sche Buchhandlung, Vienna.

Vogl, A. E. (1899) *Die wichtigsten vegetabilischen nahrungs- und genussmittel, mit besonderer berücksichtigung der mikroskopischen untersuchung auf ihre echtheit, ihre verunreinigungen und verfälschungen Berlin*, Urban & Schwarzenberg, Wien.

Vogl, A. E. (1892) *Pharmakognosie*, Carl Gerold's Sohn, Vienna.

Wagner, G. J. (1991) *Plant Physiol.*, **96**(3), 675–679.

Wagner, G. J., Wang, E., and Shepherd, R. W. (2004) *Ann. Bot.*, **93**(1), 3–11.

Wasicky, R., Fischer, R., Fuchs, L., *et al.*, (1936) *Leitfaden für die pharmakognostischen Untersuchungen in Unterricht und in der Praxis*. in ed. R. Wasicky, Franz Deuticke: Leipzig und Wien; p. 420.

Wiesner, J. (1900) Die Rohstoffe des Pflanzenreiches.

Wigand, A. (1887) *Lehrbuch der Pharmakognosie*, 4th edn, August Hirschwald, Berlin.

Wiggers, A. (1847) *Grundriss der Pharmacognosie*, 2nd edn, Bandenhoeck und Ruprecht, Göttingen.

Winton, A. L. and Moeller, J. (1906) *The Microscopy of Vegetable Foods, with Special Reference to the Detection of Adulteration and the Diagnosis of Mixtures*, Wiley, New York.

Wu, Q., Merchant, F., and Castleman, K. (2008) *Microscope Image Processing*, Academic Press, Elsevier.

Xiang, C.-L., Dong, Z.-H., Peng, H., *et al.* (2010) *Flora*, **205**(7), 434–441.

Zörnig, H. (1912) *Tabelle zur mikroskopischen Bestimmung der offizinellen Drogenpulver. J*, Berlin, Springer.

Thin-Layer Chromatography, with Chemical and Biological Detection Methods

Aurélie Urbain[1] and Claudia Avello Simões-Pires[2]

[1]*Pharmacognosy and Bioactive Natural Products, University of Strasbourg, Illkirch, France* and
[2]*School of Pharmaceutical Sciences, University of Geneva, University of Lausanne, Geneva, Switzerland*

1 INTRODUCTION

Natural sources represent a unique reservoir of bioactive molecules: more than 200,000 natural secondary metabolites have been identified to date, exhibiting a large variety of scaffolds. It is not in doubt that plants are the privileged raw material for the development of high added-value products, with applications in the cosmetic, pharmaceutical, fragrance, and food industries. However, it is estimated that less than 1% of plant species have been thoroughly investigated for their potential use as drugs, and many species have never been studied even from a phytochemical point of view (Payne *et al.*, 1992; Rates, 2001). Therefore, the search for novel bioactive chemical entities is more pressing than ever and requires simple and rapid screening procedures.

Thin-layer chromatography (TLC) is a planar chromatographic technique that is strongly related to the origins of chromatography itself. The chromatographic process was initially invented to separate plant pigments (carotenes and chlorophyll), using a column of calcium carbonate as an adsorbent and a mixture of ethanol/petrol ether as an eluent (Tswett, 1905). This separation technique was first adapted to a thin-layer solid support by Izmailov and Schraiber for the rapid separation of pharmaceuticals (Izmailov and Schraiber 1938; Stahl, 1962). Since then, a number of studies have been conducted to expand the uses of TLC in terms of performance and detection, notably via derivatization or by coupling with spectroscopic and biological methods. Nowadays, TLC is a rapid and nonexpensive method used worldwide for the phytochemical analysis of plant extracts. It is easy to perform and cost effective, as it does not require any heavy equipment. Moreover, TLC is adaptable to extracts of various polarities, and the possibility to simultaneously analyze a broad number of samples makes it a rapid screening method. Appropriate separation on TLC plates can be followed by a number of strategies that allow a comprehensive detection of highly diverse compounds. This turns planar chromatography into a potent tool for the quality control of herbal drugs, initial screening, or bioguided fractionation.

2 GENERAL PRINCIPLES OF THIN-LAYER CHROMATOGRAPHY

The separation principles in TLC are the same as those in liquid–solid chromatography, where the stationary phase is usually an activated adsorbent

such as silica gel, alumina, or cellulose fixed onto a flat support (a glass plate or aluminum foil). The activation step is achieved by heating the thin-layer adsorbent to eliminate residual water. An organic solvent or a mixture to be used as an eluent is added to cover the bottom of a developing tank. The sample is applied to the lower part of the TLC plate (just above the level of the mobile phase), which is then placed in the tank so that the eluent will carry up the sample through capillarity, as shown in Figure 1. Selective separation is given by sorption and desorption of the different compounds in the sample, which depends on their adsorption to the stationary phase and their affinity to the mobile phase. In cases where water is present at small amounts in the mobile phase or over the solid phase, the phenomenon of partition will also take place and influence the separation. After development of the TLC plate, pigments show up directly in daylight, whereas other products can be visualized either by their natural fluorescence or by quenching of fluorescence. In case they are barely detected in visible or ultraviolet light, or if additional structural information is required, universal or specific detection reagents can be used to form colored or UV-absorbing by-products. In addition to chemical data, information concerning biological activity might also be obtained by biologically oriented reactions. Indeed, natural products separated on the TLC plate can be subsequently exposed to specific biological or chemical reagents to detect a potential activity (Figure 1).

3 ADVANCES IN TLC FOR PLANT ANALYSIS

Plant extracts are highly complex mixtures and the quality of the separation is of the utmost importance to obtain meaningful results. Spotting, development, and visualization are three essential steps of TLC that interfere with the quality of results. The application of the plant extract onto the plate and the development of the chromatogram are two critical steps: the initial spot should be as regular and as narrow as possible and the development should be homogeneous and reproducible. TLC has benefited from advances in recent years, resulting in better resolution and a more accurate localization of phytochemical markers.

The development of high performance thin-layer chromatography (HPTLC) plates with smaller particles, thinner layers, and a narrow particle-size distribution offers faster separation by reducing the required distance of migration. HPTLC also appears to be more sensitive and reproducible than classical TLC. Considering that the distance of migration in HPTLC can be half that of common TLC, the cost of the plate to perform the same analysis is nearly the same for both techniques. However, HPTLC is more costly in terms of initial financial investment for acquiring the main and auxiliary apparatus and software.

It is noteworthy that automated systems have been developed for sample application and chromatogram development: devices spraying extracts as thin bands can dramatically improve the application procedure,
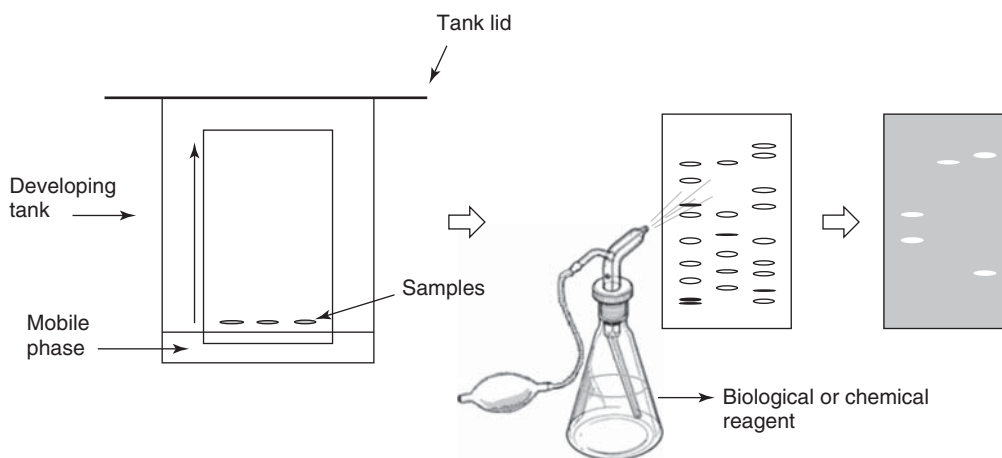
Figure 1   General principle of TLC development and detection. Samples (e.g., plant extracts) are applied to a TLC plate to be placed in a developing tank containing the mobile phase. After migration, the plate is dried and a solution containing chemical or biological reagents can be sprayed onto the TLC plate, leading to visible spots with a specific chemical characteristic or biological activity.

and irregular curve-like elutions can be avoided using an automatic developing chamber, which controls the relative humidity. One of the latest innovations in the field of TLC, the automated multiple development (AMD), enables multistep gradient elution, which dramatically improves separation power. The fully automated process consists of successive developments of the plate with solvent mixtures of decreasing polarities, running each time over a longer migration distance, and with drying steps under vacuum between each partial run. The use of AMD-HPTLC provides a spectacular enhancement of separation and resolution, leading to the detection of dozens of compounds within a shorter migration distance (Yan *et al.*, 2010).

Structural data related to the separated compounds can also be assessed. The coupling of TLC or HPTLC with mass spectrometry (MS) has been relatively useful in determining the molecular weight of separated compounds, providing more information than an HPLC-UV-DAD analysis (Morlock and Schwack, 2010). Usually, the TLC foils or plates are dried and initially analyzed by UV detection or even a bioautographic assay to locate the bands of interest. For UV detection, selected bands can be cut from the whole sheet and placed directly under the excited helium stream of a time-of-flight mass spectrometer equipped with an ion source of direct analysis in real time (DART-TOF-MS) (Kim *et al.*, 2010). An online extractor can also transfer the regions of interest from the plate directly into an ESI-MS

by means of an appropriate solvent to obtain an HPTLC-ESI-MS spectrum (Figure 2). The advantage of the extraction procedure is that it can even be employed directly following antimicrobial autobiographic techniques. In this context, the chosen organic solvent (e.g., methanol) will extract the compound of interest and precipitate proteins from the biological medium. Using the same interface, compounds can also be desorbed from the layer into a vial. After thorough removal of the solvent under a stream of nitrogen, the product can be dissolved in deuterated solvent for subsequent nuclear magnetic resonance (NMR) spectroscopic measurements (Gössi *et al.*, 2012). The direct TLC-MS coupling and the indirect hyphenation with NMR, therefore, provides a full set of spectral data, enabling the structural elucidation of natural products separated by TLC.

In addition to spectral or biological information, semiquantitative data can be retrieved from TLC by comparing the light that is scattered, reflected, or generated from a spot, with that of the background, that is, a part of the plate devoid of any compound. This can be carried out directly on the raw TLC plate or after derivatization with a chemical or biological reagent, for example, to estimate a biological activity. Over the past 15 years, the scanning of TLC plates employing optical instrumentation has been increasingly used, notably in the field of quality control of herbal drugs, where both fingerprint and quantitative assessment are required (Apers *et al.*, 2006; Bhandari *et al.*, 2007; Bhope *et al.*, 2011). Herbal raw
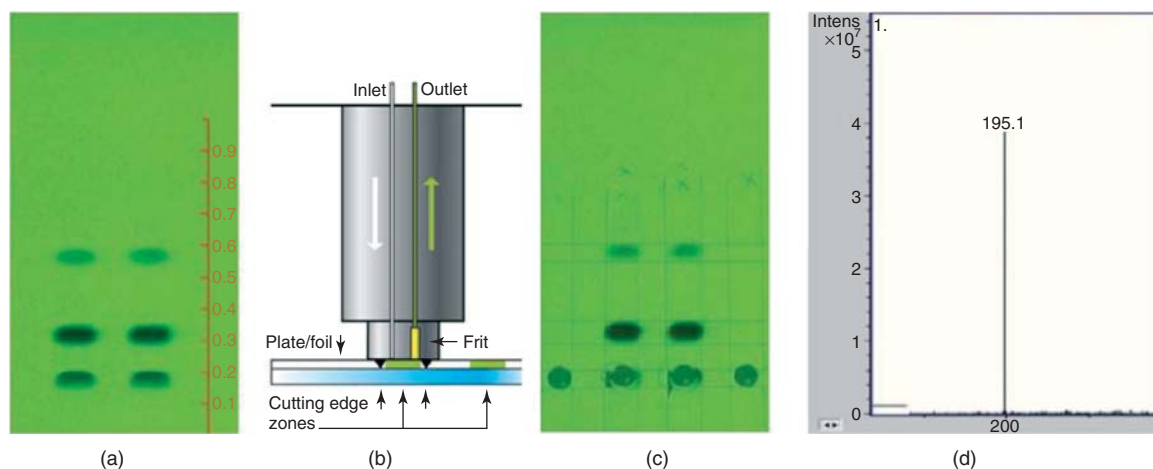
Figure 2   HPTLC-ESI-MS analysis of caffeine. (a) Chromatogram with 4-mm bands; (b) scheme of extraction piston; (c) same plate after extraction of zone at hRf 15; (d) extracted zone identified as caffeine, based on the mass signal at *m/z* 195 in positive mode. (*Source*: Reproduced by permission of Gertrud Morlock, Camag (CBS 102, March 2009).)

material or herbal drugs might be contaminated by toxins released into the environment, such as pesticides. These herbal compounds can also be adulterated with synthetic drugs or contaminated with other herbal drugs. The control of markers by TLC using various detection methods has been developed for all classes of natural products and reference books are helpful for the work of a plant analyst (Wagner and Bladt, 1996). In an attempt to provide semiquantitative data valuable in the evaluation of markers in plant extracts, densitometric measurements have been applied to common TLC to provide an image that can be treated for dose-response purposes. One example of the application of a simple TLC method is the quantification of the fluorescent alkaloid berberine in batches of *Argemone mexicana* L. (Papaveraceae), a plant from which the decoction has shown clinical efficacy in Africa (Willcox *et al*., 2007). TLC followed by the detection of fluorescence has been suggested as a cheap quality control method for an herbal medicine approved in Mali for the treatment of uncomplicated malaria. A fluorescence image of the plate containing a calibration curve of the standard berberine and the extracts of *A. mexicana* was treated by image conversion to a gray scale to provide densitometric values, enabling an estimation of the berberine content (Figure 3) (Simões-Pires, 2009; Simões-Pires *et al*., 2009b).

4 TLC BIOAUTOGRAPHIC ASSAYS

The first attempts to obtain a biological response from planar chromatography were made in 1946 to detect whether natural products could affect the growth of microorganisms (Goodall and Levi, 1946). As this type of assay characterizes a direct set of active compounds against living organisms, it was therefore named bioautography. Nowadays, this term has been widely accepted to refer to any method used to assign bioactivity to compounds by combining chromatographic separation before a bioassay. Indeed, although the term TLC bioautography should strictly involve a living organism, it was extended by consensus to designate any TLC separation followed by the application of a reagent (biological or chemical) to detect a biological response. Owing to its numerous advantages, TLC bioautography has spread throughout the world as a simple and useful "bench top" bioassay to track bioactivity through all the fractionation processes. Currently, bioautographic assays

are often used to detect antimicrobial compounds, and they are also widely applied to the discovery of antioxidants or free radical scavengers and the development of bioautographic assays to detect enzyme inhibition is promising.

4.1 Antimicrobial Bioautographic Assays

Historically, the first experiments of what we currently call bioautographic assays used to involve bacteria for the detection of antibiotics or growth factors such as vitamin B derivatives (Goodall and Levi, 1946; Lewin and Marcus, 1965; Matsuda and Goto, 1952; Picken and Bauriedel, 1950). Bioautography on TLC is particularly well adapted for the detection of antimicrobial compounds that affect the growth of bacteria, fungi, or protozoa (Betina, 1973). Three main variants of bioautographic assays are encountered. The most commonly used is direct bioautography, where a microorganism suspension is applied directly onto the plate after chromatographic separation. Agar diffusion and immersion prevailed in the 1960s but are now used only when direct bioautography is impossible.

The main challenge of any TLC bioautographic assay is the visual detection of activity. Thus, it makes sense that the first tested organisms were spore-producing fungi, such as *Aspergillus* or *Cladosporium* species that form easily observable green-to-black spores. Subsequently, antimicrobial bioautography spread to a wide range of organisms because of the development of reagents enabling microbial growth visualization, such as tetrazolium salt derivatives, which are reduced to purple formazan in living cells. In both cases, zones of growth inhibition appear as pale areas on a colored background.

4.1.1 Direct TLC Bioautography

Direct bioautography involves the application of a microorganism suspension directly onto the developed TLC plate, either by spraying or dipping. This method is therefore applicable to any non-pathogenic strain able to grow on a TLC layer. If experiments deal with pathogenic bacteria or fungi, they have to be carried out with extreme caution to avoid unwanted contaminations, by working, for

Argemone mexicana

$$y = 0.3781x + 90.55$$
$$R^2 = 0.9125$$

Batch	Concentration in tea (μg mL^{-1})*	% in plant (w/w)*
AM 8	18.85 ± 0.83	0.02 ± 0.002
AM 9	17.57 ± 1.12	0.03 ± 0.002
AM 11	19.54 ± 0.88	0.03 ± 0.001

* Average ± SD.
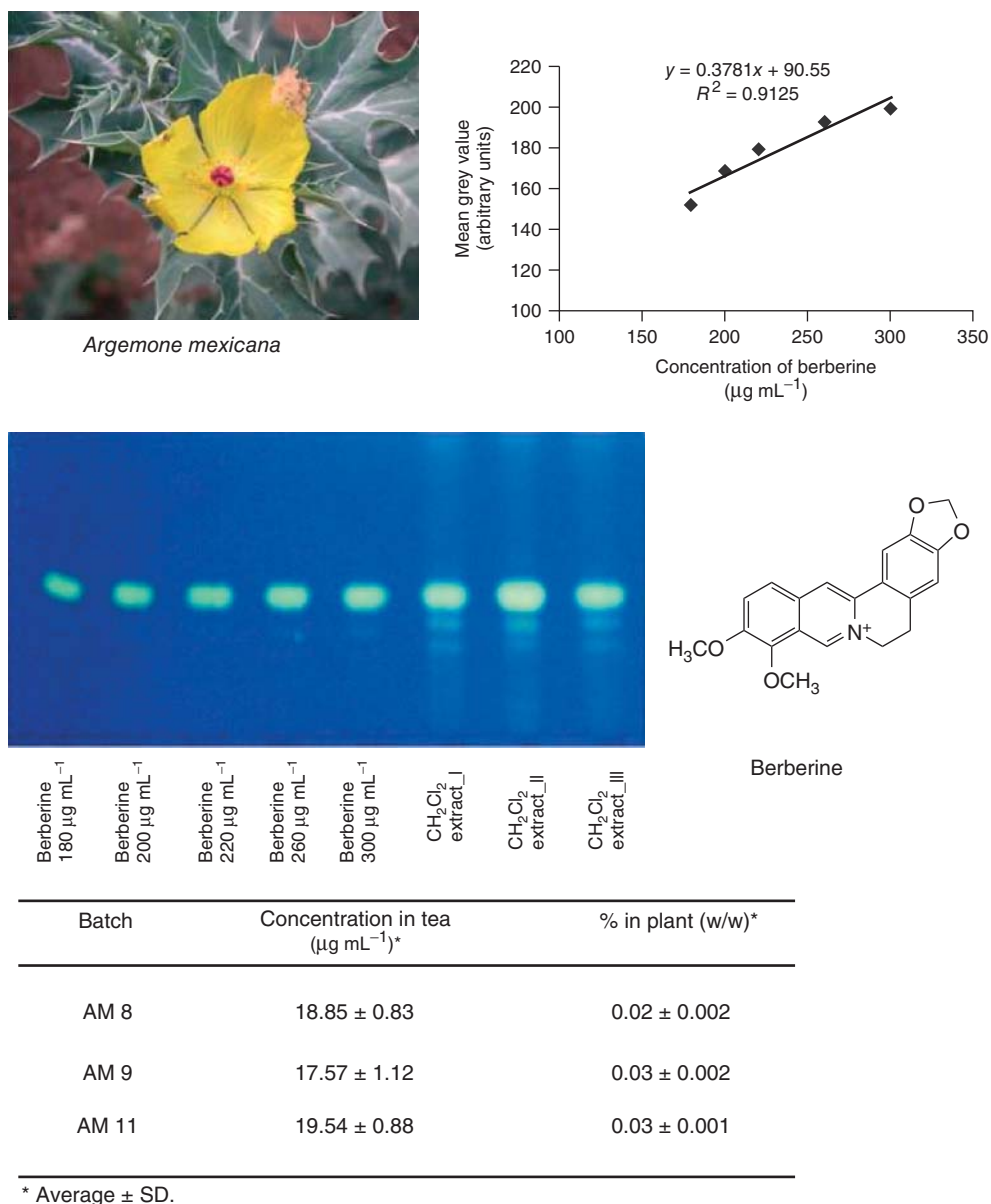
Figure 3 Semiquantitative TLC method for the determination of berberine in *Argemone mexicana* decoction. The density of fluorescent spots of standard berberine was measured by the public domain software ImageJ at different concentrations to produce a calibration curve. The decoction of clinical batches of the plant was extracted with dichloromethane. Sample concentrations were calculated using the linear equation provided by the calibration curve and were extrapolated to the total volume of decoction. (*Source*: C. Simões-Pires (2009). Reproduced with permission © C. Simões-Pires.)

example, in a biosafety Class II cabinet. Contrary to immersion, direct bioautography is performed without agar: immediately before use, microorganisms are usually transferred from their solid culture medium to a sterile fresh broth containing all required nutrients for proper microbial growth and having a viscosity compatible with spraying if necessary. After appropriate incubation, zones of growth inhibition are observed naturally for spore-producing fungi and pigmented bacteria or revealed via tetrazolium
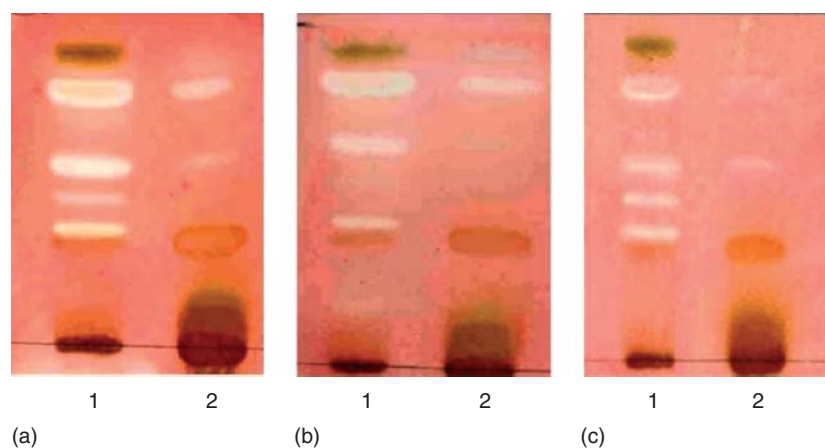
Figure 4 Antimicrobial TLC bioautography using *p*-iodonitrotetrazolium violet as an indicator of microbial growth. Comparison of the antimicrobial chemical components in leaves (1) and stem bark (2) of *Curtisia dentata*. Acetone extracts of leaves and barks were analyzed on TLC against *Escherichia coli* (a), *Staphylococcus aureus* (b), and *Candida albicans* (c). Clear bands indicate growth inhibition of the microorganisms. (*Source*: Reprinted from S. Afr. J. Bot., 75, Shai L.J., McGaw L.J., Eloff J.N., Extracts of the leaves and twigs of the threatened tree *Curtisia dentata* (Cornaceae) are more active against *Candida albicans* and other microorganisms than the stem bark extract, 363–366, Copyright 2009, with permission from Elsevier.)

salts (Figure 4) (Shai *et al*., 2009). An alcoholic solution can be subsequently sprayed onto the entire layer to fix microorganisms.

4.1.2 Agar Diffusion Bioautography

Agar diffusion or contact bioautography is rarely used nowadays. In this method, compounds are firstly separated on TLC, and after complete removal of the solvent, the TLC layer is applied face-down directly onto an agar medium previously inoculated with a given microorganism (Wagman and Bailey, 1969). The layer is left for 30 min to 1 h to allow diffusion, before the TLC is removed and the agar gel is incubated in suitable conditions. After an appropriate time, zones of growth inhibition are compared to the TLC to locate antimicrobial compounds. The disadvantage of this method is the irregularity of the diffusion and the possible dilution of substances during transfer to the agar gel.

4.1.3 Immersion Bioautography

Immersion bioautography is an intermediate technique between agar diffusion and direct bioautography. In this approach, the developed TLC is overlaid with a seeded agar gel, enabling

a uniform contact between the chromatogram and the inoculated medium. After solidification of the gel, plates are incubated and subsequently observed after development when needed (Haouat *et al*., 2013; Hostettmann and Marston, 1994; Rahalison *et al*., 1991). This procedure is a good alternative when direct bioautography is not possible, for example, when blockage of the spray head with mycelia occurs or when spraying should be avoided to minimize microbial spreading and contamination within the laboratory. However, as for contact bioautography, the sensitivity is quite low because of the dilution of active substances throughout the agar gel.

4.2 Toxicity Bioautographic Assays

Bioluminescent bacteria such as *Vibrio fischeri* or *Photobacterium phosphoreum* are microorganisms that mostly occur in marine environments, which are characterized by a readily visible light emission. This glowing process is due to a luciferase-catalyzed chemiluminescent reaction releasing chemical energy in the form of blue-green light and is closely related to other metabolic processes. When these bacteria are exposed to poisonous substances, the metabolic activity is affected, resulting in a proportional decrease in luminescence. This phenomenon can be exploited in TLC bioautographic assays to

detect natural toxins (Eberz *et al*., 1996; Sherma, 2010). Even if this bioassay involves bacteria, it cannot be considered as an actual antimicrobial bioautographic test, as it does not show a genuine growth inhibition but rather a disruption of the respiratory chain. In practice, the developed TLC plate is exposed to a solution of luminescent bacteria by dipping, and if toxic compounds are present, they will interfere with metabolism and appear after few minutes as dark spots against a luminescent background.

Currently, there are two bioluminizer systems available for the measurement of luminescence in TLC plates: Bioluminex™ (ChromaDex) and BioLuminizer® (Camag). These have been applied to detect natural toxins such as mycotoxins and heavy metals and also to evaluate the toxicity of natural secondary metabolites (Klöppel *et al*., 2008, 2013) or to detect adulterants for the quality control of phytomedicines (Verbitski *et al*., 2008).

4.3 Seed Germination and Growth-Inhibition Bioautographic Assays

In addition to microorganisms, toxicity toward superior living organisms such as plants can be assessed directly via TLC. Some secondary metabolites synthesized by plants positively or negatively influence the growth, survival, and reproduction of other organisms. These compounds, termed allelochemicals, can be of great interest, especially as organic herbicides in agriculture.

In conventional research for allelopathic substances, the most common bioassay consists of petri-dish germination assays, in which the crude extract to be tested is contained within the agar gel. The identification of phytotoxic compounds is therefore time consuming as it involves fractionation steps routinely followed by experiments in petri dishes to track the active substances. Using the TLC bioautographic assay, the identification of phytotoxic compounds is greatly accelerated. After development and proper drying of the TLC plate, a solution of 0.5% agar is poured over the plate to obtain a 2-mm-thick layer. Then, seeds are sown in the agar all along the band corresponding to the developed extract and incubated in appropriate conditions of humidity, temperature, and photoperiod (Inoue *et al*., 1992). After a few days, the absence of germination or a significant difference in seedling growth reveals the presence of phytotoxic compounds, as illustrated

Hexane–CHCl$_3$–MeOH (4:10:1)
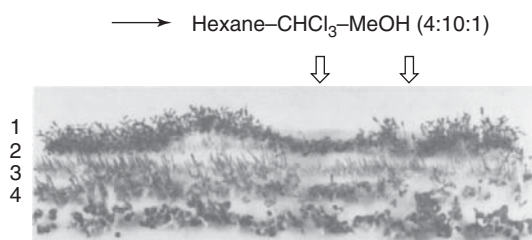
Figure 5 Growth inhibition TLC bioautographic assay of fractions from *Polygonum sachalinense* (5 days). Arrows point to areas of inhibited plant activity. (1) Green amaranth, (2) timothy grass, (3) crab grass, and (4) Chinese cabbage. (*Source*: Reproduced from J. Chem. Ecol., 18, 1992, 1833, Allelochemicals from *Polygonum sachalinense* Fr. Schm. (Polygonaceae), Inoue M., Nishimura H., Li H.H., Mizutani J. With kind permission from Springer Science and Business Media.)

in Figure 5. Different allelochemical-sensitive seeds can be used in this bioassay if they are small enough and easy to grow, such as (L.) Heynh. (Brassicaceae) seeds, lettuce (*Lactuca sativa* L., Asteraceae), or green amaranth (*Amaranthus viridis* L., Amaranthaceae) seeds. To date, this bioautographic assay has rarely been used, although it is relatively easy to perform and rapidly provides information about germination or growth inhibitors. For example, anthraquinone and naphthoquinone derivatives have been proved to be potent germination inhibitors using this TLC bioassay (Meyer *et al*., 2007).

4.4 Antioxidant and Free-Radical Scavenging Activities

Oxidative stress is mainly associated with an increased production of reactive oxygen species (ROS) such as free radicals and peroxides. It is now well attested that oxidative stress is one of the main mechanisms involved in the development of several disorders such as cancer, neurodegenerative, vascular, and heart diseases. Furthermore, it has been observed that a high intake of vegetables and fruits reduces the risk of developing these disorders, suggesting that dietary compounds might act as potent antioxidants. Consequently, the search for natural antioxidants and free radical scavengers represents a large part of natural products research, and rapid screening assays are therefore valuable. Many *in vitro* methods are currently available, and some of these can be applied to TLC for postchromatographic derivatization (Huang *et al*., 2005).

4.4.1   The β-Carotene Bleaching Assay

The antioxidant activity of natural products can be observed by their ability to prevent the discoloration of β-carotene, induced by the oxidative degradation products of linoleic acid (Marco, 1968); β-carotene bears a long unsaturated chain resulting in an intense orange color, which disappears after reaction with linoleic acid, because of the disruption of conjugation.

For this purpose, developed and dried TLC plates are sprayed with an emulsion of 0.03% β-carotene in chloroform and 0.01% linoleic acid in ethanol (1 : 2 v/v) (Philip, 1974). After 2- to 3-h exposure to daylight, the yellow-orange background progressively bleaches, whereas the coloration remains in zones containing antioxidant products. Bleaching of β-carotene can also occur without linoleic acid, by simple exposure to daylight, although the reaction will proceed more slowly. In this case, TLC plates are sprayed with a solution of β-carotene only, and left in sunlight for at least 12 h until bleaching of the background occurs (Marston, 2011).

4.4.2   The ABTS•+ Scavenging Assay

The occurrence of radical scavengers can be checked using 2,2′-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS). This colorless compound exists in a relatively stable radical cationic form, ABTS•+, which exhibits a green-blue color (Lee and Yoon, 2008). When reacting with radical scavengers, this radical returns to its colorless neutral form, and antioxidant activity can be directly related to the fading of color (Miller *et al*., 1993; Miller and Rice-Evans, 1997). This assay is often used in the food and beverage industries, where the antioxidant activity of products is compared to that of Trolox, an analog of vitamin E, and is therefore referred to as the *Trolox equivalent antioxidant capacity* (TEAC) assay. It can be monitored in spectrophotometer cuvettes following the decrease in absorbance at 415 or 734 nm, and it can also be adapted to TLC plates, where radical scavengers will appear as colorless spots against a green background. However, when seeking radical scavengers from plant extracts, the DPPH radical is preferred over ABTS•+ as it gives a more stable coloration and better defined spots.

4.4.3   The DPPH• Scavenging Assay

DPPH• (2,2-diphenyl-1-picrylhydrazyl) is a stable free radical with a deep violet color due to an absorption band at 517 nm. However, when DPPH• is neutralized by radical scavengers, there is a modification of its UV spectrum and the dark purple color progressively turns pale yellow (Figure 6) (Blois, 1958). This discoloration enables the monitoring of radical scavenging activity, which can be followed either by spectrophotometric measurements at 517 nm (Sharma and Bhat, 2009) or by TLC assay (Glavind and Holmer, 1967; Takao *et al*., 1994). In the latter, an alcoholic solution of DPPH• (generally 0.2% w/v) is sprayed onto the developed plate, leading within 30 min to pale yellow spots against a purple background, where substances demonstrating scavenging properties toward the DPPH radical are located. When DPPH• assays are carried out on reversed-phase TLC (RPTLC), the background color is relatively unstable and tends to fade within 3 min
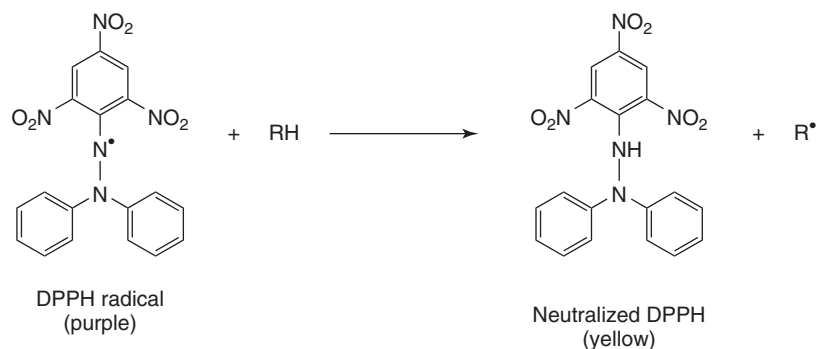
Figure 6   Neutralization of the 2,2-diphenyl-1-picrylhydrazyl radical (DPPH•) by radical scavengers.

after dipping, leading to blurred active spots (Yrjönen *et al*., 2003). It should be kept in mind that the nature of the TLC layer might greatly influence the chemical or biological response. This point is discussed further (see Section 4.8).

4.5 Enzymatic Bioautographic Assays

As enzymes represent important pharmacological targets, the development of suitable and rapid screening methods for enzyme inhibitors is a key element in the field of drug discovery. Along with antimicrobial and antioxidant compounds, natural enzyme inhibitors can be detected by TLC. Enzymatic bioautographic assays require the successive application of the enzyme and a suitable substrate on the entire TLC layer, which react together to produce a colored product, except in areas containing inhibitory compounds because of the inactivation of the enzyme, and consequently to the absence of the chain reaction. To date, such assays exist to detect cholinesterase, glucosidase, lipase, tyrosinase, and xanthine oxidase inhibitors. The main difficulty in the development of these bioautographic assays is to find an appropriate substrate that will give a product exhibiting a different color after reaction to enable the visual detection of inhibitory activity. Furthermore, when studying enzyme activities, it is important to bear in mind that near physiological conditions are required to recover maximum activity. For this reason, TLC plates should not be dried thoroughly with a hair dryer or using too high temperatures after spraying the enzyme solution to not denaturate it. Similarly, it is also better to preincubate the enzyme at 37 °C or an optimal temperature in a humid chamber before adding the substrate to ensure that the enzyme will display full activity.

4.5.1 Acetylcholinesterase and Butyrylcholinesterase Inhibition

Alzheimer's disease (AD) is the main form of dementia among the elderly and is becoming more prevalent with the increase in mean life expectancy, especially in developed countries where AD is one of the most costly diseases to society. The cognitive impairments in AD are mainly due to a cholinergic deficit associated with neuronal loss. As the neurotransmitter acetylcholine is regulated in brain synapses by the hydrolytic action of acetylcholinesterase (AChE, EC 3.1.1.7), cognitive dysfunctions can be partially improved by the use of AChE inhibitors, leading to an increase in acetylcholine levels. This is the case for galanthamine, an alkaloid isolated from some Amaryllidaceae species such as snowdrop (*Galanthus nivalis* L.), currently approved to restore the cholinergic deficit caused by AD. In addition, recent studies have shown that butyrylcholinesterase (BuChE) could be further considered as a significant therapeutic target for the development of new drugs. To accelerate the discovery of new cholinesterase inhibitors among plant extracts, effective and rapid assays are needed (Cieśla, 2012; Hostettmann *et al*., 2006). In this framework, several colorimetric bioassays have been developed. These are all based on the same principle of the coupling of two successive reactions: the enzymatic cleavage of an ester followed by the subsequent reaction of the end product with a chromogenic agent. In these bioautographic assays, AChE inhibitors, therefore, appear as colorless spots on a colored background.

The first bioautographic assay developed for AChE inhibitors was based on Ellman's reaction, which was initially conducted in large cuvettes for kinetic measurements of acetylcholinesterase activity (Ellman *et al*., 1961). In this method, acetylthiocholine (ATC) is cleaved by AChE to form thiocholine, which in turn reacts with 5,5′-dithiobis-(2-nitrobenzoic acid) (DTNB) to give a yellow 5-thio-2-nitrobenzoate anion as depicted in Figure 7.

This method was adapted for TLC plates 30 years later for the qualitative screening of AChE inhibitors (Kiely *et al*., 1991). However, samples were simply spotted onto the TLC plate in this study but were not developed, so further adjustments were made to validate the effectiveness of this bioautographic assay for the screening of natural AChE inhibitors from crude plant extracts. Rhee *et al*. (2001) have applied the methodology to pure galanthamine as a positive control, together with Amaryllidaceae extracts known to contain galanthamine and other active derivatives. After migration of the extracts with an appropriate solvent system, a solution of ATC and DTNB is first sprayed onto the silica layer, and then a solution of AChE is applied. A pale yellow coloration should appear within about 5 min on the TLC plate. This reaction does not occur when inhibitory compounds are eluted, and these active products are consequently revealed as white spots on the yellow background. In
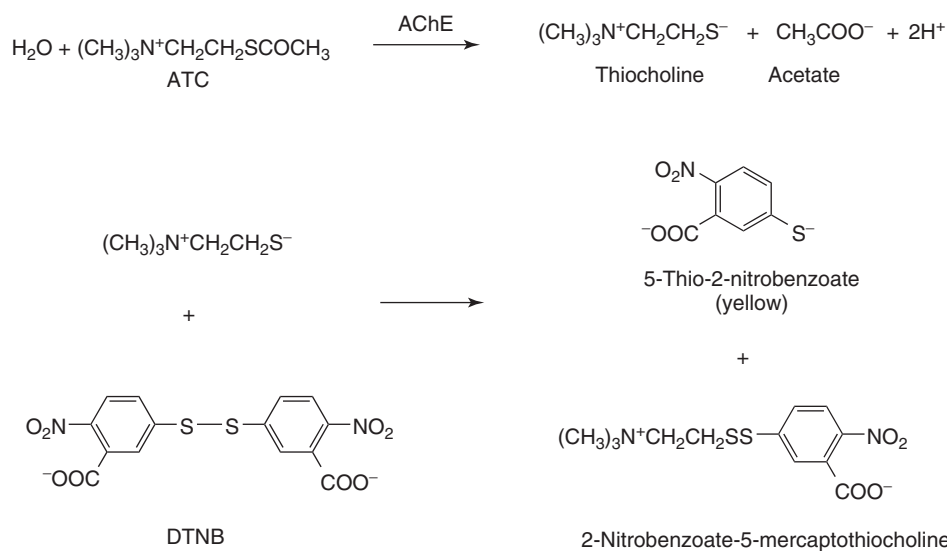
Figure 7 Ellman's reaction.

this study, the limits of detection (LOD) were found to be in the range 0.01–0.20 μg for several known inhibitors (e.g., 0.01 μg for physostigmine), proving that this bioassay is convenient for screening of plant extracts at quantities usually loaded onto TLC plates. These limits were even much lower for some compounds than for those with UV detection at 254 nm or with Dragendorff's reagent and enabled the identification of natural inhibitors in the studied extracts.

The enzyme concentration is set at $3\,U\,mL^{-1}$ to obtain an acceptable contrast without consuming too much of the enzyme, which is quite expensive. Indeed, spots can be difficult to distinguish because of the weak yellow color intensity, and this is one of the main drawbacks of the bioautographic assay based on Ellman's reaction. Moreover, the interpretation of results should be performed carefully, as some compounds (mostly aldehydes and amines) are known to give false-positive reactions by inhibiting the reaction between thiocholine and DTNB (Rhee *et al.*, 2003). However, this method was, and still is, often used for preliminary screening or for the bioguided isolation of natural AChE inhibitors (Adsersen *et al.*, 2006; Beedessee *et al.*, 2013; Dall'Acqua *et al.*, 2010; Houghton *et al.*, 2004).

In the meantime, another method was developed, inspired by a combination of studies, including a study of housefly esterases by colorimetric means (van Asperen, 1962), as well as methods developed

to detect organophosphorus pesticides or carbamates, because these act by inhibiting AChE (Mendoza *et al.*, 1968; Weins and Jork, 1996). In this bioassay, 1-naphthyl acetate is used as initial substrate and Fast Blue B salt as a chromogenic agent (Marston *et al.*, 2002). The test was applied to known anticholinesterase alkaloids and to different Amaryllidaceae extracts and extended to butyrylcholinesterase. After elution of the plant extracts, the TLC plate is dried and sprayed with the enzyme solution and incubated for a few minutes at 37 °C. A freshly prepared solution of 1-naphthyl acetate and Fast Blue B salt is then sprayed onto the plate. Active enzymes convert 1-naphthyl acetate into 1-naphthol, which subsequently couples to Fast Blue B salt to give a purple diazonium dye visible after a few minutes (Figure 8). As the contrast is greatly accentuated compared to the bioassay based on Ellman's reaction, the purple coloration enables an easier detection of bioactive compounds, which appear as white spots, as observed in Figure 9 (Adhami *et al.*, 2013; Urbain *et al.*, 2004, 2005). This probably explains why the detection limit is lower for some compounds (e.g., 1 ng for physostigmine). Similar results were obtained using butyrylcholinesterase, even though sensitivity to the inhibitory compounds can vary slightly.

Further modifications have been made to this bioautographic assay to improve both sensitivity and
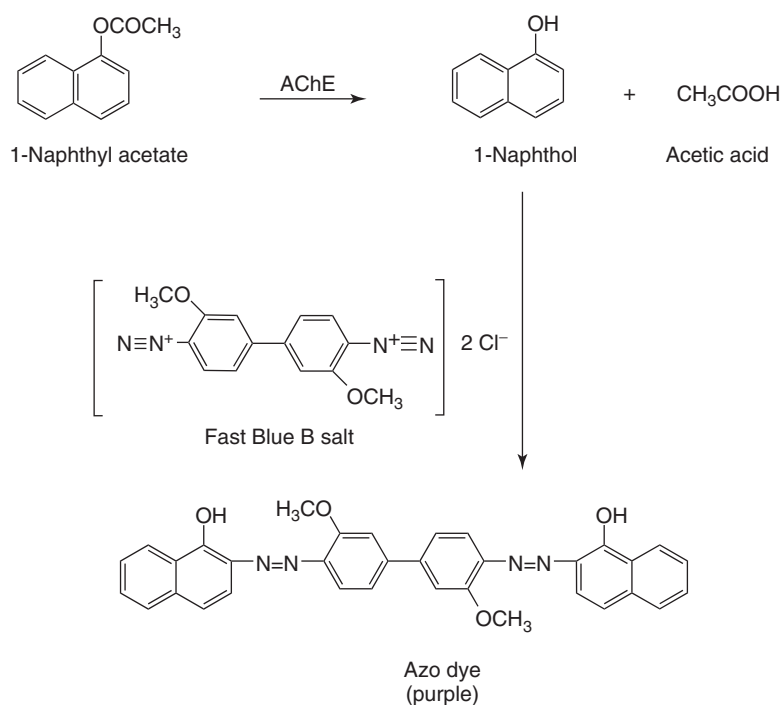
Figure 8 Reaction of acetylcholinesterase with 1-naphthyl acetate and subsequent diazotization.

| A5 | A6 | Chelidonine | A5 | A6 | Chelidonine |
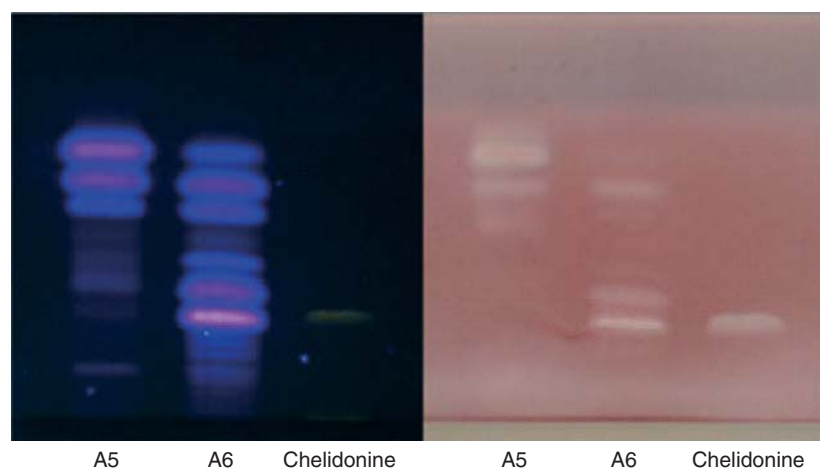
Figure 9 TLC of fractions A5 and A6 after vacuum liquid chromatography (VLC) of a dichloromethane extract from galbanum. Detection at UV 366 nm (left). Detection via the bioautography assay with AChE inhibition (right). Mobile phase: CHCl₃–EtOAc–MeOH (100 + 10 + 2). Chelidonine served as a positive control. (*Source*: H-R. Adhami, *et al*. (2013). Reproduced by permission of John Wiley & Sons, Ltd.)

cost. Attempts were notably made to reduce the consumption of the enzyme, which is quite expensive. In Marston's protocol, the acetylcholinesterase concentration is $6.7\,\mathrm{U\,mL^{-1}}$. Yang *et al*. (2009) reduced

this concentration to $1\,\mathrm{U\,mL^{-1}}$, but to maintain an intense purple coloration, simultaneously decreased the amount of Fast Blue B salt and increased the concentration of 1-naphthyl acetate. Otherwise, the

coupling of 1-naphthol with the stain is incomplete and produces an unstable azo-product still containing one free diazoamino group, which gives a pale and transient coloration. Therefore, the modified method enables the enzyme consumption to be reduced by 85%, while dramatically improving the sensitivity, as the lowest amount of physostigmine required to observe a white inhibition spot is 0.01 ng, that is, 100-fold less than with Marston's method and 1000-fold less than with the Ellman's method modified by Rhee.

Similarly, different parameters such as reagent concentration, incubation time, and even the nature of the TLC plate were tested to increase sensitivity and simultaneously reduce reagent consumption (Mroczek, 2009). Even if 1-naphthyl acetate has a stronger affinity to AChE (van Asperen, 1962), and is also less expensive, 2-naphthyl acetate was chosen as the esterified substrate, as it leads to a more pronounced coloration, which appears more rapidly (within 1 min) and remains longer (at least 24 h). According to this optimized protocol, the TLC layer is sprayed with a solution of AChE at 3 U mL^{-1}, incubated for 10 min at 37 °C in a humid atmosphere, and subsequently sprayed with a Fast Blue B salt water solution at 1.25 mg mL^{-1}. The initial substrate is not sprayed over the layer in this assay, but added directly to the elution solvent system, with an optimum concentration of 2-naphthyl acetate set at 1.5 mg mL^{-1} of mobile phase. This provides a homogeneous and adequate substrate repartition on the plate and greatly increases the sensitivity of the bioautographic assay (the LOD for galanthamine is reduced to 0.36 ng). Another notable aspect of this study is evidence of the influence of the stationary phase: sensitivity was significantly higher when the assay was performed on aluminum oxide plastic plates, extending the limit of detection of galanthamine to 4.5 pg, possibly because of the spherical distribution of the spots on the Al$_2$O$_3$ sheets. This modified Fast Blue B method was further used for the qualitative assessment of various extraction techniques applied to Amaryllidaceae alkaloids (Mroczek and Mazurek, 2009).

Although many improvements in terms of sensitivity and cost have been brought to the method using naphthyl acetate and Fast Blue B salt, the development of novel TLC bioassays remains challenging. The most recent method involves 4-methoxyphenyl acetate. This ester is a suitable substrate of AChE, leading to a deep blue coloration following subsequent reaction with a specific chromogenic agent.

The active enzyme converts 4-methoxyphenyl acetate into its phenol derivative, which in turn reacts with a mixed solution of potassium ferricyanide (K$_3$Fe(CN)$_6$) and iron chloride hexahydrate (FeCl$_3$·6H$_2$O) to produce an aquamarine blue color on the TLC plates. Against this background, acetylcholinesterase inhibitors appear not as white but as light yellow spots (Yang *et al.*, 2011). The weakness of this study is that the chromogenic agent is not sprayed onto the layer but soaked onto filter paper subsequently applied to the TLC surface for 1 min. After processing in this way, the background does not appear to be homogeneous and inhibition spots are not clearly distinct. Furthermore, attempts to establish the LOD estimated a minimum inhibitory amount of 1 ng for physostigmine. Thus, this method is 100-fold less sensitive than the modified Marston's method previously published by Yang *et al.* (2009) and does not appear to offer any advantage compared to existing bioautographic assays. Nevertheless, further assays should be attempted by spraying or dipping the layer, instead of dabbing it with soaked paper: a proper and uniform application of the chromogenic solution might improve the resolution of this bioassay.

The complete range of TLC bioautographic methods developed to detect AChE inhibitors is referenced in Table 1.

4.5.2 Lipase Inhibition

The prevalence of obesity is increasing alarmingly in developed countries and if not managed can result in secondary diseases such as type 2 diabetes and cardiovascular diseases. Even though physical exercise and a balanced diet are the first necessary controlling steps, medication can help reduce or control weight. Only one drug is currently approved by the FDA for the long-term treatment of obesity: orlistat, a derivative of the natural product lipstatin, a potent natural inhibitor of pancreatic lipases (EC 3.1.1.3). Lipase inhibitors alter the gastrointestinal absorption of fats: triglycerides are not hydrolyzed into absorbable free fatty acids and are excreted in feces, resulting in a reduction in calorie intake.

Lipases belong to the group of carboxylic ester hydrolases (EC 3.1.1), as do cholinesterases. For this reason, it is possible to apply methods developed from the study of cholinesterase inhibitors to the inhibition of lipases. Using Marston's protocol, orlistat

Table 1 Main features of TLC bioautographic assays developed for acetylcholinesterase inhibition.

Method	Substrate	Chromogenic agent	AChE concentration $(U\,mL^{-1})$	LOD Physostigmine	LOD Galanthamine
Rhee *et al.* (2001)	Acetylthiocholine (ATC)	5,5′-Dithiobis-(2-nitrobenzoic acid)	3	10 ng	10 ng
Marston *et al.* (2002)	1-Naphthyl acetate	Fast Blue B salt	6.7	1 ng	10 ng
Mroczek (2009)	2-Naphthyl acetate	Fast Blue B salt	3.0	/	0.36 ng 4.5 pg[a]
Yang *et al.* (2009)	1-Naphthyl acetate	Fast Blue B salt	1.0	0.01 ng	/
Yang *et al.* (2011)	4-Methoxyphenyl acetate	Potassium ferricyanide + iron chloride hexahydrate	1.0	1 ng	/

[a] On Al_2O_3 plastic plates.

can be detected as a white inhibition spot at a concentration of up to 0.01 µg, and two plant extracts known to contain lipase inhibitors, when tested, led to the emergence of several inhibition spots after postchromatographic derivatization (Hassan, 2012). Similarly, a patent was registered for a TLC bioautographic assay applied to the screening of lipase inhibitors, using Fast Blue B salt as a chromogenic agent. The sole noteworthy novelty is the use of 2-naphthyl myristate as an initial substrate (Wu *et al.*, 2013).

4.5.3 α- and β-Glucosidase Inhibition

Glucosidases (EC 3.2.1.20) are considered as important therapeutic targets in type 2 diabetes, HIV infection, metastatic cancer, and lysosomal storage disease (Asano, 2003; Mehta *et al.*, 1998). Particularly, the enzyme α-glucosidase is responsible for the hydrolysis of oligosaccharides into monomers and for the hydrolysis of heterosides of α-glucose, providing an aglycone and a glucose unit. The inhibitors of this enzyme are of particular interest in type 2 diabetes, because they are able to slow down the release of glucose from oligosaccharides, lowering postprandial levels of glucose in diabetic patients (Lebovitz, 1998). The search for α-glucosidase inhibitors from nature led to the discovery of miglitol and acarbose, currently used in therapeutics (Grabley and Thiericke, 1999). α-Glucosidase is also involved in the glucosylation process of viral membrane proteins responsible for cell adhesion, contributing to viral infection. Thus, the investigation of α-glucosidase inhibitors is considered important in the search for antiviral compounds (Chen *et al.*, 2004;

Sanchez-Medina *et al.*, 2001). One such inhibitor is the polyhydroxyalkaloid castanospermine, which inhibits the growth of HIV *in vitro* (Walker *et al.*, 1987).

An enzymatic assay, based on the detection of *p*-nitrophenol (400 nm) released by hydrolysis of *p*-nitrophenyl-α-D-glycopyranoside, has been used to identify α-glucosidase inhibitors in solution (Ali *et al.*, 2002). A TLC bioautographic assay has been adapted for the screening of plant extracts using 2-naphthyl-α-D-glucopyranoside as a substrate, followed by the reaction of the released naphthol with the Fast Blue B salt to provide a purple background (Simões-Pires *et al.*, 2009a). In this method, α-D-glucosidase at $10\,U\,mL^{-1}$ is sprayed onto a developed TLC plate, which is then incubated at room temperature for 1 h. For subsequent detection of the active enzyme, a mixture of equal volumes of ethanolic 2-naphthyl-α-D-glucopyranoside $(2\,mg\,mL^{-1})$ and aqueous Fast Blue B Salt $(2.5\,mg\,mL^{-1})$ is sprayed over the plate to give a purple background after 2–5 min (Figure 10). Inhibitors are identified as white spots. This method was also extrapolated to β-glucosidase using the respective enzyme and substrate, with a 20-min incubation at 37 °C and a substrate/Fast Blue B salt ratio of 1 : 4. Another assay for β-glucosidase has been previously developed, based on the conversion of esculin into esculetin, which then reacts with $FeCl_3$ to provide a brown complex (Pandey *et al.*, 2013; Salazar and Furlan, 2007). This assay was notably applied to chemically engineered plant extracts and led to the identification of a semisynthetic β-glucosidase inhibitor (Ramallo *et al.*, 2012). However, the method uses a natural product as an enzymatic substrate, which might be
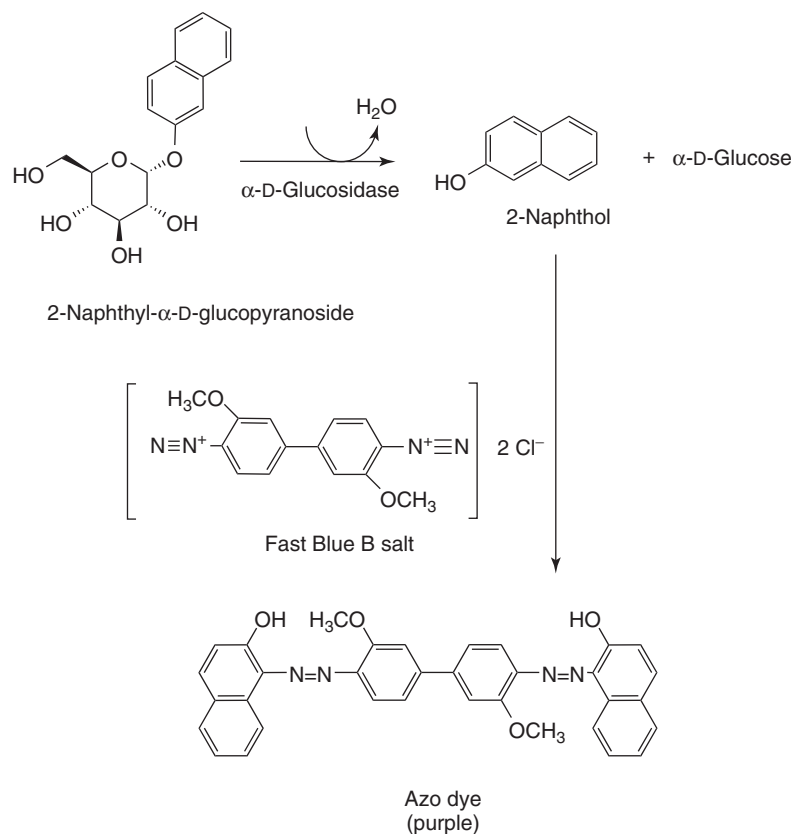
Figure 10   Reaction of α-D-glucosidase with 2-naphthyl-α-D-glucopyranoside to give 2-naphthol and subsequent diazotization.

difficult for the detection of inhibitors from plant extracts containing compounds of the same type as esculetin (coumarins).

4.5.4   Tyrosinase Inhibition

Tyrosinase (EC 1.14.18.1) is the rate-limiting enzyme that controls the production of melanin in animals from tyrosine. It is also responsible for the browning observed when fungi or plant tissues are injured. This oxidase catalyzes the hydroxylation of monophenols to *o*-diphenols and their subsequent oxidation to *o*-quinones, which will subsequently undergo several reactions. Tyrosinase inhibitors are widely used as whitening agents in cosmetics and can be used as antibrowning agents in fruit products. A large number of tyrosinase inhibitors have already been isolated from natural products and the search for new compounds is topical.

A large variety of *in vitro* assays have been developed, most of them based on spectrophotometric methods. Vanni and coworkers developed one of the best-performing reaction mixtures, where tyrosinase, L-DOPA, and the inhibitor are combined in solution. The formation of dopachrome is then measured by spectrophotometry in 96-well plates, and the tyrosinase inhibition rate can be calculated over a blank solution (Vanni *et al.*, 1990). However, this assay is not very convenient for the preliminary screening of crude extracts. Thus, a TLC bioautographic version has been proposed, in which a tyrosinase solution is sprayed onto the developed TLC plate and is immediately followed by a spray of L-tyrosine solution (Wangthong *et al.*, 2007). After 10 min, the background turns into a stable brownish-purple color because of the oxidation of L-tyrosine, and tyrosinase inhibitors are visualized as white areas.

This assay was validated with different known inhibitors such as kojic acid. The concentration of

both enzyme and substrate solutions was expressed as a suitable coverage on the plate (ca. $4\,U\,cm^{-2}$ and $10^{-5}\,mmol\,cm^{-2}$, respectively), but the appropriate quantity to use is not convenient in routine lab procedures. Moreover, the contours of the inhibition areas are not clearly distinguishable, even if activity can be detected into the nanogram range. To improve the contrast between background and active spots, some modifications can be applied, such as spraying the substrate (L-tyrosine) before the enzyme solution. This was performed, for example, to study the activity of *Sideroxylon inerme* L., a South-African tree whose stem bark is widely used as a skin lightener. The tyrosinase TLC bioassay was used for the preliminary validation of activity and for the bioguided isolation that resulted in the identification of epigallocatechin gallate and procyanidin B1 as tyrosinase inhibitors (Momtaz *et al.*, 2008). A better contrast can also be obtained with slight changes in the substrate. Using the TLC bioassay with a mixture of L-DOPA and L-tyrosine as a substrate, several tyrosinase inhibitors could be clearly detected from some other African plants traditionally used for skin lightening (Kamagaju *et al.*, 2013). Interestingly, some compounds led to an activation of tyrosinase, noticeably as dark-brown spots on the plate, and were further confirmed by an increased melanogenesis in melanoma human cell lines (Figure 11). This autographic assay can, thus, detect both tyrosinase inhibitors and activators.

The assay was also applied to sandalwood oil, indicating that α-santalol inhibits tyrosinase activity. This inhibition was subsequently confirmed by conventional spectrophotometric microplate measurements (Misra and Dey, 2013).

To date, the bioautographic assays for tyrosinase inhibition are only based on the mushroom enzyme, which contains an extra allosteric site compared to human tyrosinase. As this secondary site might give false-positive results (Dubois *et al.*, 2012), this bioautographic technique should rather be used as a rapid screening procedure to detect the presence of tyrosinase inhibitors in crude extracts. Owing to the intrinsically low enzymatic rate of human tyrosinase, a visible background would hardly be obtained using TLC.

4.5.5 Xanthine Oxidase Inhibition

The enzyme xanthine oxidase (XO, EC 1.17.3.2) catalyzes the oxidation of hypoxanthine to xanthine and subsequently to uric acid, while generating ROS. Therefore, the inhibition of xanthine oxidase decreases oxidative stress involved in the development of several pathologies such as cancer, inflammation, and aging. It also reduces hyperuricemia and related disorders such as gout and kidney stones.

In the bioautographic assay developed for the detection of XO inhibitors, an agar solution
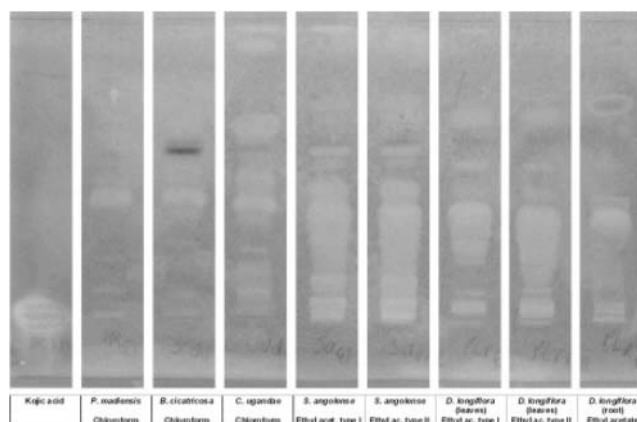
Figure 11 Inhibition of mushroom tyrosinase on a TLC-chromatoplate at $1\,mg\,mL^{-1}$. Visualization of tyrosinase inhibition by all chloroform and ethyl acetate extracts. Whitish spots indicate areas of mushroom tyrosinase inhibition with kojic acid as a positive control. Dark spots indicate tyrosinase activation. (*Source*: Reprinted from J. Ethnopharmacol., 146, Kamagaju L., Morandini R., Bizuru E., Nyetera P., Nduwayezu J.B., Stévigny C., Ghanem G., Duez P., Tyrosinase modulation by five Rwandese herbal medicines traditionally used for skin treatment, 824–834, Copyright 2013, with permission from Elsevier.)

containing XO ($68\,mU\,mL^{-1}$) and nitroblue tetrazolium chloride (NBT) is layered onto the plate, which is dipped into a xanthine solution following solidification. The reaction with the active enzyme leads to the production of superoxide radicals that reduce the pale yellow tetrazolium salt to purple formazan (Figure 12) (Ramallo *et al.*, 2006). Inhibitors such as allopurinol, for which the detection limit is 5 ng, subsequently appear as pale spots. It should be noted that radical scavengers that do not act directly on XO activity also give positive results in this bioassay, as the chromogenic response involves reaction with free radicals. Therefore, this assay requires a further test to discriminate genuine XO inhibitors from superoxide scavengers. This can be performed using a mixture of riboflavine and NBT dissolved in agar, which is subsequently layered onto the TLC plate: after solidification and exposure for a few minutes to daylight, pure XO inhibitors do not produce any change, whereas radical scavengers appear as clear spots against the dark purple background. As light can influence the activity of xanthine oxidase, these assays (solidification of the agar solution and subsequent reactions) must be performed in the dark.

4.6 Yeast Estrogen Screen (YES) Bioautographic Test

To detect estrogenic compounds, a bioassay was developed with genetically engineered yeasts as indicator organisms. The genome of *Saccharomyces cerevisiae* was transformed with a plasmid encoding the human estrogen receptor, along with estrogen-responsive sequences controlling the expression of a reporter gene encoding β-galactosidase. In the presence of estrogen-like compounds, β-galactosidase is secreted and subsequently reacts with chlorophenol red β-D-galactopyranoside, which turns from yellow to red in the presence of active substances (Routledge and Sumpter, 1996). The transposition of the original microplate assay to TLC was facilitated by the fact that *S. cerevisiae* is easy to grow on TLC layers. Furthermore, the sensitivity of the bioassay was improved by the use of the coumarin derivative, 4-methylumbelliferyl-β-D-galactopyranoside, instead of chlorophenol red β-D-galactopyranoside, leading to the fluorescent 4-methylumbelliferone (Figure 13) (Coldham *et al.*, 1997; Müller *et al.*, 2004). Indeed, the contrast between the yellow background and red active spots is not sufficient to clearly

Figure 12 Enzymatic oxidation of xanthine to uric acid and subsequent formation of formazan.
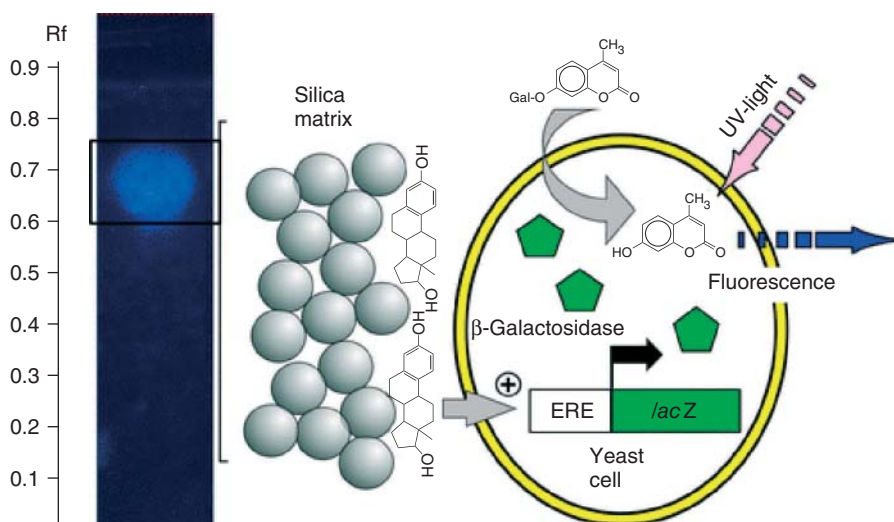
Figure 13 Principle of the TLC bioautographic assay using recombinant *Saccharomyces cerevisiae* and 4-methylumbelliferyl-β-ᴅ-galactopyranoside for the detection of estrogenic compounds. (*Source*: Reprinted with permission from Buchinger S, Spira D, Broeder K, Schluesener M, Ternes T, Reifferscheid G. 2013. Direct coupling of thin-layer chromatography with a bioassay for the detection of estrogenic compounds: applications for effect-directed analysis. Anal. Chem. 85, 7248–7256. Copyright 2013 American Chemical Society.)

detect weak levels of estrogen-like substances, either visually or with a TLC scanner for densitometric assessments. In practice, the developed TLC plates are dipped into a suspension of recombinant yeasts and incubated at 32 °C for 26 h in a sterilized tank containing humid paper. A $0.1 \, g \, L^{-1}$ solution of 4-methylumbelliferyl-β-ᴅ-galactopyranoside is then sprayed onto the layer, which is incubated for a further 3 h. To increase the fluorescence of 4-methylumbelliferone, the TLC plates are exposed to ammonia vapor for 5 min. In these conditions, the minimum amount of 17β-estradiol detectable at 366 nm on the layer (without chromatographic separation) is 1.3 pg and the limit of quantification is less than 3 pg (Buchinger *et al.*, 2013). This sensitive bioassay might be a valuable tool for the discovery of natural products that possess estrogen-like activities, but surprisingly, there are no reports that cite the application of this TLC bioautographic assay in the field of plant analysis. Even using the microplate version, only two studies have been published concerning estrogenic properties of natural products, notably some dietary flavonoids (Breinholt and Larsen, 1998; Zhang *et al.*, 2006).

4.7 Other TLC Bioautographic Assays

4.7.1 DNA-Binding

Despite progress in early diagnosis and new therapeutic treatments, cancer remains the leading cause of death in developed countries. Currently, most of the anticancer drugs approved for chemotherapy are small DNA-binding molecules. Targeting DNA as a cancer therapy was shown to be successful, and thus, the search for such antitumor drugs is still relevant. Cell-based assays are the main strategy to identify DNA-binding compounds. However, a screening approach based on TLC was developed to analyze the capacity of secondary metabolites to bind to DNA.

This test is based on the difference in migration distance on the TLC plate, depending on whether substances are linked to DNA. To observe the differences in *Rf* values, single compounds are spotted twice, either as a pure sample or together with DNA before chromatography. The affinity for DNA is then expressed by the Rf_2/Rf_1 ratio, where Rf_1 represents the *Rf* value of the pure compound and Rf_2 represents the *Rf* value of the substance cospotted with DNA (Maier *et al.*, 1999). The assay, validated with known intercalating drugs such as

doxorubicin, is run on RP-18 silica gel plates eluted with 4 : 1 (v/v) methanol:1 M aqueous ammonium acetate solution to maintain a physiological pH value. To observe a complete interaction, the appropriate amounts to spot are 5 μg for secondary metabolites and 4 μg for DNA. Detection can be performed by classical means, such as UV extinction at 254 nm or following derivatization with staining agents. For complex mixtures such as crude extracts, the TLC bioautographic assay must be adjusted by the use of two-dimensional thin-layer chromatography (2D-TLC). In the first dimension, samples are spotted without DNA and eluted with 1 : 3 (v/v) methanol:0.5 M aqueous ammonium acetate solution. Following this initial separation, the DNA solution is spotted as a thin straight line just above the separated extract and the second separation step is run with 4 : 1 (v/v) methanol:0.5 M aqueous ammonium acetate solution. The application of this bioautographic assay to some microbial extracts enabled the characterization of several compounds as new potential DNA-binding agents (Maul *et al.*, 1999).

The main limitations of this bioassay are the poor resolution due to the use of reverse-phase TLC plates, and also the 2D-TLC procedure, which prevents the simultaneous screening of numerous extracts.

4.7.2 Hemolytic Activity

Hemolytic activity, as well as the inhibition of seed germination, is usually carried out in petri dishes containing agar gel, and only crude extracts or fractions are tested this way. This procedure generates no information about the compounds responsible for the observed activity in contrast to a TLC bioautographic assay. This procedure was applied to a specific *Bacillus subtilis* strain to detect potential hemolytic activity. Following elution of the microbial extract, the silica gel plate was air-dried and subsequently placed face down onto the surface of Columbia blood agar. After 2 h of incubation at room temperature, some hemolytic halos were clearly observed, enabling the localization of active substances on the TLC plate (Hofemeister *et al.*, 2004).

4.8 Pitfalls of TLC Bioautographic Assays

To obtain a clear fingerprint of biological activity, the development of the TLC plate is of the utmost importance, together with the quality of the subsequent bioassay. The application of microbial suspensions, enzymes, chemicals, or agar solutions is a key step in the bioautographic assay, which can be performed by spraying or dipping the layer. The latter is the best method to obtain a uniform impregnation, which is essential for semiquantitative measurements. This requires the immersion of the developed TLC plate in a tank containing the solution to apply, either manually or using an automatic immersion device. The solvents used to prepare the solution must be carefully selected to not dissolve the compounds from the layer itself, even if the immersion time is relatively short (from a few seconds to 1 min). The main drawback of this method is the large volume of solution usually required to immerse the TLC plate (200 mL for an automatic immersion device), which is clearly not possible with expensive reagents such as enzymes. In these cases, spraying can be a suitable alternative, if the solution is applied with a uniform movement (up-and-down and side-to-side) and at an appropriate distance (15–20 cm) until the layer is entirely covered. Excessive saturation should be avoided as it leads to drips that distort the response areas. To prevent polar compounds from dissolving during dipping or spraying, another application technique has been recently developed. This method, used for the bioluminescent assay with *Vibrio fischeri*, applies a rolling device, similar to a rolling pin, impregnated with the test solution, over the chromatographic plate (Baumgartner *et al.*, 2011). In this way, zones corresponding to polar compounds, which tend to show tailing or blurring after dipping, are better defined. Highly lipophilic compounds might also be problematic when applying aqueous solutions, but for other reasons. In some cases, the question arises whether there is a genuine effect or whether the solution "slides" on the lipophilic spot because of apparent repulsion between water and nonpolar substances, avoiding any contact between these compounds and the biological target. For this reason, it is essential to develop parallel bioassays to identify false positive or negative results.

In addition, even if both development and reagent application are successfully carried out, the obtained results might vary significantly according to the nature of the layer and the reagents and solvents used (Litwinienko and Ingold, 2003; Yrjönen *et al.*, 2003). To investigate the influence of the adsorbent on the bioautographic response in depth, different layers

were tested with the DPPH$^\bullet$ bioautographic assay. Approximately 20 phenolic compounds were applied to each layer as free radical scavenger standards. The TLC plates were dipped immediately into a methanolic DPPH$^\bullet$ solution, without any chromatographic elution, and the plates were photographed every 5 min for 1 h (Cieśla *et al*., 2012). Images were then processed using specific software to extract quantitative data from the discoloration spots. As expected, the global radical scavenging activity increases with the number of phenolic groups. Nevertheless, significant differences according to the nature of the adsorbent are notable. The first observation concerns reaction kinetics, which differ from one coating to another; spots tend to appear immediately after staining on silica plates, whereas it requires several minutes to detect them on CN-silica plates. Furthermore, reaction kinetics also vary according to the compound tested; some require more time to develop a stable yellow spot. However, the most striking difference is the variation in activity according to the nature of the adsorbent. For example, protocatechuic acid (compound **5** in Figure 14) exhibits a strong free-radical scavenging activity on silica gel and diol plates, whereas almost no activity is distinguishable on aluminum oxide, cyano, or reversed-phase silica plates. The overall activity is weaker on these three coatings than on silica gel and diol-silica. These two phases possess functional groups that are able to interact with phenolic and carboxylic groups by the formation of hydrogen bonds, leading to a protonation of phenolic compounds. In this case, hydrogen atom transfer during the DPPH$^\bullet$ neutralization is accelerated, resulting in an enhancement of the free-radical scavenging activity. For the same reason, the nature of the solvent used to prepare the DPPH$^\bullet$ solution greatly influences the response (Musialik *et al*., 2009). Therefore, as both adsorbent and solvent might interfere with the TLC-DPPH$^\bullet$ assay,

to obtain accurate results, it would be preferable to carry out the bioautographic assay on a nonspecific adsorbent such as RP-18-modified silica and to prepare the DPPH$^\bullet$ solution with nonprotic solvents.

Finally, the main limitation of TLC bioautographic assays is the restricted number of conceivable targets, as many pharmacological assays are not compatible with TLC conditions. As previously observed, TLC assays require both a genuine reaction and above all, visualization of the activity directly on the plate, which can be problematic. In addition, considering TLC bioautography as a screening procedure to discover hit compounds, the same problems encountered with conventional biological *in vitro* strategies might exist. The risk to miss prodrugs or minor active compounds or the loss of potential synergistic effects also arises with TLC bioassays. Furthermore, in a few cases, TLC bioautographic assays and conventional microtiter plate assays do not lead to the same results. This can be illustrated with acetylcholinesterase inhibitors, comparing Marston's TLC bioassay to the 96-well plate assay based on Ellman's method. Even when 83% of the tested compounds gave similar results with both tests, 15% were found to inhibit AChE in solution but not in TLC and only 2% were active in TLC but inactive with Ellman's microplate assay. Additional experiments tend to explain these results by a probable interaction of the silica layer with the enzyme or the test substances, as previously mentioned, resulting in an altered affinity of acetylcholinesterase for the compounds (Di Giovanni *et al*., 2008). Therefore, it can be useful to validate positive results obtained from TLC with microplate assays, especially to quantify activity via IC$_{50}$, for example, even if densitometric measurements can be performed from highly resolved HPTLC plates.
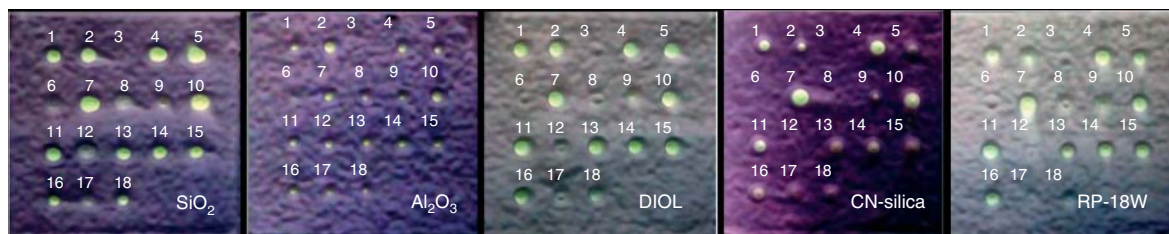
Figure 14   Effect of the type of adsorbent on the DPPH$^\bullet$ scavenging activity of phenolic compounds. (*Source*: Reproduced by kind permission of Łukasz Cieśla.)

5 CONCLUSION

TLC has long been considered as a routine technique for the rapid analysis of plant extracts, as it is cost effective and easy to perform. A few decades ago, TLC was considered outdated, compared to the recent advances in the field of other liquid chromatographic techniques, especially regarding hyphenated systems. Despite these prejudices, many innovations and enhancements have been made within the last few years to improve the TLC performance and to expand its field of application (Poole, 2003). Through the development of advanced automated devices and high resolution stationary phases, TLC is now a multipurpose technique, which can be applied for the separation, quantification, and subsequent structural elucidation of secondary metabolites from crude plant extracts, with minimal sample preparation. Furthermore, the coupling with bioautographic assays provides useful *in situ* information concerning the biological activity of separated compounds.

Modern TLC, therefore, remains the technique of choice in the field of plant analysis, either for the quality control of herbal drugs or in the early stage of drug development, as a powerful screening method and a useful tool throughout all purification processes.

ACKNOWLEDGMENTS

REFERENCES

Adhami, H.-R., Scherer, U., Kaehlig, H., *et al*. (2013) *Phytochem. Anal.*, **24**, 395–400.

Adsersen, A., Gauguin, B., Gudiksen, L., *et al*. (2006) *J. Ethnopharmacol.*, **104**, 418–422.

Ali, M. S., Jahangir, M., Hussan, S. S., *et al*. (2002) *Phytochemistry*, **60**, 295–299.

Apers, S., Naessens, T., Pieters, L., *et al*. (2006) *J. Chromatogr., A*, **1112**, 165–170.

Asano, N. (2003) *Glycobiology*, **13**, 93R–104R.

Baumgartner, V., Hohl, C. and Schwack, W. (2011) *J. Chromatogr., A*, **1218**, 2692–2699.

Beedessee, G., Ramanjooloo, A., Surnam-Boodhun, R., *et al*. (2013) *Chem. Biodiv.*, **10**, 442–451.

Betina, V. (1973) *J. Chromatogr., A*, **78**, 41–51.

Bhandari, P., Kumar, N., Gupta, A. P., *et al*. (2007) *J. Sep. Sci.*, **30**, 2092–2096.

Bhope, S. G., Kuber, V. V., Ghosh, V. K., *et al*. (2011) *J. Liq. Chromatogr. Relat. Technol.*, **34**, 579–590.

Blois, M. S. (1958) *Nature*, **181**, 1199–1200.

Breinholt, V. and Larsen, J. C. (1998) *Chem. Res. Toxicol.*, **11**, 622–629.

Buchinger, S., Spira, D., Broeder, K., *et al*. (2013) *Anal. Chem.*, **85**, 7248–7256.

Chen, H. M., Yan, X. J., Lin, W., *et al*. (2004) *Pharm. Biol.*, **42**, 416–421.

Cieśla, L. (2012) *Med. Chem.*, **8**, 102–111.

Cieśla, L., Kryszen, J., Stochmal, A., *et al*. (2012) *J. Pharm. Biomed. Anal.*, **70**, 126–135.

Coldham, N. G., Dave, M., Sivapathasundaram, S., *et al*. (1997) *Environ. Health Perspect.*, **105**, 734–742.

Dall'Acqua, S., Maggi, F., Minesso, P., *et al*. (2010) *Fitoterapia*, **81**, 1208–1212.

Di Giovanni, S., Borloz, A., Urbain, A., *et al*. (2008) *Eur. J. Pharm. Sci.*, **33**, 109–119.

Dubois, C., Haudecoeur, R., Orio, M., *et al*. (2012) *ChemBioChem*, **13**, 559–565.

Eberz, G., Rast, H. G., Burger, K., *et al*. (1996) *Chromatographia*, **43**, 5–9.

Ellman, G. L., Courtney, K. D., Andres, V. Jr., *et al*. (1961) *Biochem. Pharmacol.*, **7**, 88–95.

Glavind, J. and Holmer, G. (1967) *J. Am. Oil Chem. Soc.*, **44**, 539–542.

Goodall, R. R. and Levi, A. A. (1946) *Nature*, **158**, 675.

Gössi, A., Scherer, U. and Schlotterbeck, G. (2012) *Chimia*, **66**, 347–349.

Grabley, S. and Thiericke, R. (1999) The impact of natural products on drug discovery, in , Drug Discovery from Nature, eds. S. Grabley and R. Thiericke, Springer, Berlin, pp. 3–37.

Haouat, A. C., El Guendouzi, S., Haggoud, A., *et al*. (2013) *J. Med. Plants Res.*, **7**, 1015–1021.

Hassan, A. M. S. (2012) *Phytochem. Anal.*, **23**, 405–407.

Hofemeister, J., Conrad, B., Adler, B., *et al*. (2004) *Mol. Genet. Genomics*, **272**, 363–378.

Hostettmann, K., Borloz, A., Urbain, A., *et al*. (2006) *Curr. Org. Chem.*, **10**, 825–847.

Hostettmann, K. and Marston, A. (1994) *Pure Appl. Chem.*, **66**, 2231–2234.

Houghton, P. J., Agbedahunsi, J. M. and Adegbulugbe, A. (2004) *Phytochemistry*, **65**, 2893–2896.

Huang, D., Ou, B. and Prior, R. L. (2005) *J. Agric. Food Chem.*, **53**, 1841–1856.

Inoue, M., Nishimura, H., Li, H. H., *et al*. (1992) *J. Chem. Ecol.*, **18**, 1833–1840.

Izmailov, N. A. and Schraiber, M. S. (1938) *Farmatsiya*, **3**, 1–7.

Kamagaju, L., Morandini, R., Bizuru, E., *et al*. (2013) *J. Ethnopharmacol.*, **146**, 824–834.

Kiely, J. S., Moos, W. H., Pavia, M. R., *et al*. (1991) *Anal. Biochem.*, **196**, 439–442.

Kim, H., Jee, E., Ahn, K., *et al*. (2010) *Arch. Pharm. Res.*, **33**, 1355–1359.

Klöppel, A., Grasse, W., Brümmer, F., *et al*. (2008) *J. Planar Chromatogr. Mod TLC*, **21**, 431–436.

Klöppel, A., Brümmer, F., Schwabe, D., *et al*. (2013) *J. Mar. Biol.*, **2013**, 1–9.

Lebovitz, H. E. (1998) *Diabetes Rev.*, **6**, 132–145.

Lee, C. and Yoon, J. (2008) *J. Photochem. Photobiol. A*, **197**, 232–238.

Lewin, L. M. and Marcus, N. (1965) *Anal. Biochem.*, **10**, 96–100.

Litwinienko, G. and Ingold, K. U. (2003) *J. Org. Chem.*, **68**, 3433–3438.

Maier, A., Maul, C., Zerlin, M., *et al*. (1999) *J. Antibiot.*, **52**, 945–951.

Marco, G. J. (1968) *J. Am. Oil Chem. Soc.*, **45**, 594–598.

Marston, A., Kissling, J. and Hostettmann, K. (2002) *Phytochem. Anal.*, **13**, 51–54.

Marston, A. (2011) *J. Chromatogr., A*, **1218**, 2676–2683.

Matsuda, T. and Goto, F. (1952) *Repts. Govt. Ind., Res. Inst., Nagoya*, **1**, 80–82.

Maul, C., Sattler, I., Zerlin, M., *et al*. (1999) *J. Antibiot.*, **52**, 1124–1134.

Mehta, A., Zitzmann, N., Rudd, P. M., *et al*. (1998) *FEBS Lett.*, **430**, 17–22.

Mendoza, C. E., Wales, P. J., McLeod, H. A., *et al*. (1968) *Analyst*, **93**, 34–38.

Meyer, J. J. M., Van der Kooy, F. and Joubert, A. (2007) *S. Afr. J. Bot.*, **73**, 654–656.

Miller, N. J., Rice-Evans, C., Davies, M. J., *et al*. (1993) *Clin. Sci.*, **84**, 407–412.

Miller, N. J. and Rice-Evans, C. A. (1997) *Free Radical Res.*, **26**, 195–199.

Misra, B. B. and Dey, S. (2013) *Nat. Prod. Commun.*, **8**, 253–256.

Momtaz, S., Mapunya, B. M., Houghton, P. J., *et al*. (2008) *J. Ethnopharmacol.*, **119**, 507–512.

Morlock, G. and Schwack, W. (2010) *J. Chromatogr., A*, **1217**, 6600–6609.

Mroczek, T. (2009) *J. Chromatogr., A*, **1216**, 2519–2528.

Mroczek, T. and Mazurek, J. (2009) *Anal. Chim. Acta*, **633**, 188–196.

Müller, M. B., Dausend, C., Weins, C., *et al*. (2004) *Chromatographia*, **60**, 207–211.

Musialik, M., Kuzmicz, R., Pawłowski, T. S., *et al*. (2009) *J. Org. Chem.*, **74**, 2699–2709.

Pandey, S., Sree, A., Dash, S., *et al*. (2013) *BMC Microbiol.*, **13**, 55–60.

Payne, G. F., Bringi, V., Prince, C. L., *et al*. (1992) The quest for commercial production of chemicals from plant cell culture, in *Plant Cell and Tissue Culture in Liquid Systems*, eds. G. F. Payne, V. Bringi, L. C. Prince and M. L. Shuler, Carl Hanser Verlag, Munich, pp. 1–10.

Philip, F. (1974). An Investigation of the Antioxidants in a Textured Vegetable Protein Product from Soy Products, Ph.D. Thesis. West Lafayette, IN: Purdue University.

Picken, J. C. Jr., and Bauriedel, W. R. (1950) *Proc. Soc. Exp. Biol. Med.*, **75**, 511–515.

Poole, C. F. (2003) *J. Chromatogr., A*, **1000**, 963–984.

Rahalison, L., Hamburger, M., Hostettmann, K., *et al*. (1991) *Phytochem. Anal.*, **2**, 199–203.

Ramallo, I. A., Sierra, M. G. and Furlan, R. L. E. (2012) *Med. Chem.*, **8**, 112–117.

Ramallo, I. A., Zacchino, S. A. and Furlan, R. L. E. (2006) *Phytochem. Anal.*, **17**, 15–19.

Rates, S. M. (2001) *Toxicon*, **39**, 603–613.

Rhee, I. K., van Rijn, R. M. and Verpoorte, R. (2003) *Phytochem. Anal.*, **14**, 127–131.

Rhee, I. K., van, d M M., Ingkaninan, K., *et al*. (2001) *J. Chromatogr., A*, **915**, 217–223.

Routledge, E. J. and Sumpter, J. P. (1996) *Environ. Toxicol. Chem.*, **15**, 241–248.

Salazar, M. O. and Furlan, R. L. E. (2007) *Phytochem. Anal.*, **18**, 209–212.

Sanchez-Medina, A., Garcia-Sosa, K., May-Pat, F., *et al*. (2001) *Phytomedicine*, **8**, 144–151.

Shai, L. J., McGaw, L. J. and Eloff, J. N. (2009) *S. Afr. J. Bot.*, **75**, 363–366.

Sharma, O. P. and Bhat, T. K. (2009) *Food Chem.*, **113**, 1202–1205.

Sherma, J. (2010) Bioluminescence: detection in TLC, in , Encyclopedia of Chromatography, 3rd edn, ed. J. Cazes, CRC Press, pp. 234–237.

Simões-Pires, C. (2009) Investigation of antiplasmodial compounds from various plant extracts, Ph.D. Thesis. Geneva: University of Geneva, 197 pp.

Simões-Pires, C. A., Hmicha, B., Marston, A., *et al*. (2009a) *Phytochem. Anal.*, **20**, 511–515.

Simões-Pires, C. A., Diop, E. A., Ioset, J. R., *et al*. (2009b) *Planta Med.*, **75**, PD22.

Stahl, E. (1962) Dünnschicht-Chromatographie, in , Modern Methods of Plant Analysis, ed. Springer , Springer, Berlin, pp. 214–229.

Takao, T., Kitatani, F., Watanabe, N., *et al*. (1994) *Biosci. Biotechnol. Biochem.*, **58**, 1780–1783.

Tswett, M. S. (1905) *Proc. Warsaw Soc. Nat.*, **14**, 20–39.

Urbain, A., Marston, A., Queiroz, E. F., *et al*. (2004) *Planta Med.*, **70**, 1011–1014.

Urbain, A., Marston, A. and Hostettmann, K. (2005) *Pharm. Biol.*, **43**, 647–650.

van Asperen, K. (1962) *J. Insect Physiol.*, **8**, 401–416.

Vanni, A., Gastaldi, D. and Giunta, G. (1990) *Ann. Chim. (Rome)*, **80**.

Verbitski, S. M., Gourdin, G. T., Ikenouye, L. M., *et al*. (2008) *J. AOAC Int.*, **91**, 268–275.

Wagman, G. H. and Bailey, J. V. (1969) *J. Chromatogr.*, **41**, 263–264.

Wagner, H. and Bladt, S. (1996) *Plant Drug Analysis: A Thin Layer Chromatography Atlas*, Springer, Berlin.

Walker, B. D., Kowalski, M., Goh, W. C., *et al*. (1987) *Proc. Natl. Acad. Sci.*, **84**, 8120–8124.

Wangthong, S., Tonsiripakdee, I., Monhaphol, T., *et al*. (2007) *Biomed. Chromatogr.*, **21**, 94–100.

Weins, C. and Jork, H. (1996) *J. Chromatogr., A*, **750**, 403–407.

Willcox, M. L., Graz, B., Falquet, J., *et al*. (2007) *Trans. R. Soc. Trop. Med. Hyg.*, **101**, 1190–1198.

Wu, T, Tang, J and Tao, H. (2013) Screening method of lipase inhibitor using thin layer chromatography-bioautography. China patent CN102980969A.

Yan, Y.-Z., Xie, P.-S., Lam, W.-K., *et al.* (2010) *J. AOAC Int.*, **93**, 1384–1389.

Yang, Z.-D., Song, Z.-W., Ren, J., *et al.* (2011) *Phytochem. Anal.*, **22**, 509–515.

Yang, Z., Zhang, X., Duan, D., *et al.* (2009) *J. Sep. Sci.*, **32**, 3257–3259.

Yrjönen, T., Li, P., Summanen, J., *et al.* (2003) *J. Am. Oil Chem. Soc.*, **80**, 9–14.

Zhang, C. N., Zhang, X. Z., Zhang, Y., *et al.* (2006) *J. Ethnopharmacol.*, **105**, 223–228.

HPLC and Ultra HPLC: Basic Concepts

Veronika R. Meyer

Swiss Federal Laboratories for Materials Science and Technology, Laboratory for Protection and Physiology, St. Gallen, Switzerland

1 INTRODUCTION

1.1 The Invention of Chromatography

Chromatography, one of the most widespread methods in today's analytical science, was invented and developed in plant analysis science. At the beginning of the twentieth century, the chemical investigation of plants usually started with liquid–liquid extraction of the samples. However, Michael (or Mikhail) Semenovich Tswett, a young plant physiologist of Russian origin and by then living in Geneva (Switzerland), was not satisfied with the results and started partition experiments with a broad range of fine-grained adsorbents. He packed these materials into glass tubes with a stopcock at the lower end, wetted them with an organic solvent, applied the raw plant extract solution, and "developed" it with the continuous addition of solvent. He termed the packed tube a "column" and the obtained pattern of plant pigments a "chromatogram." After numerous trials, he found a promising separation system for the pigments of green leaves: a calcium carbonate column developed with carbon disulfide or benzene. Tswett proved that chlorophyll consists of two components (nowadays chlorophylls a and b, by then α and β) with different visible spectra. In the same run, the xanthophylls separated into three or more yellow bands (Tswett, 1906, 1990). Richard Willstätter, the plant pigment authority of these decades, claimed that Tswett's new method would be an *odd way*

and stuck to the conventional liquid–liquid extraction (Willstätter and Stoll, 1913).

Soon after Twett's publications, an important discovery in plant science was made by Leroy Sheldon Palmer. In his PhD thesis at the University of Missouri (USA), he found using the new chromatographic method that the carotenes in cow's milk come from the plants a cow is eating and that they are not synthesized by the animal (Ettre and Wixom, 1993).

1.2 Chromatography Today

Tswett had developed what is called today "column liquid chromatography". Its feature is the partition of the analytes between a packed stationary phase and a liquid mobile (moving) phase. Nowadays, chromatography is used in a variety of designs. The stationary phase can be a particulate matter of $1–100\,\mu m$ size within a column or a thin film at the interior wall of a capillary. Particular stationary phases can also be fixed on a plate, which results in thin-layer chromatography (TLC). The mobile phase can be a liquid, a gas [this approach is called *gas chromatography* (GC), usually in open capillaries], or a supercritical fluid [supercritical fluid chromatography (SFC)].

This article describes liquid chromatography in columns, packed with particles of $1–10\,\mu m$ in diameter. Owing to the fine packing, high pressure is needed to force the mobile phase through the column. Therefore, the method was first known as

high-pressure liquid chromatography (HPLC). Later, when the problem of forcing a constant flow by a high-pressure pump through the column was more or less solved, the abbreviation was interpreted as high-performance liquid chromatography (HPLC). In the decade between 2000 and 2010, its pressure range could be expanded to 600–1000 bar. This new variety is known as *ultra high-performance liquid chromatography* (UHPLC; see Section 8).

In plant research, HPLC in all its varieties is one of its most important analytical tools (Wolfender *et al*., 2011).

2　THEORY OF LIQUID CHROMATOGRAPHY

2.1　The Liquid Chromatographic Process

Chromatography takes place if the different analytes that are present in an extract or any other mixture have different partition behavior between the mobile and the stationary phases. They all must be soluble in the mobile phase (or eluent). However, they must have different partition coefficients between these two phases; otherwise, a separation is not possible. The sample solution is injected into the flowing stream of the liquid mobile phase; thus, it is transported to the column packing (or chromatographic bed). As soon as the analyte molecules are in contact with both phases, they start partitioning according to their physicochemical properties (dipole moment, possible charges at their surface, shape, size, etc.). Analytes with a high solubility in the eluent will wander fast through the column and will soon be eluted from its outlet. Those analytes with high "adsorptivity" (in a broad sense) to the stationary phase will stay longer within the column and will be eluted late.

The chromatographic quantity for the description of the partition between the phases is the retention factor k:

$$k = \frac{n_{\text{stat}}}{n_{\text{mob}}} \qquad (1)$$

where n is the number of moles of an analyte in the stationary and the mobile phases, respectively, under conditions of equilibrium. It is linked to the distribution coefficient K that relies on concentrations instead of moles, but which is not of importance here. (Formerly, k was known as the capacity factor k' but

this nomenclature is out of date.) In many cases, it is theoretically possible to predict the approximate value of k in a given phase system from the molecular structure of the analyte, but this approach cannot be used for the investigation of unknown and/or complex mixtures. It is up to the skill and experience of the analyst to find promising separation systems for the problem at issue within short time.

2.2　The Chromatogram

In an ideal case, the different analytes or chemical compounds of the mixture under investigation are eluted one after the other from the HPLC column. Their concentration profiles are theoretically of Gauss shape, although in reality they often show more or less asymmetry. The concentration profiles need to be registered by a detector: first, because most analytes are not colored and cannot be observed by eye, and second because their concentration is very low in a typical HPLC experiment. The detector (e.g., with UV lamp and UV-sensitive diodes for compounds absorbing in the UV) transforms the registered physical effect (e.g., the absorbance) into a voltage, which then will be displayed on a computer screen (could also be a recorder). During the chromatographic run, the observer discovers the evolving pattern of Gauss-shape signals, the so-called peaks, on the screen. The peak pattern is the chromatogram. Figure 1 shows the simplest possible chromatogram with only two peaks and its numbers of merit.

The mobile phase does not interact with the stationary phase and needs the minimum time for its flow through the column, known as *breakthrough time* (sometimes also as dead time) t_0. This time mark is sometimes, but not necessarily, visible in the chromatogram as a small deviation of the baseline. The individual peaks elute at their retention times, here at t_{R1} and t_{R2}. Retention times have a limited informative value because they depend on the flow rate of the eluent. A more scientific number is the retention factor k already defined by Equation (1), because it describes not only the partition equilibrium but also the normalized retention time of an analyte in a given separation system:

$$k = \frac{t_R - t_0}{t_0} \qquad (2)$$
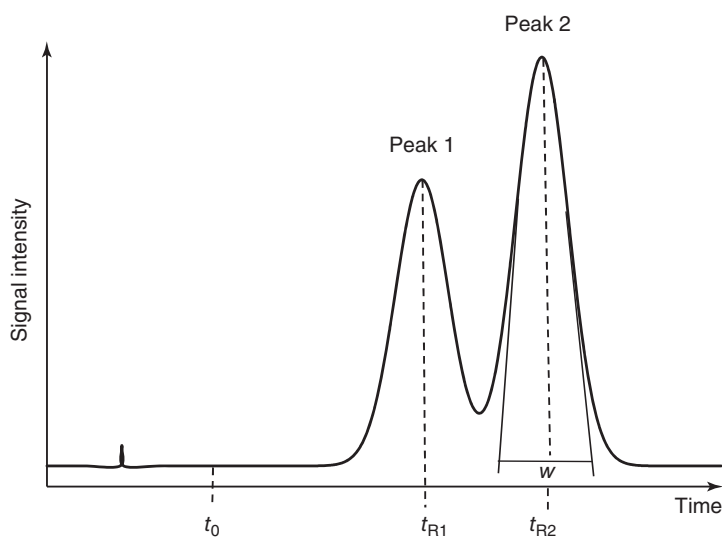
Figure 1 The simplest possible chromatogram with a small signal at the breakthrough time t_0 and two incompletely resolved peaks with their retention times t_R. In the second peak, the inflection tangents are drawn, which define its width w at the position of the baseline.

The ratio of two retention factors is known as the *separation factor* α (formerly as relative retention α):

$$\alpha = \frac{k_2}{k_1} \quad \text{(always } \alpha \geq 1.0\text{)} \qquad (3)$$

Peaks should be as narrow as possible, but the mass transfer and diffusion processes within the column lead to their broadening. The peak width w is given by the intersection of the inflection tangents (the straight lines that touch the peak curve at its sidewise inflection points) with the baseline (w can also be calculated from the peak width at half height, $w = 1.7$ $w_{1/2}$). The resolution, R, of two peaks is defined by:

$$R = 2\frac{t_{R2} - t_{R1}}{w_1 + w_2} \qquad (4)$$

It is important to note that the two peaks with resolution $R = 1.0$ still have some overlap because a Gauss function is not triangular. For baseline resolution of two peaks of equal size, R must be at least 1.5; this number must be higher for peak pairs of unequal size. In many cases, incompletely resolved peaks cannot be quantified accurately (Meyer, 1995).

The retention time and width of a peak allow the calculation of the number of theoretical plates, N, of a column:

$$N = 16\left(\frac{t_R}{w}\right)^2 = 5.54\left(\frac{t_R}{w_{1/2}}\right)^2 \qquad (5)$$

The theoretical plate concept is indeed a theoretical, but definitely a helpful one. A theoretical plate can be imagined as the space within the column, which is needed for the establishment of a complete equilibrium of the analytes between the mobile and the stationary phases. In a well-packed column and a separation system with fast mass transfer, a theoretical plate has a height H, which corresponds to about 2–3 particle diameters of the stationary phase. Consequently, an excellent column of 10 cm length, packed with a 4-μm phase, can yield more than 10,000 theoretical plates. Note that Equation (5) is only valid for isocratic elution but not for gradient separations (see Section 7.1).

The theoretical plate height is simply calculated from N and the column length L_c:

$$H = \frac{L_c}{N} \qquad (6)$$

and is usually given in micrometer. The smaller the H is, the more efficient is the column.

2.3 How to Obtain Chromatographic Resolution

Equation (4) is the geometrical description of the resolution of two peaks. The true value of R comes from the fact that it can also be described by the interplay of the separation factor α, the theoretical plate number N, and the retention factor k:

$$R = \frac{1}{4}(\alpha - 1)\sqrt{N}\frac{k}{1+k} \qquad (7)$$

These three influence parameters show how the resolution of two peaks (and of a whole chromatogram) can be influenced. The separation factor is the most powerful point of action: If α can be increased from, for example, 1.05 to 1.10, the resolution R is doubled. Unfortunately, it can be difficult to influence α in the desired way because good knowledge of the physicochemical properties of both phases and analytes is needed (or good luck).

The theoretical plate number can be increased using a longer column or one with a finer packing. The resulting effect, although easy to understand, is limited because it is only the square root of N that influences the resolution.

The retention factor (whereby it is necessary to distinguish among k_1, k_2, or the mean k, strictly speaking) is only a powerful point of action if it is small. Increasing k from 1.0 to 3.0 (i.e., doubling the retention time from $2t_0$ to $4t_0$) increases R by 50%. However, when the initial k is 10 and then increased to 20, the effect on R is very limited. The simplest possibility to influence k is by changing the composition of the mobile phase, and the correlations are easy to understand.

2.4 The van Deemter Curve

In HPLC, the analyte mixture is usually injected into the eluent flow as a very small volume of typically 10 μL. During the travel through the chromatographic bed, the individual analytes are (hopefully) separated but their bands undergo peak broadening. The eluted peaks are dissolved in a larger volume than the initial one was; this dilution is a drawback of chromatography and should be kept as low as possible. The dilution effect has three reasons:

- Eddy diffusion and flow distribution of the eluent within the column packing – the A term of the van Deemter curve.
- Self-diffusion, thus dilution of the analytes in the surrounding eluent – the B term.
- Mass transfer of the analytes from the mobile to the stationary phase and back to the mobile phase, whereas the latter has moved some small distance toward the column outlet – the C term.

These three terms bring about a nonlinear relationship between the theoretical plate height and the flow rate of the eluent, known as the *van Deemter curve* (van Deemter, Zuiderweg, and Klinkenberg, 1956) shown in Figure 2. There is an optimum flow rate that yields the minimum plate height, thus the maximum number of theoretical plates. It is obvious that too slow an eluent flow rate is detrimental because the separation takes longer than necessary and band broadening (loss of resolution) is striking. Unfortunately, the optimum is at rather low flow rates. Increasing the flow reduces not only the analysis time but also the resolution. The effect on band broadening by higher flow rates is dominated by the mass transfer (the C term). Modern stationary phases should be (and often are) designed in such a way that the mass transfer is fast, which results in a low increase of H in the right part of the van Deemter curve. In this
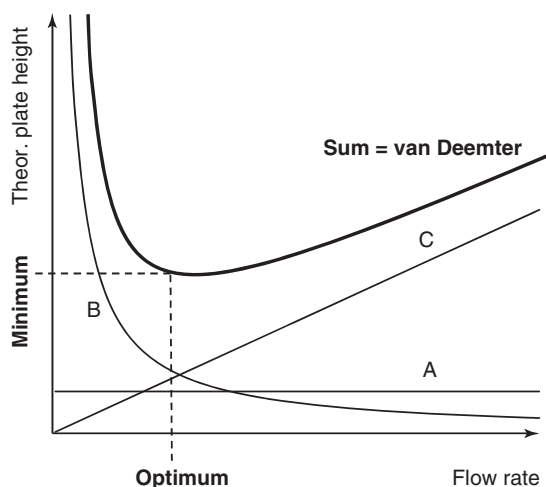
Figure 2 The van Deemter curve with its flow rate optimum. A Influence of eddy diffusion and flow distribution. B Influence of self-diffusion of the analytes in the mobile phase. C Mass transfer of the analytes between mobile and stationary phases.

case, fast separations are possible with only a small detrimental effect on the separation.

2.5 Peak Capacity and Statistical Resolution Probability

Analysts are not really interested in theoretical peak numbers but in peak capacity n. This is the number of peaks that would consecutively fit into a chromatographic run with mutual resolution 1.0 (which is not baseline separation). For isocratic separations (see Section 7.1), n is given by:

$$n_{\text{isocratic}} = 1 + \frac{\sqrt{N}}{4} \ln(1 + k_{\text{max}}) \qquad (8)$$

The value of k_{max} is the free decision of the analyst, that is, how long one accepts to wait for the completion of the chromatographic run. In practical isocratic HPLC, k_{max} will rarely be higher than 20 (the 21-fold time of t_0) because the peaks are the broader and the less high the later they are eluted. The alternative to isocratic elution is gradient separation (see Section 7.1) with changing composition of the mobile phase. In a first approximation, gradient-eluted peaks are of constant width w (linked to the theoretical plate number of the column and to the gradient steepness), and n is much higher than predicted by Equation (8); it is simply a geometric question how many peaks will fit into the gradient run time t_{Rmax}:

$$n_{\text{gradient}} = \frac{t_{\text{Rmax}} - t_0}{w} \qquad (9)$$

Peak capacity is an important number because it must be high for the separation of complex mixtures. In a given separation system, the analytes will be eluted with a retention factor, which is dictated by their physicochemical properties. For the analyst, however, a complex mixture will yield a complex chromatogram that can only be predicted theoretically to a limited degree (e.g., polar compounds will be earlier eluted than nonpolar ones on many stationary phases). In reality, the peaks will appear as a random pattern, and many will overlap, which is detrimental for their detection and quantitation. The statistical probability P that a certain analyte is resolved in the chromatogram with $R \geq 1.0$ is (Davis and Giddings, 1983):

$$P \approx e^{-2m/n} \qquad (10)$$

with $n =$ peak capacity and $m =$ number of analytes (peaks) in the mixture. It is obvious that the higher n is, the larger is P. Note that in an unknown mixture, m is unknown as a matter of course. The probability P' that all analytes of a mixture will be resolved with $R \geq 1.0$ is (El Fallah and Martin, 1991):

$$P' = \left(1 - \frac{m-1}{n-1}\right)^{m-2} \qquad (11)$$

P' is a frustratingly low number if m is 10 or higher and n is less than 100.

2.6 The Limits of Column Liquid Chromatography

If high peak capacity is needed, the first thing the analyst needs is a column with a high number of theoretical plates N. Unfortunately, successful separations of complex mixtures can only be obtained in return for long analysis times and/or high pressure. N increases with column length and with decreasing particle diameter of the stationary phase. Both approaches (and their combination, of course) require high pressure between 100 and 1000 bar. A low flow rate would reduce the pressure, but the van Deemter curve (Figure 2) makes clear that the flow must not be smaller than the curve's optimum. Working at the van Deemter optimum is always best with regard to pressure and also to time and solvent consumption if both column length and particle diameter are optimized for the separation problem (Meyer, 2010a) although it is usually not made.

The fundamental limitation of all chromatographic techniques comes from the diffusion and mass transfer processes of the analytes in the mobile and stationary phases. Diffusion needs time; therefore, also chromatographic separations need it; the x-axis of chromatograms is time. Many separations benefit from higher temperature because diffusion is faster and the eluent is less viscous; both effects allow shorter analysis times with a given pressure (Plumb *et al.*, 2007). Peak broadening is less pronounced at elevated temperature. Unfortunately, not all analytes will tolerate higher temperature, the handling of the eluent can be more demanding, and some peak pairs (especially in enantioselective chromatography) may show lower separation factors.
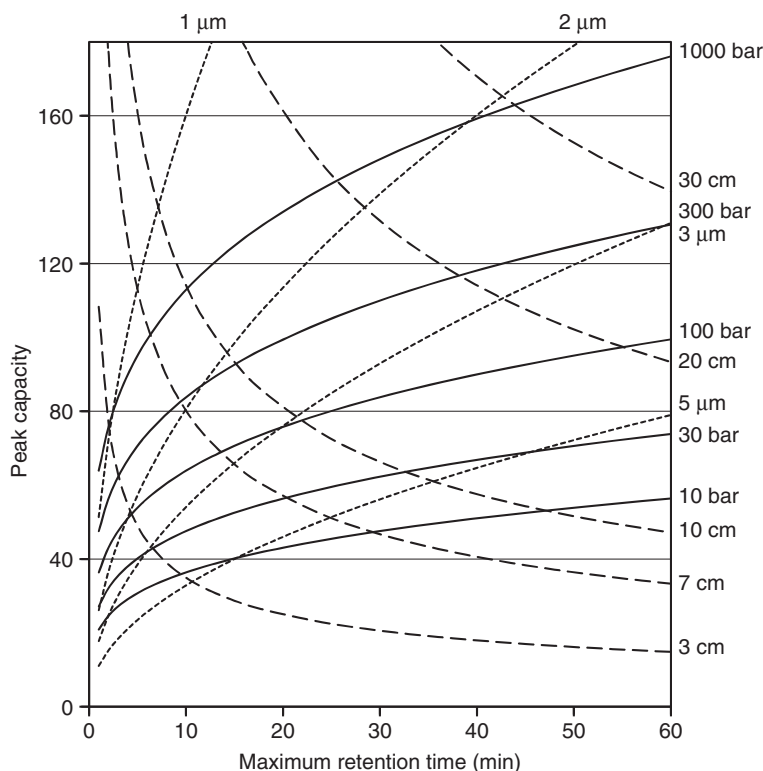
Figure 3   Peak capacity nomogram for isocratic separations of small molecules with an eluent of viscosity 0.9 mPa s, performed at the van Deemter optimum. Solid lines: pressure; dashed lines: column length; dotted lines: particle diameter of the stationary phase.

The practical limitation of HPLC comes from the available pressure delivered by the eluent pump. For decades, the available pumps could operate at a maximum pressure of about 350 bar. Now instruments with an upper limit of 600 or even 1000 bar are available on the market, which expand the possible peak capacity range, unfortunately not by a factor of 2 or 3 but to a lower degree only.

The dependence of peak capacity from analysis time, particle diameter, column length, and pressure is shown in Figure 3. It is valid for typical HPLC columns run with an isocratic mobile phase of viscosity 0.9 mPa s (e.g., water at 25 °C) and for analytes of low molecular mass (not for proteins or polymers). The full background of such nomograms can be found elsewhere (Meyer, 2008). For lower viscosity eluents, as used in adsorption chromatography, and for gradient separations, the situation is more beneficial, that is, the possible peak capacity with a given pressure is higher.

The nomogram of Figure 3 represents HPLC columns that are operated at their van Deemter optimum. Faster chromatography, above the optimum, is possible at the expense of higher pressure and a more or less decrease of the theoretical plate number and thus of the peak capacity, depending on the C term of the phase system. In any case, the van Deemter optimum allows the best use of the available pressure (Meyer, 2010a). Within the full area of the nomogram, the isocratic separation ends at $k_{max} = 20$. The particle size is represented with dotted lines, the column length with dashed lines, and the pressure with solid lines. The nomogram is read as follows:

- If the maximum retention time is around 10 min, one will obtain a modest peak capacity of about 30 with a 3-cm column packed with a 5-μm material and operated at 10 bar.
- A peak capacity of 80 within 10 min is obtained with 7 cm/2 μm and less than 300 bar.

- With 1000 bar and 10-min analysis time, the peak capacity is in the range of 110 using a 10-cm/2.5-μm column.
- By expanding the analysis time to 1 h and operating at 1000 bar, the peak capacity is almost 180. A column of around 40 cm length with a 2.2-μm phase is needed.

All these data depend on the specific conditions of the separation, which may result in a higher or lower peak capacity and in a different column length/particle size combination. However, Figure 3 shows that a pressure increase from 300 to 1000 bar brings a rather modest increase of peak capacity, although this higher separation power may be extremely helpful for the resolution of demanding mixtures.

Pressures higher than 1000 bar are theoretically possible but such separations would come along with practical problems such as the compressibility of the eluent or disadvantageous thermic effects (Martin and Guiochon, 2005).

For a more detailed description of HPLC theory, see textbooks (Meyer, 2010b; Snyder, Kirkland, Dolan, 2010).

3 COLUMNS, STATIONARY PHASES, FITTINGS, AND CAPILLARIES

The book by Neue, although dating from 1997, is still the classic of this topic (Neue, 1997).

Figure 4 shows some items discussed in this article.

3.1 Columns for HPLC

The common type of HPLC columns is a stainless steel tube of 2–4 mm inner diameter d_c and 5–30 cm length L_c. For decades, most columns had 4.6 mm inner diameter, which resulted in high solvent consumption. The necessary volumetric flow rate for a certain linear flow rate (or t_0) increases with d_c^2. Therefore, a 4.6-mm column needs twice the mobile phase volume compared to a 3.2-mm column without offering any advantage such as better resolution. Two-millimeter columns are still more eluent saving, which is favorable with regard to the (certainly not falling) prizes of organic solvents and the amount of waste generated by chromatographic analyses.

Figure 4 The heart of any HPLC instrument. From top to down: two columns of different length and with different end fittings, a connecting capillary, another type of end fitting, and a precolumn holder.

One-millimeter columns are demanding concerning the volume of the HPLC instrument flow path. The connecting capillaries between injector and column as well as between column and detector must not add noticeable to peak broadening (see Section 3.4).

So far, packed capillaries with, for example, 100 μm or even less inner diameter did not find widespread use in HPLC. They need special instrumentation and are used if the sample volume is extremely limited (Hernández-Borges *et al.*, 2007).

Besides the solvent economy, small-diameter columns or capillaries yield higher peaks if a constant amount of sample is injected because the analyte band is less diluted within them. The peak maximum concentration c_{max} at the outlet of the column is given by:

$$c_{max} = \frac{c_i \cdot V_i}{V_R} \qquad (12)$$

where c_i is concentration of the analyte in the sample solution that is injected, V_i the injected volume, and V_R the retention volume, that is, the eluent volume needed for the elution of the peak. V_R itself is proportional to the column diameter, thus c_{max} is inversely proportional to the square of d_c:

$$c_{max} \propto \frac{1}{d_c^2} \qquad (13)$$

Therefore, small-diameter columns are especially advantageous in trace analysis or if the amount of sample is limited.

Other materials than steel are sometimes used for HPLC columns: tantalum, glass, Peek (polyether-etherketone), or polyethylene. They do not release transition metal cations into the eluent that may react with critical analytes or may change the conformation of biopolymers. In such cases, the whole flow path of the eluent within the HPLC instruments must be constructed from nonsteel materials.

Precolumns are short columns mounted between injector and column. They can be packed with the identical stationary phase as the one in the column or with a cheaper (coarser, less specific) material. They retain noneluting analytes or particulate matter present in the sample; therefore, they protect the expensive column from fast deterioration. Precolumns are replaced whenever necessary (e.g., after a large number of samples of low purity) or in regular intervals.

3.2   Stationary Phases

The classical HPLC phase is a spherical, fully porous material of 2–10 μm diameter (see Figure 5a). The pore width is 6–200 nm, which results in a specific surface area of several $100 \, m^2 \, g^{-1}$. Most phases are based on silica (SiO_2 covered with Si–OH groups at the surface). In its pure form, it is used as an adsorbent with nonpolar eluents (normal-phase liquid chromatography, see Section 6.1).

However, derivatized silica is of much greater importance. The most widespread HPLC material is the so-called C_{18}– or ODS (octadecyl) silica with a surface of covalently bonded and stable n-octadecyl groups:

$$-SiOH + Cl–Si(CH_3)_2–(CH_2)_{17}–CH_3 \longrightarrow$$
$$-Si–O–Si(CH_3)_2–(CH_2)_{17}–CH_3 + HCl$$
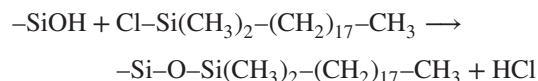
The reaction must be performed under anhydrous conditions. ODS phases are used with aqueous eluents, resulting in the ubiquitous reversed-phase chromatography (RP) (see Section 6.2).

Other functional groups can be bonded to silica with similar chemical reactions: shorter n-alkyl chains such as C_8H_{17}, cyclohexyl, nitrile, nitro, or amino groups, ion exchanger moieties, chiral groups for the separation of enantiomers, and countless other functionalities.

Silica is chemically stable in the pH range from 1 to 8. The same is generally true for bonded phases on silica base; however, using highly pure silica and sophisticated bonding chemistry, some phases can be used up to pH 10 or even higher.

Besides silica, other materials are used for HPLC, although most of them are niche products: pure or derivatized alumina, titania, zirconia, porous graphitic carbon, hydroxylapatite, and organic materials such as polystyrene (also derivatized) or methacrylate gels. Unfortunately, it seems that many analysts are not aware of the potential of this cornucopia filled with interesting phase systems

(a)   (b)
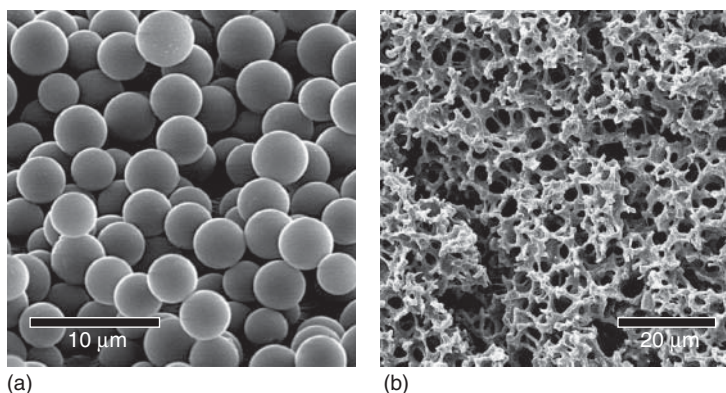
Figure 5   Appearances of stationary phases for HPLC. (a) Spherical silica 3.5 μm. (*Source*: Kromasil, photo supplied by AkzoNobel, used with permission.) (b) Monolithic structure. (*Source*: Photo supplied by Bei Nu, Waters Chemistry Operations Research & Development Analytical Laboratory, used with permission.)

that offer different separation modes, but they try to separate all their mixtures on ODS phases in the reversed-phase mode.

The smaller a totally porous stationary phase particle is, the shorter is the maximum diffusion path for an analyte molecule that enters the particle, undergoes mass transfer to the stationary phase, and finally diffuses out of the particle. Short diffusion pathways mean short residence time of the molecule within the particle. Consequently, faster chromatography is possible with fine-grain packings such as 2 or 3 µm. Even faster separations are obtained if the packing is not fully porous. The so-called core-shell phases have a solid core and a porous outer layer of only 0.25–0.5 µm (Fekete, Oláh, and Fekete, 2012). Core-shell particles are 1.7–2.7 µm in diameter, thus their pressure resistance is high, and the necessary high flow rates of the eluent, dictated by the van Deemter curve, add an extra increase to the pressure demands. If used in a UHPLC system, they offer new dimensions of peak capacity.

The packing of these fine-grain materials into HPLC columns is not trivial. A slurry of the stationary phase in a suitable liquid is packed at high pressure (higher than the later operating pressure) into the tube. Details are a secret of the column producers. Although it is possible that analysts pack their own columns, it is no longer practiced with phases that are available on the market.

An interesting approach for using HPLC at lower pressures is the monolithic stationary phases (see Figure 5b). Instead of a particulate packing, the column contains a single rod of porous material. Although both the solid structure and the channels within it have a diameter in the 1–5 µm range, the permeability of monoliths is astonishingly high, so they can be operated at rather low pressure (Guiochon, 2007). The monoliths can be synthesized separately and then cladded with a pressure-resistant column wall material although this manufacturing process is demanding. However, monomers can also easily be synthesized within the column and most interestingly even within capillaries (Li *et al.*, 2012). Suitable monomers are, for example, alkyl dimethacrylates (such as dodecanediol dimethacrylate) dissolved in methanol–dodecanol mixture that contains an initiator. The solution is placed in UV transparent fused-silica capillaries whose inner surface was treated before in a suitable manner, then the polymerization is performed by irradiation with UV light for some minutes (Liu, Tolley, and Lee, 2012a). Such

capillaries offer high peak capacity at moderate pressure. They must be used with special instrumentation.

3.3 Fittings and Frits

All items in the flow path of the HPLC instrument, from the eluent feed line to the detector outlet, are connected to each other with fittings, most of them are resistant to high pressure. Fittings include screws, nuts, ferrules, column heads (usually constructed with nuts on the side remote from the column), and unions (Batts, 2011). In most cases, they are made from stainless steel. The advantage of plastic fittings is that they can be installed and used without tools – they are "fingertight" – but their pressure resistance is limited and some types cannot be reused often. Fittings must not be overtightened! Unfortunately, there is no standard for fittings and in many cases brand A is not compatible with brand B or C. Intermixing different brands may lead to extra-column volumes or to damage of the threads or of the connected parts.

The large fittings that act as column closures contain a frit that retains the particulate packing (for monolithic columns, no frits are needed). Most frits are porous stainless steel disks, in some cases surrounded by a plastic ring.

3.4 Connecting Capillaries

The various modules of the HPLC instrument and the column are connected with capillaries. Their common outer diameter is 1/16 in. (1.6 mm). The inner diameter must be so small that the capillary volume from the injector outlet to the detector inlet does not add notably to peak broadening. Depending on the column volume (especially its inner diameter) and its performance, capillaries have an inner diameter of 0.25 mm or less. For 3-mm and wider columns, capillaries of 0.25 mm inner diameter can be used, with 2-mm columns, they should not be wider than 0.18 mm and the requirements with 1-mm columns are even stricter. They are made from stainless steel or from polyether-etherketone (Peek). Peek does not release metal cations into the eluent, but it is less pressure resistant than steel and cannot be used with all organic solvents or concentrated acids.

4　INSTRUMENTATION: FROM THE SOLVENT RESERVOIR TO THE COLUMN

4.1　The HPLC Instrument

The basic design of an HPLC instrument is shown in Figure 6. It consists of the eluent reservoir, the high-pressure pump, the injector, the column (the smallest part in Figure 6), the detector, and the data system that is also the control unit for all individual modules. In principle, the modules can be from different manufacturers, but today's HPLC systems are purchased from one single company. The modules are usually piled up to a tower.

The basic design as shown in Figure 6 can be expanded. There can be more than one eluent bottle, which is necessary for gradient elution. In this case, the two or more liquids are mixed before the pump (low pressure gradient) or there are as many pumps as bottles and the mixing occurs behind the pumps (high-pressure gradient). The injector can be a simple hand-operated valve or an autosampler with vial rack, if necessary with a cooling unit for the samples. The column can be placed in the interior of an oven (sometimes also with cooling possibility) to allow the analyses to be performed at elevated and/or constant temperature. The detector of Figure 6 is a UV detector but here a great variety is possible (see Section 5), and more than one detector can be used in series or in parallel.

For troubleshooting HPLC instruments, see the respective textbooks (Dolan and Snyder, 1989; Sadek, 1999).

4.2　Requirements for Mobile Phases

Water and organic solvents used for HPLC must be of high purity, although "high" depends on the intended use and the detection method. HPLC qualities are sold with detailed certificates, and the users should check if the guaranteed data meet their requirements. If a detection method such as refractive index (RI) or conductivity is used, a high UV transparency is not of interest. On the other hand, a special "low UV grade" is needed if the detection must be performed at 210 nm because a typical "UV grade" is perfect for higher wavelengths but maybe not for the lower ones. Special qualities are often needed for mass spectrometric detection.

The same requirements are valid for additives such as buffer salts, acids, bases, and all kinds of reagents.

In general, for gradient elution and for trace analysis, the mobile phase must be purer than for isocratic separations or for analyses well apart from the detection limit.

The mobile phase should be degassed. This measure improves the repeatability of quantitative analyses, decreases baseline noise, and may prevent from extra peaks in the chromatogram. Degassing can be done by helium sparging of the eluent in the reservoir (helium is less soluble in solvents than nitrogen and oxygen) or by on-line vacuum treatment in a special module placed in front of the pump.

4.3　Pumps

The HPLC pump must deliver a constant flow, selected by the analyst, despite the fact that the pressure depends on the column properties (length and particle diameter) and on the flow rate. During a typical gradient elution, for example, with water and acetonitrile, the viscosity of the eluent mixture changes over time, but the flow rate must remain constant anyway. It becomes clear that such a pump is a high-performance instrument, which must be handled carefully. The parts that can (and in many cases
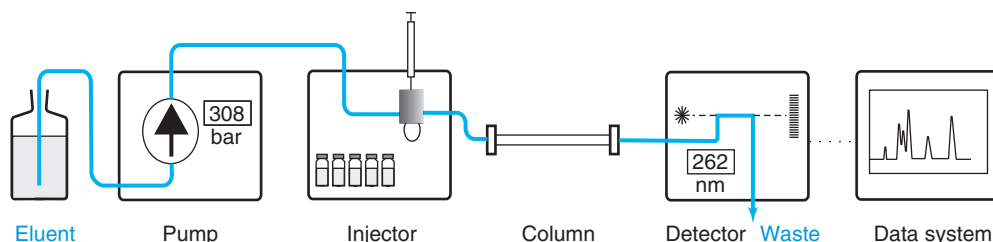
Figure 6　The modules of a basic HPLC instrument. The flow of the mobile phase is shown in blue.

must) be replaced are the piston made from sapphire or ceramics, its seal, made, for example, from Teflon and graphite, and the check valves with ruby balls and sapphire seats. To prevent damage of these delicate items, two rules must be observed:

- A pump must not run dry!
- Do not switch off the pump overnight or longer if it contains buffer or another salt solution! The salts might crystallize. Either the solution is replaced by pure water or a low flow (e.g., $0.1 \, \text{mL} \, \text{min}^{-1}$) is maintained until the next separations will be run.

4.4 Injectors and Autosamplers

The heart of the injector is a multiport valve, usually with six ports, which allows to fill a loop with a defined volume by a syringe (or similar device) and to switch the loop into or out of the mobile phase flow path. The loop has a typical volume of $1–100 \, \mu\text{L}$ (but can also be less or more). It should be filled with the sample solution to either not more than 50% of its volume or with its multifold volume (e.g., fivefold). Filling a loop exactly to its volume may lead to poor precision because of imperfect flow profiles. This valve can be operated manually or by the control unit; the interior rotor seal of a six-port valve is turned by $60°$, thereby opening or closing the way of the eluent through the loop. Its wearing parts are the syringe seal and the rotor seal, a disk with fine channels usually made from vespel. For applications in the basic range over pH 10, tefzel seals are needed, which have poorer mechanical properties.

An autosampler is an instrument module with a built-in multiport valve, a rack for the sample vials, and some robotics for moving either the valves or the syringe or both. Many autosamplers can be equipped with a cooling element if temperature-sensitive samples must be analyzed.

5 DETECTION AND INTEGRATION

5.1 General Requirements

Analysts wish rugged and cheap detectors with a wide linear range (i.e., a linear relationship between analyte mass or concentration and signal intensity).

They should be either selective and register only the analytes of interest, such as the fluorescing ones, or nonselective, which means that such instruments register any change in the composition of the column eluate. For trace analysis, their detection limit should be low. Their electronic time constant must be adjustable and must match the requirements of the actual separation, that is, of the peak width. The same is true for the detector cell volume because it must not dilute the eluting peaks. The narrower the peaks are, the shorter must be the time constant and the smaller the cell volume. However, too low a time constant results in increased baseline noise, and too small a cell volume results in a poor detection limit.

Most HPLC detectors are concentration sensitive, which means that they register the actual concentration of an analyte in the eluent but not its mass. Mass-sensitive devices are the coulometric and light-scattering detectors; concerning the mass spectrometers used as detectors, their sensitivity depends on the ionization and measuring principle.

Besides the detectors mentioned in the following sections, there exist also instruments with other detection principles: infrared, electrical conductivity, photoconductivity, radioactivity, or viscosity.

5.2 UV Detection

The most widespread HPLC detectors are UV detectors that are relatively cheap, easy to operate, and rugged, and which have a wide linear range of four concentration decades (Lendi and Meyer, 2005). They are equipped with a deuterium lamp emitting a continuous spectrum from <200 to 340 nm. In addition, also a tungsten-halogen lamp can be built in, which expands the spectrum into the visible range up to 850 nm. The detection wavelength can be selected within the offered range; it has a spectral bandwidth of approximately 10 nm (too narrow a bandwidth would result in low light intensity). A conventional detector cell has a volume of $8 \, \mu\text{L}$ at a length of 10 mm, but for narrow-bore columns of highest performance, the volume must be smaller.

Diode array instruments are a special type of UV detectors in which the cell is not irradiated with a selected wavelength but with the full lamp spectrum. After passing through the cell, the light is divided spectrally by a grating and the resulting light cone falls on an array of individual light-sensitive diodes.

The spectral resolution is 1 nm or even less. The most interesting feature of a diode array detector is its ability to obtain the UV spectrum of a peak in real time; therefore, it is the simplest possibility to couple HPLC with spectroscopy. If the spectra of the expected analytes are known, peaks can be identified by them although UV spectra have limited informative value. The diode array detector can also be used for peak purity tests, although it can only prove that a peak is not pure; the opposite result can be a deception (Papadoyannis and Gika, 2004).

UV and UV–Vis detectors register all analytes absorbing in the UV or Vis, that is, most organic molecules. Depending on the analytical task, this low selectivity can be an advantage or a disadvantage.

5.3 Fluorescence Detection

Some classes of compounds, such as the aflatoxins, show fluorescence when irradiated with the appropriate UV wavelength. The emitted wavelength is higher than the excitation wavelength. The fluorescence light that comes out of the cell is detected at a 90° angle to the incident light beam. Fluorescence detection gives markedly lower (better) detection limits than UV detection, and its selectivity is a most valuable tool for the search or analysis of fluorescing molecules.

5.4 Refractive Index Detection

If analytes do not absorb in the UV, the RI detector can be used. Its advantage is also its drawback, namely that it detects all RI fluctuations in the eluate, be they from an analyte, from a composition change of the mobile phase (e.g., owing to fluctuations in the gradient mixing system), or from a change in temperature. RI detectors have higher (i.e., worse) detection limits than UV detectors. They are rather cheap, and their operation is simpler than with the alternative light-scattering detector.

5.5 Light-Scattering Detection

The evaporative light-scattering detector (ELSD) is used for nonvolatile analytes without UV absorption.

The eluate from the column is nebulized in a stream of inert gas. Subsequently, the droplets are evaporated and the produced solid analyte particles transported into a light beam. Finally, the scattered light is detected. Unfortunately, the relationship between analyte mass and signal is not linear. The mobile phase and all additives (buffer salts, etc.) must be volatile but the composition of the eluent may change during the chromatographic run, in contrast to RI detection. Light-scattering detection is a valuable tool for classes of compounds such as saponines, oligosaccharides, or lipids.

5.6 Electrochemical Detection

Some functional groups can be oxidized electrochemically: aromatic hydroxides, aromatic amines, indoles, phenothiazines, mercaptans, and others. The working electrode that is in contact with the eluate has a positive potential. (Reduction of suitable analytes is also possible but less convenient.) The mobile phase must be conductive but not necessarily aqueous. The electrochemical conversion rate is usually approximately 10%; if it is 100%, the technique is called coulometric detection. Owing to its selectivity for a limited number of classes of compounds, electrochemical detection can be the method of choice for these analytes.

5.7 Mass Spectrometric Detection

In principle, HPLC and mass spectroscopy (MS) are difficult to couple because the first method is liquid based, whereas the second one is vacuum based. When liquids are evaporated, a large amount of vapor is generated, which must be removed before it enters the mass analyzer. What was a great challenge decades ago has been developed to a rather simple, convenient, small-sized, and affordable detection method (Holčapek, Jirásko, and Lísa, 2012). Various coupling and ionization systems are in use, all having their prerequisites, advantages, and limitations. Some techniques yield mass spectra of the analytes including the molecule ion $(M+H)^+$, which can be used for structure elucidation or confirmation. Others yield fragments only. With single ion detection, one gets a highly selective analytical method, which reduces the necessity of complete chromatographic resolution

because all analytes that yield ions or fragments of another mass are not detected, even if they are coeluting with the investigated molecules.

HPLC-MS is of utmost importance in research; this coupling technique is presented and discussed elsewhere in this handbook, see **LC and LC-MS: Techniques and Applications**.

5.8 Coupling of HPLC with Nuclear Magnetic Resonance

The investigation of the eluted fractions by nuclear magnetic resonance (NMR) is possible, either on-line and in real time or after storage in capillaries or other reservoirs (Sturm and Seger, 2012; Gonnella, 2012). In the latter case, the measuring times can be as long as needed. One of the mobile phase solvents can be used as in conventional HPLC (e.g., acetonitrile) because its NMR signal can be suppressed, whereas the other one (or all the others) must be deuterated (e.g., water). High radiofrequency such as 500 MHz is needed; therefore, this method is extremely expensive. As with MS, the immediate structure elucidation is possible. The combined coupling of HPLC with diode array, MS, and NMR detection is a research tool of unsurpassed investigative power.

HPLC-NMR is presented and discussed elsewhere in this handbook.

5.9 Integration

The size of chromatographic peaks is usually proportional to the amount or concentration of the analyte in the column eluate (depending on the detection method and the linear range of the detector). "Size" means peak height or area. Usually the peaks are evaluated by area; however, in trace analysis with low signal-to-noise ratio, peak heights give better repeatability than areas. Strictly speaking, the term "integration" means area determination but the measuring of heights is also subsumed under it.

For quantitative analysis, it is necessary to compare the peak size of a standard with the one of the analyte under identical chromatographic conditions. Accurate integration is in many cases not possible if the peaks are not well, that is, baseline resolved (Meyer, 1995), and if the true position of the baseline

is unknown. Integration can be influenced by a number of parameters such as data acquisition rate, time constant, and threshold (Dyson, 1998). While the default settings of the integrator should be suitable for the analyses usually performed with the respective HPLC instrument, knowledge of these parameters and how they act on the quantitative analysis is recommended. Narrow peaks must be integrated with a higher acquisition rate than broader ones because the number of data points over the width of a peak should at least be 10.

6 SEPARATION MODES

Details about the different HPLC modes mentioned here can be found in the respective chapters of the textbooks (Meyer, 2010b; Snyder, Kirkland, Dolan, 2010).

6.1 Adsorption Chromatography (Normal-Phase Chromatography)

Because chromatography was invented on adsorbents, this mode was the "normal" one, leading decades later to the term "normal-phase chromatography" (NP). Here the stationary phase is polar, usually underivatized silica but sometimes also alumina (alox, Al_2O_3). The mobile phase is nonpolar or moderately polar. The polarity of organic solvents is listed in the so-called eluotropic series with the *n*-alkanes being at the nonpolar end and water (or salt solutions) at the polar end. Typical eluents for adsorption chromatography are hexane or heptane (less toxic) and all solvents miscible with it, for example *tert*-butyl methyl ether, dichloromethane (should be avoided owing to ecological reasons), tetrahydrofuran, isopropyl alcohol, or ethyl acetate (Meyer and Palamareva, 1993). The sample solution must be nonaqueous.

Polar analytes are retained most, and less or nonpolar ones are eluted early in the chromatogram. Adsorption chromatography is well suited for the separation of geometrical isomers such as *cis–trans* or *ortho–meta–para* mixtures.

6.2 Reversed-Phase Chromatography

In this HPLC mode, the stationary phase is nonpolar, typically a C_{18} material (ODS phase, see

Section 3.2). The eluent is aqueous with water as the first component of the mobile phase and a miscible organic solvent as the second one: usually methanol, acetonitrile, or tetrahydrofuran. Consequently, polar analytes are eluted first and the less or nonpolar ones will appear later. The elution pattern is generally (but not in detail) inverted compared to adsorption chromatography; therefore, the method got the name "RP". The sample solution should be aqueous.

The viscosity of RP solvent mixtures is higher than in the case of purely organic solvents as used in adsorption chromatography; therefore, the flow resistance is higher and diffusion is slower. This leads to higher pressures than in NP HPLC; on the other hand, the van Deemter optimum is at lower flow rates although this fact is often ignored and separations are performed at too high flow rates (which is of minor importance if the C term of the van Deemter curve is not steep).

RP HPLC is of utmost importance for the analysis of aqueous samples such as biological and environmental samples. It can be used for almost all classes of compounds. Its strength, although rarely utilized, is the separation of homologs. Its most important feature is the fact that bonded stationary phases on silica can easily be used with changing composition of the mobile phase, a technique known as gradient elution (see Section 7.1). This allows the separation of mixtures that contain a vast polarity range of analytes.

Users of RP HPLC must be aware of the fact that there is an immense number of stationary phases available on the market with the designation C_{18} or ODS. Their properties may differ considerably owing to differences in the silica used, the bonding chemistry, the bonding density, and so on (see Figure 7). Consequently, one must specify exactly which column was used for a certain separation. If one wants to repeat a published separation, the same stationary phase should be used; if it is not at hand or cannot be purchased, one has to check which column is similar and could be used instead. An excellent list is published by the U.S. Pharmacopoeial Convention (USP Chromatographic Columns, 2011) based on comprehensive research data (Snyder, Dolan, and Carr, 2007).

6.3 Hydrophilic Interaction Chromatography

Hydrophilic interaction chromatography (or hydrophilic interaction liquid chromatography (HILIC)) is
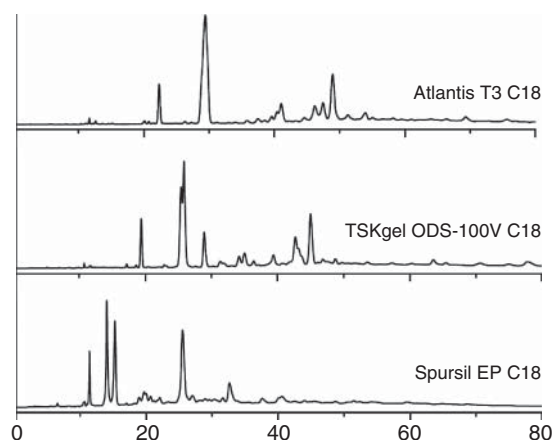
Figure 7 Stationary phases of the same general class, here RP *n*-alkane $C_{18}H_{37}$ materials of different manufacturers, can have very different separation properties. Sample: extract from *Radix isatidis* (used in traditional Chinese medicine). Stationary phases: as indicated. Column dimensions: 15 cm length, 4.6 mm inner diameter in all cases. Mobile phase: water/acetonitrile, both with 0.1% trifluoroacetic acid, 0.85 mL min^{-1}. Gradient: from 0% to 5% acetonitrile within 0–60 min, then isocratic. Detection: UV 280 nm. Temperature: 30 °C. (*Source*: J. Zeng, *et al*. (2010). Reproduced from WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.)

a more recent separation mode of HPLC (Buszewski and Noga, 2012). It allows the separation of polar analytes on polar or moderately polar stationary phases using an aqueous mobile phase. The stationary phase can be underivatized silica as well as bonded amino, nitrile, amide, or polyethylene glycol moieties and numerous other chemistries. They all should be especially designed for HILIC because conventional silica etc. may give poor results. The eluent is based on an aqueous buffer solution or an alcohol; the second component, important for gradient separations, is, for example, acetonitrile. The content of this less polar solvent decreases during gradient elution (which is the opposite situation compared to reversed-phase HPLC, see Section 7.1). Less polar analytes are eluted earlier than the more polar ones, as in adsorption chromatography.

The strength of HILIC lies in the separation of very polar analytes such as carbohydrates, glycosylated compounds, peptides, and many more, which are not retained in RP systems (because their solubility in the eluent is high) and not eluted in NP systems (because the nonaqueous mobile phase is not polar enough).

6.4 Medium-Polar Bonded Stationary Phases

The classical bonded phases on silica are the ones with diol, amino, nitrile, or nitro groups although in principle any functionality can be prepared. These materials can be handled in the NP or in the RP mode, that is, with a nonpolar or a polar eluent. They can offer high selectivity for certain classes of compounds:

Diol: for steroids, organic acids, tetracyclines, surfactants, biopolymers, and so on, namely for analytes that form hydrogen bonds with the diol function.

Amino: for sugars and glycosylated compounds. It is also a weak anion exchanger.

Nitrile (*cyano*): for analytes with double bonds, the best phase for tricyclic antidepressants.

Nitro: for aromatic compounds, especially also for polycyclic aromatic hydrocarbons and heterocyclic analytes.

6.5 Ion-Exchange Chromatography

Stationary phases with acidic groups are cation exchangers, and those with amino groups are anion exchangers. They interact with ionic or ionizable analytes. The mobile phase is a buffer of suitable pH. Silica-based ion exchangers cannot be used over the full pH range, especially not at high pH. Therefore, polystyrene-based HPLC materials are of widespread use although their pressure stability is limited. In principle, there are four classes of ion exchangers:

Strong cation exchangers (SCX) with sulfonic acid $-SO_3^-$ groups, charged over the full pH range.

Weak cation exchangers (WCX) with carboxylic acid $-COO^-$ groups, not charged at low pH, fully charged above pH 6, and partially charged in-between.

Strong anion exchangers (SAX) with quaternary amino groups $-NR_3^+$, charged over the full pH range.

Weak anion exchangers (WAX) with e.g. diethylamine groups $-NH(CH_2-CH_3)_2^+$, fully charged below pH 7, not charged at high pH, and partially charged in-between.

Only charged groups can act as ion exchangers. However, the chromatographic retention is also governed by the counterions (the type of ions present in the mobile phase) of the buffer in use. An important tool to obtain resolution is pH gradients because the analytes will be not, partially, or fully charged, depending on their pK_S value, at changing pH of the eluent. Similarly, the stationary phase will change its retentive properties during a pH gradient if it is a weak exchanger.

While classical ion exchange is suited for organic analytes whose pK_S values are not extreme, that is, whose charge depends on the pH of the eluent, it fails for the separation of the ions of strong acids and bases, such as nitrate, phosphate, sulfate, or ammonium and sodium. For such analyses, a variant called *ion chromatography* was developed (Fritz and Gjerde, 2009). Complex mixtures with anions and cations can be separated by ion-pair chromatography where detergent-like counterions are added to the eluent (Cecchi, 2009).

6.6 Size-Exclusion Chromatography

Porous stationary phases for HPLC have a well-defined pore size and pore size distribution. It is obvious that large macromolecules cannot enter, for example, a material with 10 nm (100 Å) pores, whereas small analyte molecules (as well as the ones of the mobile phase) can fully penetrate its pore system. Molecules of medium size can enter the pore system, but the volume that is available to them is limited. Consequently, properly designed porous materials will separate a mixture of molecules of different size (which may be proportional to their molecular mass) according to their size. Large molecules will be eluted first, small ones last; the separation is not influenced by adsorption (retention) phenomena of any kind if the mode is pure size-exclusion chromatography (SEC) (Striegel *et al.*, 2009).

A certain stationary phase with its pore size can only separate a certain molecular mass range; usually 1–2 ranges of magnitude. Therefore, many types of SEC phases are commercially offered in various pore sizes. SEC is classified into two sub-modes, GPC and GFC: in GPC, the eluent is nonpolar, thus this technique is especially used for the investigation of synthetic polymers (plastics). In GFC,

the eluent is aqueous; therefore, it is used for biopolymers.

6.7   Affinity Chromatography

This separation mode, in many cases used in low-pressure column chromatography or even in membranes, uses highly specific biochemical interactions (Hage, 2006; Zachariou, 2007). The stationary phase contains functional moieties such as antibodies or enzymes (maybe only the part with the active site). These groups retain the matching analytes, whereas all other compounds present in a sample will be eluted without retention. After a change in the mobile phase composition (pH, ionic strength, counterion, etc.), the retained molecules will be washed out of the column. Therefore, affinity chromatography is a purifying method that can be used not only for preparative isolation but also for the undisturbed quantification of a wide range of interesting analytes.

7   SEPARATION STRATEGIES

7.1   Gradient Elution

If the eluent composition remains constant during a liquid chromatographic separation, the method is called *isocratic elution*. The peaks will become broader over time because the theoretical plate number of the column is constant (at least in principle). In many cases, an isocratic separation is not expanded over approximately $k = 20$ (or 21 times the breakthrough time t_0); later eluting peaks would be difficult to integrate. Strongly retained peaks will not be eluted at all.

The way out of this so-called general elution problem is found when the elution strength of the mobile phase is increased over time. This mode is known as *gradient elution* (Snyder and Dolan, 2007), and it is most popular in RP chromatography although it can be used in all HPLC modes (even adsorption chromatography). Gradient separations open the door to large peak capacities.

In a typical RP gradient separation, the mobile phase composition at the beginning is, for example, 90% water or buffer (weak or A solvent)/10%

acetonitrile (strong or B solvent). Over the separation time, the acetonitrile content is continuously increased. Because the organic solvent is a stronger eluent than water in this phase system, its increasing concentration changes the partition equilibrium of the analytes between stationary and mobile phases toward the latter. Compounds with an initial preference of the stationary phase are then gradually eluted as sharp peaks. Figure 8 shows an unsatisfactory isocratic separation of carotenoids and its improvement by gradient elution.

Gradient profiles can be steep or shallow, linear, curved, or composed with parts of different slope, including isocratic time intervals anywhere in the chromatogram. If a simple linear gradient from x% to y% B does not provide a satisfactory separation, it needs to be improved, which can be time-consuming even for experienced analysts. Computer simulation software is an almost indispensable help (Jayamaran *et al.*, 2011). Although most gradient separations are performed with two eluents, it is possible to use three or even four different ones for very demanding analytical problems with a large number of different classes of compounds. The HPLC instrument must be equipped with either the necessary low-pressure or high-pressure gradient modules (see Section 4.1). RP stationary phases are well suited for gradient separations, thanks to their fast re-equilibration to the starting conditions after the end of a run.

7.2   Comprehensive Two-Dimensional Separations

Many (probably most) samples of biological origin contain an incredibly high number of different molecules. The peak capacity of even the best chromatographic columns under the most sophisticated elution modes is limited (see Section 2.6). Consequently, there is always the risk of coeluted analytes that will not be discovered. Therefore, two-dimensional chromatography was developed: The eluate from the first (usually long) column is separated again on a second (usually short) column of different selectivity. If this procedure is performed on-line and over the full time range of the first separation, the method is called "*comprehensive two-dimensional chromatography*" (Dugo *et al.*, 2008; Carr, Stoll, and Wang, 2012).
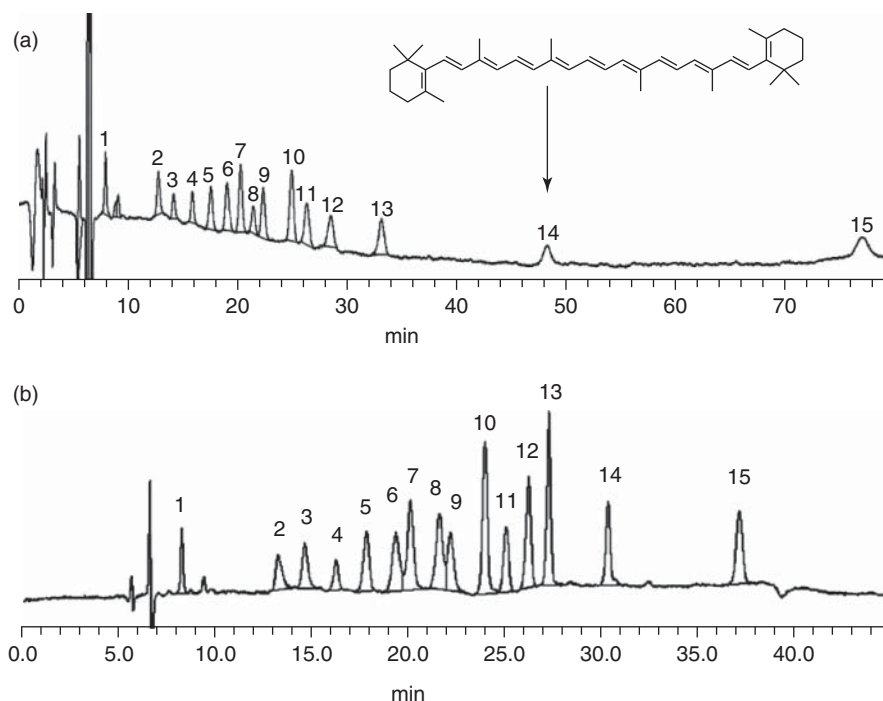
Figure 8   Separation of 15 all-*E*-carotenoids. Stationary phase: Suplex pKb-100 5 μm (alkylamide phase on silica). Column dimensions: 25 cm length, 4.6 mm inner diameter. Mobile phase with acetonitrile (ACN), *tert*-butyl methyl ether (*t*BME), and water (H_2O), 0.5 mL min^{-1}. Detection: VIS 410 nm. Temperature: 20 °C. (a) Isocratic separation, the last analyte is eluted after 77 min. Mobile phase: ACN:*t*BME:H_2O 70 : 20 : 10. (b) Optimized gradient separation with shorter analysis time. Eluent A: ACN:*t*BME:H_2O 70 : 20 : 10. Eluent B: ACN:*t*MBE 70 : 30. Gradient profile: 0–10 min 100% A; 10–14 min 0–60% B; 14–28 min 60% B; 28–45 min 100% A. Analytes: 1, astaxanthin dimethyl disuccinate; 2, astaxanthin; 3, adonirubin; 4, canthaxantin; 5, capsanthin; 6, adonixanthin; 7, β-apo-8′-carotenal; 8, asteroidenone; 9, citranaxanthin; 10, β-apo-8′-carotenoic acid ethyl ester; 11, lutein; 12, zeaxanthin; 13, echineone; 14, lycopene; 15, β-carotene. (*Source*: K. Mitrowska, *et al*. (2012). Reproduced from Elsevier.)

The separation time on the second column must be short compared to the analysis time on the first one, for example, 1 min, which means that every minute the next portion of eluate is sent to the second column. Both involved mobile phases must be miscible with each other. The stationary phases must be selected such that their orthogonality is high, that is, their separation modes are different (Gilar *et al*., 2005). Interesting column pairs can be reversed-phase and ion exchange, gel filtration and reversed-phase, octadecyl and nitrile phase, and so on. Figure 9 presents the two-dimensional separation of steroid glycosides on C18 and polyamine phases (Liu *et al*., 2012b).

Usually the first dimension is performed in the gradient mode, whereas the second separation is isocratic. For highly demanding separations, gradients are also necessary in the second dimension, an approach that is not simple because the gradient times on the second column are short (Jandera, 2012).

8   ULTRA HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY

As explained in Section 2.6, the separation speed and the peak capacity of all chromatographic modes are limited by diffusion phenomena. This physical barrier cannot be removed, and all possibilities to shorten analysis times and/or to increase peak capacity lead to higher pressures needed. In the course of the 1970s, reliable HPLC pumps were developed to marketability; they could be used up to a maximum of 300–350 bar. For decades, this limit restricted the work of the analysts especially in research. Only after
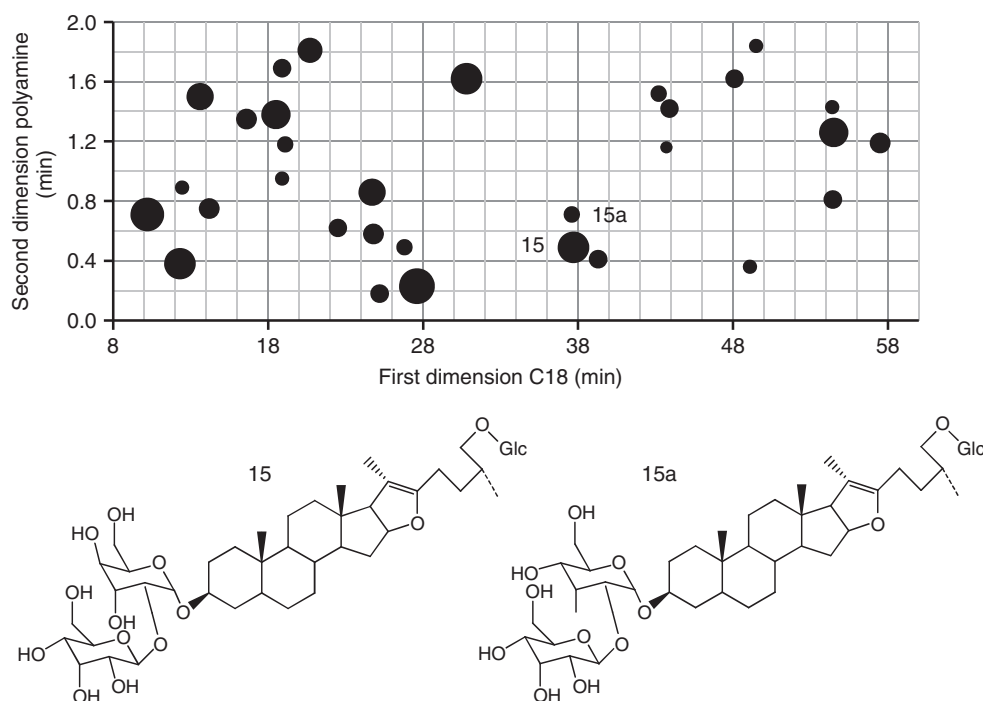
Figure 9 Comprehensive two-dimensional separation of *Anemarrhena asphodeloides* extract, first on a reversed-phase column (Waters Acquity BEH C$_{18}$), second on a polyamine column (YMC-Pack Polyamine II). Mobile phase: water/acetonitrile with 0.1% formic acid. First dimension with gradient elution, second dimension isocratic. Time-of-flight MS detection. The size of the circles is proportional to the corresponding peak area. The structures of two isomers are shown: 15 timosaponin B-III, 15a timosaponin C. They cannot be resolved on the first column alone because their retention times are identical. (*Source*: Adapted from Z. Liu *et al*. (2012). © John Wiley & Sons, Ltd.)

the year 2000, commercial pumps with a pressure limit of 600–1000 bar were available and the new term UHPLC was coined, with U for ultra (Wu and Clausen, 2007; Neue *et al*., 2010). (Actually, the first systems were developed by Waters, and this company created the term UPLC®, which must not be used by other manufacturers.) Figure 3 contains a 1000-bar line that shows the expansion of the (isocratic) separation performance when working at such high pressure. It is obvious that the combination of UHPLC with gradient elution leads to even higher peak capacities, unreachable in the twentieth century with commercial equipment.

For successful UHPLC, one cannot just buy a UHPLC pump and use it with traditional HPLC instrumentation. First, the packing materials must be especially designed for the ultra-high pressures. Second, columns with, for example, 1.7-μm packings (and of considerable length) fully utilize the

expanded possibilities; however, they generate narrower peaks than traditional columns do; therefore, the data acquisition and processing must be rapid. Third, all parts of the HPLC system such as mixers, injectors, capillaries, or fittings must tolerate the pressure maximum generated by the pump.

Figure 10 shows an impressive example of how the peak capacity can be improved from HPLC to UHPLC (Wolfender *et al*., 2011). The peaks in the second chromatogram are much narrower because the UHPLC column has the threefold number of theoretical plates compared to the original one. In research and method development, but even in routine analysis, UHPLC is frequently coupled with MS detection, thus generating a huge amount of information. The coupling is not trivial but manageable if done properly (Rodriguez-Aller *et al*., 2013).

HPLC and UHPLC are still in full development, and column liquid chromatography is a method well
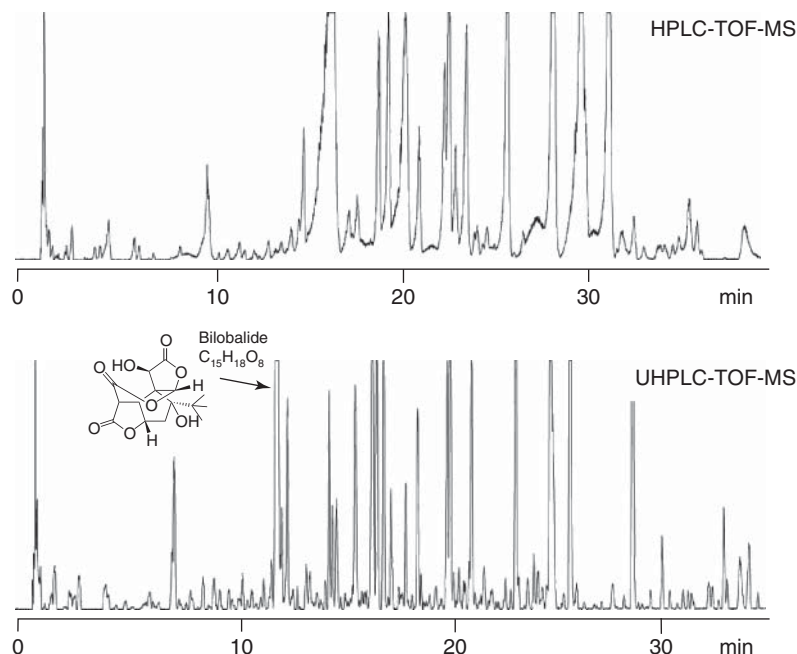
Figure 10 Comparison of HPLC and UHPLC separations of *Gingko biloba* extract. Gradient: 5–40% acetonitrile in 60 min. Detection with time-of-flight mass spectrometry (TOF-MS). HPLC – column dimensions: 15 cm length, 4.6 mm inner diameter. Stationary phase: 5 μm. Flow rate: 1 mL min^{-1}. UHPLC – column dimensions: 15 cm length, 2.1 mm inner diameter. Stationary phase: 1.7 μm. Flow rate: 0.35 mL min^{-1}. (*Source*: J.L. Wolfender, *et al.* (2011). Reproduced by permission of the Swiss Chemical Society.)

prepared to cope with future analytical demands (Carr *et al.*, 2011).

REFERENCES

Batts, J. W. (2011) *All About Fittings*, 3rd printing edn, Oak Harbor, Idex Health & Science.

Buszewski, B., and Noga, S. (2012) *Anal. Bioanal. Chem.*, **402**, 231–247.

Carr, P. W., Stoll, D. R., and Wang, X. (2011) *Anal. Chem.*, **83**, 1890–1900.

Carr, P. W., Davis, J. M., Rutan, S. C., *et al.* (2012) Principles of online comprehensive multidimensional liquid chromatography, in *Advances in Chromatography*, eds. E. Grushka and N. Grinberg, CRC Press, Boca Raton, vol. **50**, pp. 139–235.

Cecchi, T. (2009) *Ion-Pair Chromatography and Related Techniques*, CRC Press, Boca Raton.

Davis, J. M., and Giddings, J. C. (1983) *Anal. Chem.*, **55**, 418–424.

Dolan, J. W., and Snyder, L. R. (1989) *Troubleshooting LC Systems*, Chester, Aster.

Dugo, P., Cacciola, F., Kumm, T., *et al.* (2008) *J. Chromatogr. A*, **1184**, 353–368.

Dyson, N. (1998) *Chromatographic Integration Methods*, 2nd edn, Cambridge, Royal Society of Chemistry.

El Fallah, M. Z., and Martin, M. (1991) *J. Chromatogr.*, **557**, 23–37.

Ettre, L. S., and Wixom, R. L. (1993) *Chromatographia*, **35**, 659–668.

Fekete, S., Oláh, E., and Fekete, J. (2012) *J. Chromatogr. A*, **1228**, 57–71.

Fritz, J. S., and Gjerde, D. T. (2009) *Ion Chromatography*, 4th edn, Weinheim, Wiley-VCH.

Gilar, M., Olivova, P., Daly, A. E., *et al.* (2005) *Anal. Chem.*, **77**, 6426–6434.

Gonnella, N. C. (2012) Chromatographic separation and NMR, in *Advances in Chromatography*, eds. E. Grushka and N. Grinberg, CRC Press, Boca Raton, vol. **50**, pp. 93–138.

Guiochon, G. (2007) *J. Chromatogr. A*, **1168**, 101–168.

Hage, D. S. ed. (2006) *Handbook of Affinity Chromatography*, Taylor & Francis, New York.

Hernández-Borges, J., Aturki, Z., Rocco, A., *et al.* (2007) *J. Sep. Sci.*, **30**, 1589–1610.

Holčapek, M., Jirásko, R., and Lísa, M. (2012) *J. Chromatogr. A*, **1259**, 3–15.

Jandera, P. (2012) *J. Chromatogr. A*, **1255**, 112–129.

Jayamaran, K., Alexander, A. J., Hu, Y., *et al.* (2011) *Anal. Chim. Acta*, **696**, 116–124.

Lendi, B. E., and Meyer, V. R. (2005) *LC GC Eur.*, **18**, 156–163.

Li, Y., Aggarwal, P., Tolley, H. D., *et al.* (2012) Organic monolith column technology for capillary liquid chromatography, in *Advances in Chromatography*, eds. E. Grushka and N. Grinberg, CRC Press, Boca Raton, vol. **50**, pp. 237–280.

Liu, K., Tolley, H. D., and Lee, M. L. (2012a) *J. Chromatogr. A*, **1227**, 96–104.

Liu, Z., Zhu, D., Qi, Y., *et al*. (2012b) *J. Sep. Sci.*, **35**, 2210–2218.

Martin, M., and Guiochon, G. (2005) *J. Chromatogr. A*, **1090**, 16–38.

Meyer, V. R., and Palamareva, M. D. (1993) *J. Chromatogr.*, **641**, 391–395.

Meyer, V. R. (1995) *J. Chromatogr. Sci.*, **33**, 26–33.

Meyer, V. R. (2008) *J. Chromatogr. A*, **1187**, 138–144.

Meyer, V. R. (2010a) *Chromatographia*, **72**, 603–609.

Meyer, V. R. (2010b) *Practical High-Performance Liquid Chromatography*, 5th edn, Chichester, Wiley.

Mitrowska, K., Vincent, U., and von Holst, C. (2012) *J. Chromatogr. A*, **1233**, 44–53.

Neue, U. D. (1997) *HPLC Columns – Theory, Technology, and Practice*, Wiley, New York.

Neue, U. D., Kele, M., Bunner, B., *et al*. (2010) Ultra-performance liquid chromatography technology and applications, in *Advances in Chromatography*, eds. E. Grushka and N. Grinberg, CRC Press, Boca Raton, vol. **48**, pp. 99–143.

Papadoyannis, I. N., and Gika, H. G. (2004) *J. Liquid Chromatogr.*, **27**, 1083–1092.

Plumb, R., Mazzeo, J. R., Grumbach, E. S., *et al*. (2007) *J. Sep. Sci.*, **30**, 1158–1166.

Rodriguez-Aller, M., Gurny, R., Veuthey, J. L., *et al*. (2013) *J. Chromatogr. A*, **1292**, 2–18.

Sadek, P. C. (1999) *Troubleshooting LC Systems*, Wiley, New York.

Snyder, L. R., and Dolan, J. W. (2007) *High-Performance Gradient Elution: The Practical Application of the Linear-Solvent-Strength Model*, Wiley, Hoboken.

Snyder, L. R., Dolan, J. W., and Carr, P. W. (2007) *Anal. Chem.*, **79**, 3254–3262.

Snyder, L. R., Kirkland, J. J., and Dolan, J. W. (2010) *Introduction to Modern Liquid Chromatography*, 3rd edn, Chichester, Wiley.

Striegel, A. M., Yau, W. W., Kirkland, J. J., *et al*. (2009) *Modern Size-Exclusion Liquid Chromatography: Practice of Gel Permeation and Gel Filtration Chromatography*, 2nd edn, Chichester, Wiley.

Sturm, S., and Seger, C. (2012) *J. Chromatogr. A*, **1259**, 50–61.

Tswett, M. S. (1906) *Ber. Dtsch. Botan. Ges.*, **24**, 384–392.

Tswett, M. S. (1990) *Chromatographic Adsorption Analysis: Selected Works*, Ellis Horwood, Chichester.

USP Chromatographic Columns. 2011. http://www.usp.org/usp-nf/new-chromatographic-columns-online-database.

Van Deemter, J. J., Zuiderweg, F. J., and Klinkenberg, A. (1956) *Chem. Eng. Sci.*, **5**, 271–289.

Willstätter, R., and Stoll, A. (1913) *Untersuchung über Chlorophyll: Methoden und Ergebnisse*, Springer, Berlin.

Wolfender, J. L., Eugster, P. J., Bohni, N., *et al*. (2011) *Chimia*, **65**, 400–406.

Wu, N., and Clausen, A. M. (2007) *J. Sep. Sci.*, **30**, 1167–1182.

Zachariou, M. ed. (2007) *Affinity Chromatography: Methods and Protocols*, Humana Press, Totowa.

Zeng, J., Guo, Z., Xiao, Y., *et al*. (2010) *J. Sep. Sci.*, **33**, 3341–3346.

9

Near-Infrared (NIR) Spectroscopy in Natural Product Research

Christian W. Huck

Leopold-Franzens University, Innsbruck, Austria

1   INTRODUCTION

Herbal medicine is not only the oldest and most widely used form of medicine in the world but also increasingly popular because of their health-promoting properties (Krüger and Schulz, 2007). Many medicinal and herbal plant properties are related to individual compounds, and their active chemical ingredients in natural products usually are found in the parts per million and/or parts per billion range (Krüger and Schulz, 2007). During the recent couple of years, different pharmaceutical companies have initiated sophisticated plant screening programs, applying biochemical high throughput techniques in order to find new drugs with distinct properties (e.g., anticancer and antibacterial properties) (Schulz, 2004).

For a long time, the use of separation techniques including thin-layer chromatography (TLC), liquid chromatography (LC), gas chromatography (GC), and capillary electrophoresis (CE) hyphenated to mass spectrometry (MS) was focused on the elucidation of isolated compounds from different plant matrices (Krüger and Schulz, 2007; Schulz, 2004). These analytical techniques have been found useful in phytochemical and physiological studies enabling recording a fingerprint and/or identifying single active compounds (Stecher *et al.*, 2003).

Spectroscopic analytical techniques using the NIR wavelength region of the electromagnetic spectrum have been used in the food industry to monitor and evaluate the composition and quality of foods (Williams *et al.*, 1985). Even though the NIR region was the first known nonvisible part of the electromagnetic spectrum and was discovered in 1800 by Herschel, only in the 1950s, first applications of near-infrared spectroscopy (NIRS) for analytical chemistry were developed. Since the 1960s, there was a steady increase of applications with the most dramatic growth in recent 25 years.

NIR spectroscopy is characterized by low molar absorptivities and scatterings. Originally, the NIR region was regarded as having little potential for analytical work; it has nowadays become one of the most promising techniques for molecular spectroscopy. Affordable end powerful computers have further supported the implementation of applications in several fields, including medical, textile, polymer, and pharmaceutical applications (Blanco *et al.*, 1998).

This contribution highlights the fundamental principles of NIR spectroscopy, its applicability, regulatory issues, advantages, and limitations in natural product research.

2   NEAR-INFRARED (NIR) SPECTROSCOPY

Infrared (IR) radiation is the region of the electromagnetic spectrum between the visible (VIS) and the microwave wavelengths (McClure, 2003). In NIR
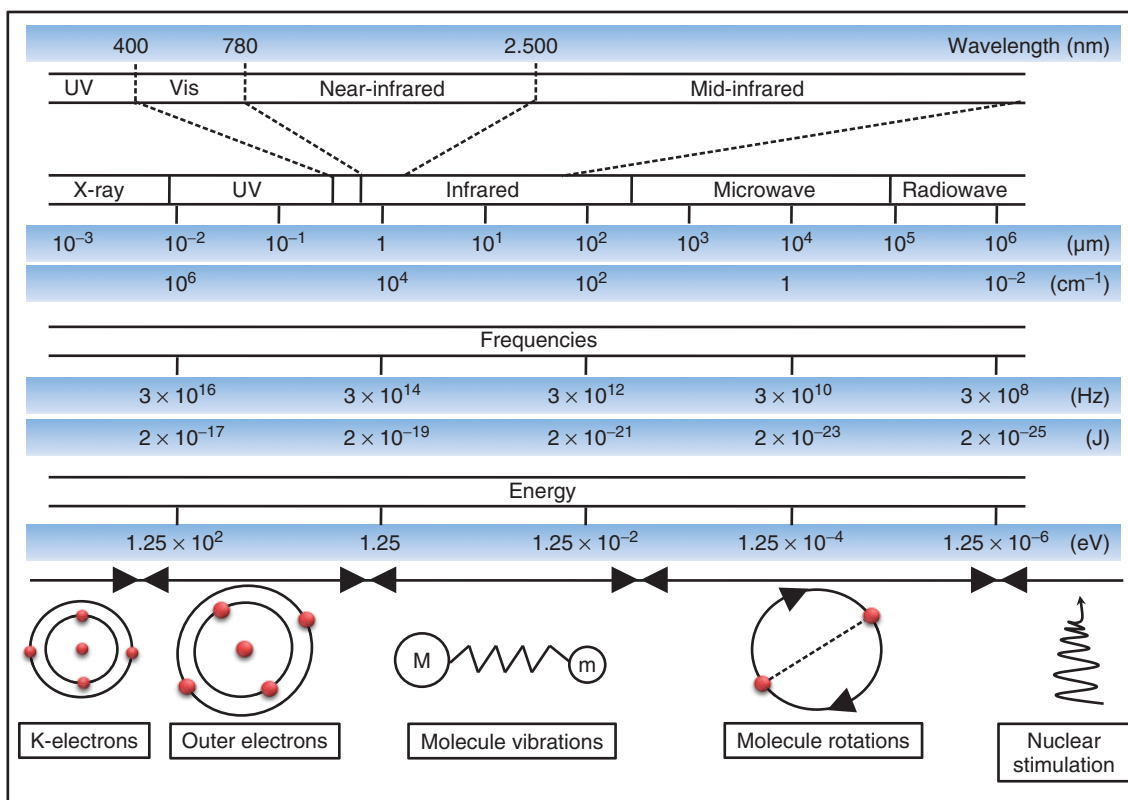
Figure 1   Electromagnetic spectrum.

spectroscopy, excitation of molecules is accomplished in a wavelength range between 750 and 2500 nm (Figure 1), corresponding to a wave number range between 4000 and 13,000 cm^{-1} (Herschel, 1800). Solid, liquid, or gaseous samples can absorb parts of the incoming IR radiation at specific wavelength resulting in a fingerprint or a spectrum (Blanco and Villarroya, 2002). In this spectral region, molecules containing C–H, C–O, C=O, N–H, and O–H functional groups are excited to perform stretching-, deformation-, and scissor-vibrations. In comparison to the mid-infrared (MIR) region, where only fundamental vibrations ("signatures") can be observed, overtones and combinations can be found in the NIR region containing a manifold of information compared to MIR (Barton, 2002). This often results in a crowded spectrum with overlapping peaks. Although NIR intensities are 10–1000 times lower than that of the MIR, highly sensitive spectrometers can be built through several means including the use of efficient detectors (McClure,

2003). The light recorded by the detector contains compositional information that can be unraveled by a computer to report multiple analyses almost instantaneously. NIR spectroscopy can provide simultaneous, rapid, and nondestructive qualitative and quantitative analyses of major components in many organic substances (Osborne, Fearn, and Hindle, 1993).

3   MODEL OF THE HARMONIC AND ANHARMONIC OSCILLATORS

The physical principle describing the observed effects both in the MIR and in the NIR region is the model of the harmonic and anharmonic oscillators. According to the inset in Figure 2, the reduced mass μ performs vibrations with the frequency ν_{osc}. In the MIR, this vibration follows the equation for the harmonic oscillator, whereas in the NIR, the equation for the anharmonic oscillator is valid, describing the

$$\mu = \frac{mM}{m+M}$$ Reduced mass

k = Force constant

$$\nu_{osc} = \frac{1}{2\pi} \sqrt{\frac{k}{\mu}}$$

Harmonic oscillator

Mid-infrared

Anharmonic oscillator

Near-infrared

Potential energy

Atomic distance

$\nu = 4$
$\nu = 3$
$\nu = 2$
$\nu = 1$
$\nu = 0$

$\Delta \nu = \pm 1$

$$E_{osc} = h\nu_{osc} \left(\upsilon + \frac{1}{2} \right)$$

Potential energy

Atomic distance

$\nu = 4$
$\nu = 3$
$\nu = 2$
$\nu = 1$
$\nu = 0$

$\Delta \nu = \pm 1, \pm 2, \pm 3, ...$

$$E_{osc} = h\nu_{osc} \left(\upsilon + \frac{1}{2} \right) - \chi h\nu_{osc} \left(\upsilon + \frac{1}{2} \right)^2$$

χ = anharmonicity constant
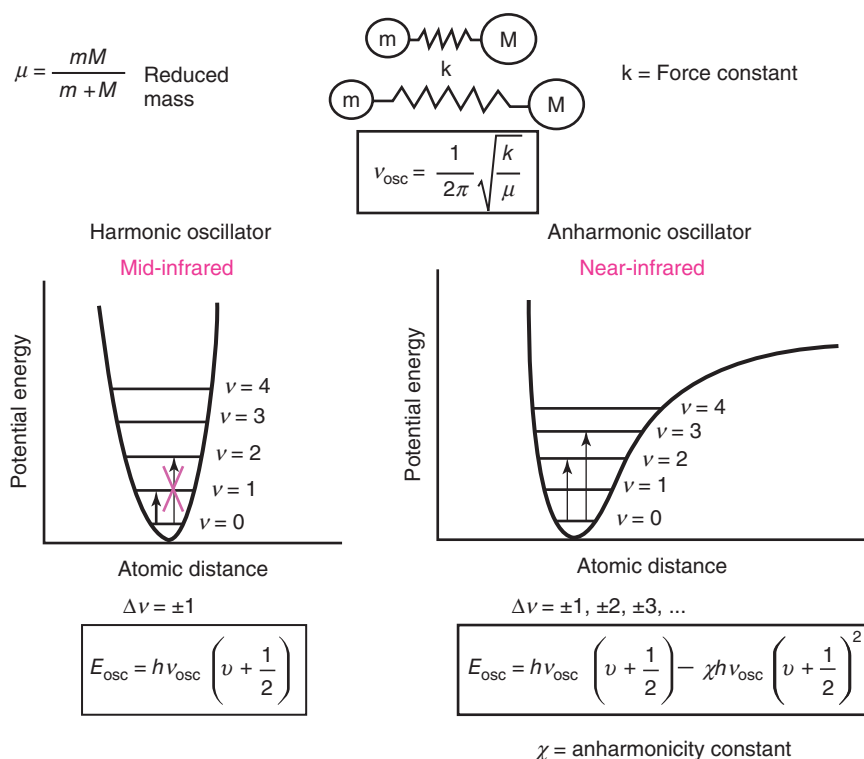
Figure 2 Model of the harmonic and anharmonic oscillators.

excitation into higher energy states. Chemometrics, a mathematical, statistical, multivariate analytical (MVA) tool, is applied for further treatment of recorded spectra (Figure 3).

4 INSTRUMENTATION AND SAMPLE PREPARATION

Owing to the vast amount of work and reviews published related to NIR spectroscopy, the purpose of this section is to give an overall introduction to the different available instruments most commonly applied in natural product analysis.

An NIR spectrophotometer consists of a light source (e.g., tungsten halogen lamp), sample presentation accessories, monochromator, detector, and optical components including lenses, collimators, beam splitters, integrating spheres, and optical fibers (Figure 3) (Williams and Norris, 1987).

One of the most frequently cited benefits of analysis by NIR spectroscopy is that little or no sample

preparation before analysis is required. In principle, transparent materials such as liquid extracts can be analyzed by transmission or transflection, whereas solid materials such as tissue can be analyzed by diffusive reflection and/or interactance mode (Figure 4). In reflection mode, light source and detector are mounted under a specific angle, for example, 45°, to avoid specular reflection. In transmission mode, the light source is positioned opposite to the detector, whereas in interactance mode, the light source and detector are positioned parallel to each other in such a way that light due to specular reflection cannot directly enter the detector (Williams and Norris, 1987).

Spectrophotometers are conveniently classified into dispersive and nondispersive instruments. For example, in a dispersive filter instrument, the monochromator is a wheel holding a number of absorption or interference filters owing the disadvantage of limited resolution (Williams and Norris, 1987). In a scanning monochromator instrument, a grating or a prism is used to separate the individual frequencies of the radiation. In Fourier transform
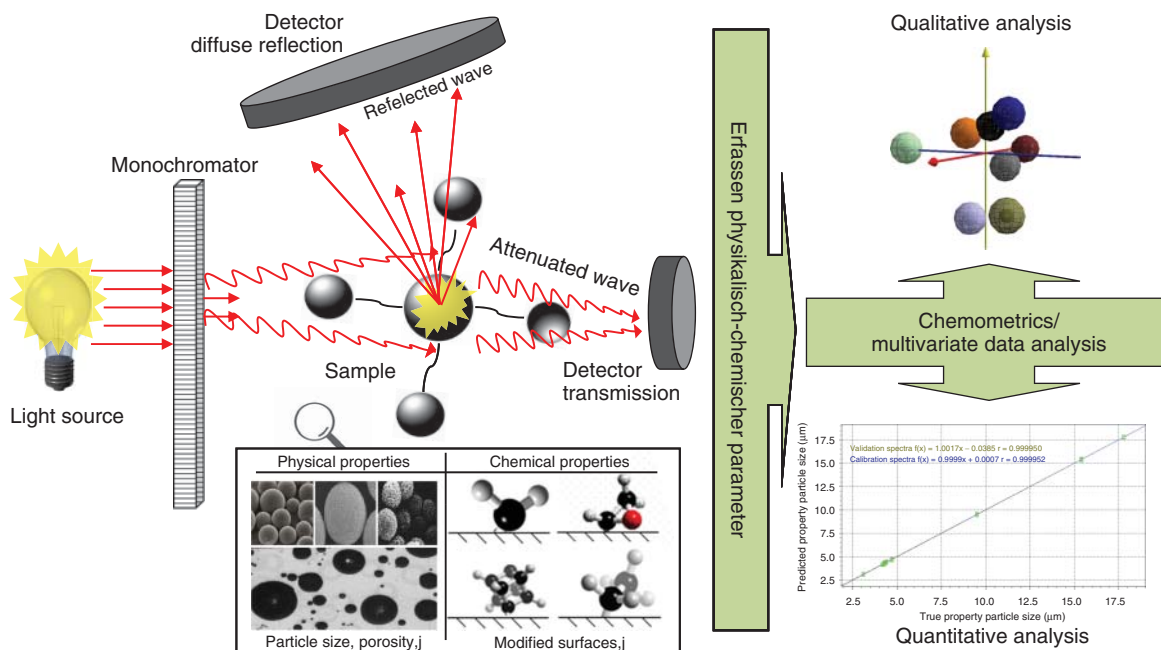
Figure 3   Measurement setup in near-infrared (NIR) spectroscopy.
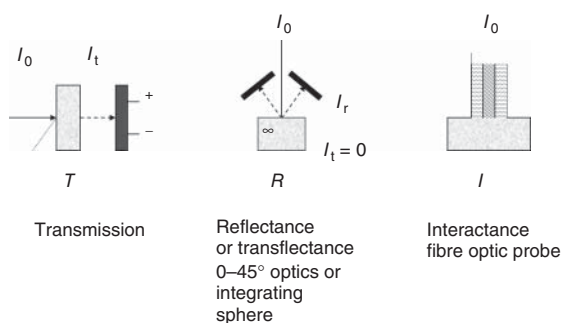
Figure 4   Measurement modes.

spectrophotometers, interferometers are used to generate modulated light and time domain signal of the light reflected or transmitted by the sample can be converted into a spectrum via a fast transformation (Faix, 1992). In most cases, a Michelson or polarization interferometer is used.

Photodiode array (PDA) spectrophotometers consisting of a fixed grating that focuses the dispersed radiation onto an array of silicon (Si, 350–1100 nm) or indium gallium arsenide (InGaAs, 1100–2500 nm) offer the advantage of high acquisition speed (between 50 ms and a few milliseconds).

Therefore, they are often mounted on online lines. Miniaturized, portable, and affordable (less than €30,000) low cost versions are available from several companies, such as Thermo Fisher (USA) or JDSU (Milpitas, CA, USA).

As an alternative laser-based systems without the necessity to use a monochromator, or acousto-optic tuneable filter instruments, using a diffraction-based optical band-pass filter can be effective (Nicolaï *et al.*, 2007). During recent years, liquid crystal tuneable filter (LCTF) instruments have been developed (Nicolaï *et al.*, 2007). In this type, a special filter is used to create constructive and destructive interferences.

Microelectromechanical systems (MEMSs) combine mechanical parts, sensors, actuators, and electronics on a common substrate using microfabrication technology.

During the recent couple of years, hyperspectral imaging systems have become popular (Salzer and Siesler, 2009). A multi- (a few wavelengths) or hyperspectral (a continuous range of wavelengths) cube is recorded consisting of spectra recorded at every 2-D spatial position. The cube is recorded by stepwise moving the object of interest under the camera by means of an actuator while at each step, a line is scanned. In latest developments,

focal plane array detectors are employed (MCT, mercury–cadmium–telluride). In combination with attenuated total reflection (ATR) spectroscopy, the maximal resolution of 1.2 μm can be reached (Salzer and Siesler, 2009; Osborne, Fearn, and Hindle, 1993).

In quantitative analysis, the amount of absorbed radiation is dependent on the Lambert–Beer's law of the concentration c of the sample, the thickness d of the sample, and its molar extinction coefficient ε.

$$E = \log \frac{I_0}{I} = \varepsilon * d * c \qquad (1)$$

where I is intensity of the transmitted light, I_0 intensity of the incident light, c concentration of the absorbing substance (unit: $\mathrm{mol\,dm^{-3}}$ or $\mathrm{mol\,L^{-1}}$), ε decimal extinction coefficient (unit: $\mathrm{mol^{-1}\,dm^2}$), and d thickness of the irritated body (unit: cm).

5 CHEMOMETRICS INCLUDING DATA PREPROCESSING

The NIR spectrum is represented by a huge number of partially overlapping overtones and combination vibrations. Additionally to these scattering effects, instrumental noise and/or sample inhomogeneities might occur (McClure, 2003). The consequence out of this circumstance is the fact that it is in many applications impossible to correctly assign the corresponding vibration bands. Therefore, NIR in combination with multivariate statistical analysis (MVA) is a powerful combination enabling the extraction of the required information from the spectrum (Blanco and Villarroya, 2002). The most appropriate chemometrical procedures include principal component analysis (PCA) for reducing the number of variables facilitating both qualitative and quantitative analyses. Data pretreatment minimizes inhomogeneities originating from the recording of the spectra and enables elimination of baseline shifts. Differences in intensity caused by different sample positioning can be eliminated by normalization algorithms. Diffusion and/or unexpected particle size effects can be compensated by multiplicative scatter correction (MSC). Spectral noise can be reduced by performing the first or second derivative of the original spectrum. Calibration development can mathematically describe the covariation between certain variables or find a mathematical function (regression model), by which the values of the dependent variables are calculated from the values of the measured variables (Siebert, 2001). In this context, it is important to select an optimum number of variables or components. If too many are used, the system can become over-fitted and the model will provide then poor prediction results. On the other hand, using too few components will cause under-fitting, and the model is too small to capture the variability in the data. This fitting effect is strongly dependent on the number of samples and in general more samples give rise to more accurate predictions (Cozzolino, 2009).

The calibration procedure of the NIR spectrometer can be summarized in five steps: (i) choice of a representative sample set; (ii) recording of the NIR spectra; (iii) measurement of the reference values; (iv) multivariate modeling in order to generate a relationship between the recorded spectra and the reference values; and (v) validation of the system. The most frequently used regression methods are principal component regression (PCR), partial least square regression (PLSR), discriminant analysis (DA), and artificial neural networks (ANNs) (Pallua *et al*., 2011).

The choice of the highest suitable regression model is based on the calculation of the following values:

1. Bias, that is, the average deviation between the predicted values (y_n) and the actual values (x_n), in the calibration-set, should be close to zero.

$$\mathrm{BIAS} = \frac{1}{N} \sum (x_n - y_n) \qquad (2)$$

2. Predicted residual error sum square (PRESS) is the sum of the square of the deviation between the predicted values and the reference values. The PRESS value of the validation set should be as small as possible and similar to that of the calibration set.

$$\mathrm{PRESS} = \sum (x_n - y_n)^2 \qquad (3)$$

3. Standard error of estimation (SEE), is the standard deviation of the differences between the reference values and the NIRS results in the calibration set.

$$\mathrm{SEE} = \sqrt{\frac{1}{N} \sum (x_n - y_n - \mathrm{BIAS})^2} \qquad (4)$$

4. Standard error of prediction (SEP) is the counterpart for the test-set samples. SEE and SEP should

be as small as possible.

$$\text{SEP} = \sqrt{\frac{1}{N} \sum (x_n - y_n - \text{BIAS})^2} \quad (5)$$

5. The correlation coefficient (R^2) should approach 1.

6 CHARACTERIZATION OF MEDICINAL PLANTS AND THEIR CONSTITUENTS, SPECIES, AND ORIGIN

In Section 6.1, the potential of NIRS for the classification of the origin of natural products and verification of authenticity is discussed. Section 6.2 points out the potential of NIRS for the quantitative characterization of medicinal plants and their constituents including secondary metabolites and leading compounds.

6.1 Qualitative Analysis

There are several reports in the literature describing the classification, discrimination, or authentication of different natural products by the use of NIR spectroscopy (Cordella *et al.*, 2002): NIRS especially shows high potential for classifying the origin of natural products on the one hand and for verifying authenticity on the other hand. Adulteration is practiced since ancient times because of profit reasons and, in some cases, mixtures can be prepared to hide the misuse. To overcome these problems, several analytical techniques including TLC, GC, LC, and coupled techniques especially to MS have been developed.

St. John's wort: Huck-Pezzei *et al.* (2013) established a procedure to discriminate between pharmaceutical formulations prepared from either *Hypericum perforatum* or *Hypericium hirsutum* originating from China (Figure 5).

Scutellariae Radix: The suitability of NIR to distinguish between 27 cultivated and 22 wild samples collected from nine different regions in China was investigated. Figure 6 shows the mean spectra for all original NIR data of the wild and cultivated varieties. Spectral differences between wild and cultivated plants were enhanced after second derivative

preprocessing. The most intense band in each spectrum could be assigned to the second overtone of the carbonyl group ($5352 \, \text{cm}^{-1}$), followed by the CH stretch and CH deformation vibration ($7212 \, \text{m}^{-1}$), the OH vibration ($4440 \, \text{cm}^{-1}$), the CH_2 ($5742 \, \text{cm}^{-1}$), and the CH_3 overtone ($5808 \, \text{cm}^{-1}$). Furthermore, PCA was used to develop a cluster model for qualitative analysis of wild and cultivated *radix scutellariae*. Figure 7 shows the 2-D cluster plot represented by the first principal component (PC1) and the second principal component (PC2). The total variance explained by the first principal component was 81%. Table 1 compares the different pretreatments and calibrations applied for quantitative analysis. In this study, the best results were obtained when models were based on the $4200–7716 \, \text{cm}^{-1}$ spectral region. Second derivative is obviously superior to other pretreatments, which is demonstrated in the similarity of the results for SEC and SEP.

Tea plant: Xiaoli and Yong analyzed 200 and 93 samples from field experiments (He, Li, Deng, 2007). Wavelet transformation (WT), PCA, and ANN were used to classify the tea samples. The ANN models developed gave good classification accuracy up to 77.3% for the varieties analyzed.

The fast analytical judgment of green, black, and Olong teas in combination with support vector machine (SVM) was reported (Chen *et al.*, 2007). Thereby, the spectral features of each category can be used to differentiate in the NIR region the three tea varieties. The best classification rates were up to 90%, 100%, and 93.33% using the calibration set and 90%, 100%, and 95% using the validation set, respectively.

Coffee: The discrimination between Arabica and Robusta pure coffee varieties and mixtures therefrom was also achieved by NIRS (Huck, Guggenbichler, and Bonn, 2005).

Tanreqing: *Tanreqing* injection is a widely used patent drug in China. It is made from five kinds of TCM (traditional Chinese medicine) extracts, namely *Radix Scutellariae*, *Forsythia Suspense*, *Flos Lonicerae*, Bear gall powder, and *Cornu gorais*, and was used chiefly in treating infection of the upper respiratory tract and serious influenza; the medicine also has satisfactory efficacy on the defense of SARS (Serious Acute Respiration Symptom) and A/H1N1 flu. In its manufacturing process, several kinds of intermediates need to be analyzed to ensure that the operation runs steadily. Li *et al.* (2010) established a qualitative model to monitor the quality of produced
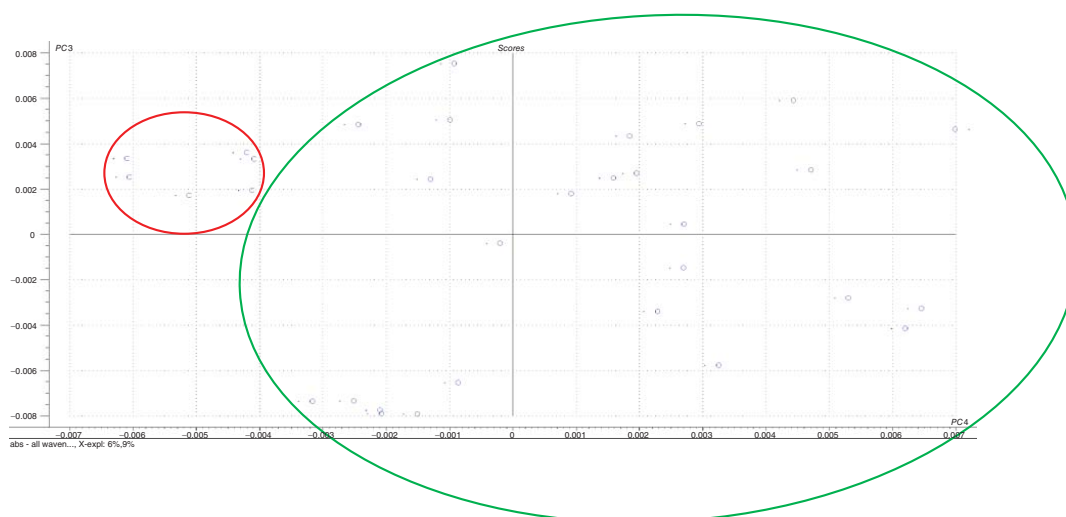
Figure 5 Principal component analysis (PCA) plot to distinguish between European and Chinese proveniences (samples containing additionally *Hypericum hirsutum*).
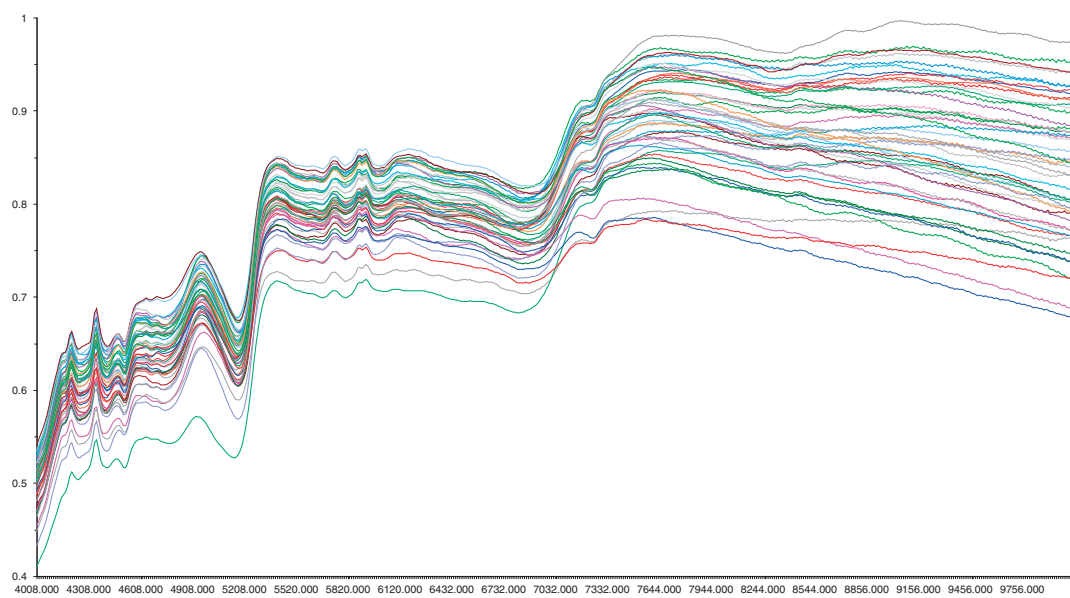
Figure 6 Original NIR spectra recorded from 49 *radix scutellariae* samples.
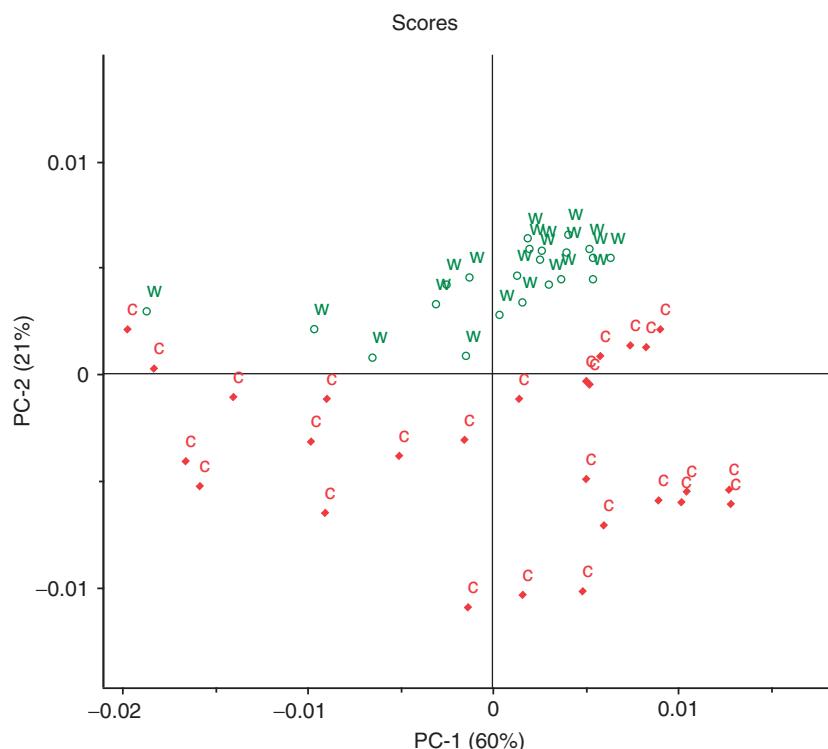
Figure 7 Principal component score values of the wild and cultivated *radix scutellariae* samples.

Table 1 The performance parameters of the established models of B via different spectra preprocessing methods.

Pretreatment	Wave number (cm^{-1})	PLS factors	SEC (%)	SEP (%)	BIAS	R^2 (c-set)	R^2 (v-set)
Without preprocessing	4200–7716	14	0.148	0.142	−0.0411	0.890	0.932
MSC	4200–7716	14	0.180	0.182	−0.0383	0.837	0.893
SNV	4200–7716	14	0.156	0.138	0.0048	0.878	0.942
First derivative	4200–7716	9	0.0768	0.0599	−0.00719	0.970	0.989
Second derivative	4200–7716	9	0.0484	0.0522	−0.0205	0.988	0.990

batches according to process analytical technology (PAT) guidelines recommended by the US FDA.

Saffron: Zalacain *et al*. developed a NIRS method to determine the chemical composition and geographical origin of 111 samples from the main producing countries: Iran, Greece, and Spain (Zalacain *et al*., 2005). Compared to UV–vis and high performance liquid chromatography (HPLC) as reference methods, NIRS was found to be appropriate for determining moisture and volatile content according to the ISO 3632 Technical Specification Normative.

Wine: Guggenbichler *et al*. (2006) investigated different vine grapes (genus *Vitis*), grape varieties, and bottled red wines using fiber-optics-based NIR transflectance and transmission spectroscopy. Recording of spectra in the transmission mode allowed distinguishing between different vine grapes. Totally, 57 bottles of *Cabernet Sauvignon, Lagrein*, and *Chianti Classico* were analyzed. PCA over a wave number range from 4500 to 10,000 cm^{-1} including data pretreatment allowed assigning each wine to a separate cluster with a *Q*-value of 0.72. PLSR additionally allowed determining the mixing ratio of grape varieties in one bottle of red wine. Therefore, this model can help the consumer check the wine for its quality and thus prevent him or her from being misled. In order to achieve a fast and simple quantitative analysis in the transflectance mode of

the carbohydrate, total acid, tartaric acid, malic acid, polyphenol content, and pH value in Weißburgunder, Chardonnay, Ruländer, Silvaner, Müller Thurgau, Gewürztraminer, Sauvignon, Lagrein, Grossvernatsch, Blauburgunder, Cabernet, and Merlot grapes, a NIRS method was established.

Other examples for the qualitative investigations of natural products include tea (He, Li, and Deng, 2007), *ganoderma lucidum* (Chen *et al*., 2008a), cinnamon, (Juliani *et al*., 2006), and several others.

6.2 Quantitative Analysis

In the following section, a summary of recent applications on the quantitative analysis of different compounds in natural products, including medicinal and aromatic plants is described.

6.2.1 *Phenolic Compounds*

Phenolics in several different plants have been investigated by NIR spectroscopy to determine total polyphenols, catechins, and others (Cozzolino, 2009; Bittner *et al*., 2013) also to get knowledge upon the antioxidative (Perron and Brumaghim, 2009), antimicrobial (Rauha *et al*., 2000), antiviral (Perez, 2008), anti-inflammatory, analgesic, antipyretic, and vasodilatory (Padilla *et al*., 2005) effects.

Wine: There are several reports about NIRS application of phenolic compounds in wine. Garde-Cerdán *et al*. (2010) measured 510 aged wines with different storage times and in different oak barrels to calculate PLS (partial least square) calibration models to predict 10 different oak volatile compounds and ethylphenols. In a recent publication, Garde-Cerdán *et al*. (2012) applied NIRS to analyze five different haloanisols and halophenols in barrelled red wine. Totally 600 wines were measured and PLS calibrations were calculated for all wines, age groups, and zone groups of the barrelled wine. Cozzolino *et al*. (2004) used NIRS in combination with multivariate data analysis to simultaneously determine the concentrations of malvidin-3-glycoside, pigmented polymers, and tannins in red wine. PLS calculations resulted in squared correlation coefficients for the calibration above 0.80. A method for the determination of the total phenolic content in red wine was introduced by Guggenbichler *et al*. (2006).

Green tea: NIRS in combination with PLS was described to quantify the epigallocatechin gallate, epicatechin, and trolox equivalent antioxidant capacity (TEAC) in green tea (*Camelia sinensis* L.) leaves. For the control of the total phenolic content in green tea, Folin-Ciocalteu was used as a reference method resulting in an RMSECV of $0.75\,\mathrm{g\,g^{-1}}$ for a calibration range 15.84–$24.39\,\mathrm{g\,g^{-1}}$. (Chen *et al*., 2010a) Employing HPLC as a reference was used for the investigation of the total polyphenol content in four varieties. Chen and coworkers compared the three algorithms PLS, interval partial least square (iPLS), and synergy interval partial least square (siPLS) to predict the total polyphenol content (Chen *et al*., 2008b). The siPLS model performed best with an RMSEP of 0.7327 (range: 15.84–24.39%) using five PLS factors. The same work group reported PLS calibrations to determine the contents of the main catechins. Using 75 samples, epigallocatechin (EGC) (0.14%, range: 2.4–5.4%), epicatechin (EC) (0.017%, range: 0.1–0.4%), epigallocatechingallate (EGCG) (0.38%, range: 7.7–14.1%), and epicatechingallate (ECG) (0.12%, range: 1.8–3.7%) were calibrated using 10–14 PLS factors. Luypaert, Zhang, and Massart (2003) utilized NIRS in combination with PLS algorithms to predict caffeine, EGCG, EC, and total antioxidant capacity. Zhang *et al*. (2004) predict the total antioxidant capacity in green tea using PCR regression with test-set validation with 100 samples in the calibration set and 23 in the validation set. In a recent publication, Chen *et al*. (2012) report about a comparison of PLS, back propagation artificial neural networks (BP-ANN), and support vector machine regression (SVMR) to measure the antioxidant activity in green tea.

Kava: Kava (*Piper methysticum*) has been traditionally used for thousands of years in the South Pacific region. It promotes a state of relaxation without the loss of mental alertness. Gautz *et al* (2006) developed PLS calibration models for kavalactones (desmethoxyangonin, dihydrokavain, yangonin, kavain, dihydromethysticin, methysticin, and total kalvalactones) using a maximum of seven PLS factors. The SEPs were 0.20% (range: 0.08–2.35%) for desmethoxyangonin, 0.31% (range: 0.10–3.33%) for dihydrokavain, 0.47% (range: 0.08–3.02%) for yangonin, 0.21% (range: 0,11–3.02%) for kavain, 0.15% (range: 0.08–2.58%) for dihydromethysticin, 0.19% (range: 0.09–2.70%) for methysticin, and 1.05 % (range: 0.54–14.68%) for total kalvalactones. Kava has become a part of the herbal pharmacopoeia

throughout the United States and Europe because of its anxiolytic properties (Gautz *et al*., 2006). Nevertheless, in 2007, BfArM cancelled the admission with an exception for homeopathic preparations. Swiss (Swissmedic), France (AFSSAPS), the Netherlands (CBG), and Great Britain (MHRA) followed came to the same decision even though other countries did not.

Grape skins: Ferrer-Gellego *et al*. (2011) analyzed phenolic compounds in grape skins and intact grapes during ripening. They established modified partial least square (mPLS) models using 56 samples. Anthocyanins, phenolic acids, flavonols, flavanols, and total phenolic compounds were quantified with standard RPDs ranging from 4.4 to 13.6. The same workgroup also worked on calibration models to determine flavanols in grape seeds (Ferrer-Gallego *et al*., 2010). Qualitatively, the possibility to discriminate between possible wine yards of origin has been evaluated. Quantitatively, 14 categories of flavanols were analyzed using 50 samples and applying 5–7 PLS factors resulting in RPDs from 2.8 to 10.7.

Blueberries: Sinelli *et al*. (2008) developed a system to determine total soluble solid (TSS), total phenols, total flavonoids, total anthocyanins, and ascorbatein blueberries (*Vaccinium corymbosum*) applying NIRS and MIR spectroscopy comparing cross validation and test-set validation, respectively.

Tocopherols: Szłyk, Szydłowska-Czerniak, and Kowalczyk-Marzec (2005) worked on a NIRS method to evaluate the natural α-tocopherol content in vegetable oils after extraction with ethanol. PLS calibrations for eight different edible oils were calculated. Precision and accuracy of the NIR method was comparable to the HPLC reference method.

Honeybush: A NIRS method for the quantification of mangiferin and hesperidin in dried green honeybush (*Cyclopia genistoides*) was established by Joubert, Manley, and Botha (2006a,b). PLS calibrations were calculated using 160 samples with four and six PLS factors resulting in SEPs of 0.46% (range: 0.70–7.21%) for mangiferin and 0.38 % (range: 0.64–4.80%) for hesperidin. RPDs were 1.96 (mangiferin) and 1.90 (hesperidin).

Roiboos: Joubert, Manley, and Botha (2006a,b) reported a NIRS method in combination with PLS to predict total polyphenols, aspalatin, notofagin, and dihydro-chalcone in dried green roiboos (*Aspalathus linearis*) and aspalathin in water extracts. More robust models could be achieved by adding dried roiboos extract powder to some of

the samples to increase the aspalathin and nothofagin contents.

Magnolia officinalis: Yu, Tong, and Huang (2007) reported about a NIRS method for quantification of phenolic compounds in *Magnolia officinalis*. PLS, mPLS, and PCR algorithms were compared with mPLS performing best with reaching correlation coefficients of 0.97 for the calibration and 0.95 for validation.

Rice grain: The applicability of NIRS in combination with multivariate data analysis to determine total phenolic and flavonoid contents, as well as antioxidant capacity of rice grain was discussed by Zhang *et al*. (2008). The PLS and mPLS algorithms were compared delivering similar low SEP and correlation coefficients above 0.84 for total phenolics and antioxidant capacity. However, the NIRS models to predict the flavonoid content were not successful with correlation coefficients < 0.4.

Snow lotus: Chen, Jiang, and Zhao (2010b) applied NIRS to measure the total flavone content in snow lotus (*Saussurea involucrata*) using interval partial least square with genetic algorithm (iPLS-GA) to select the efficient spectral regions and variables.

Bamboo Leaves: Flavonoids and phenolic acids in extracts of bamboo leaves were determined by Lu *et al*. (2011) applying NIRS in combination with PLS and least squares SVM.

Honghua oil: Wu *et al*. (2008) reported about MIR and NIRS methods for quantitative analysis of the three marker components α-pinene, methyl salicylate, and eugenol to assess the quality of Honghua oil, a TCM oil preparation, consisting of several plant essential oils.

Alfalfa: González-Martín *et al*. (2006) applied NIRS to determine tocopherols in alfalfa. Using mPLS calibration models for 60 fresh and dehydrated samples, SECVs of 0.37 mg/100 g for α-tocopherol (range: 0.55–5.16) and 0.027 for $(\beta + \gamma)$-tocopherol (range: 0.07–0.48) were achieved.

Radix puerariae: Puerarin, daidzin, and total flavonoids in *Radix puerariae* were determined by Lau *et al*. (2009) calculating PLS regression models. Correlation coefficients in the range 0.939–0.970 were reached using a maximum of five PLS factors. For qualitative analysis, linear discriminant analysis (LDA) in combination with SIMCA was applied to discriminate between the two species of *Radix puerariae*, *Pueraria lobata* and *Pueraria thomsonii*.

Ginkgo biloba: Zhou *et al*. (2007) reported about NIRS in combination with iPLS to determine
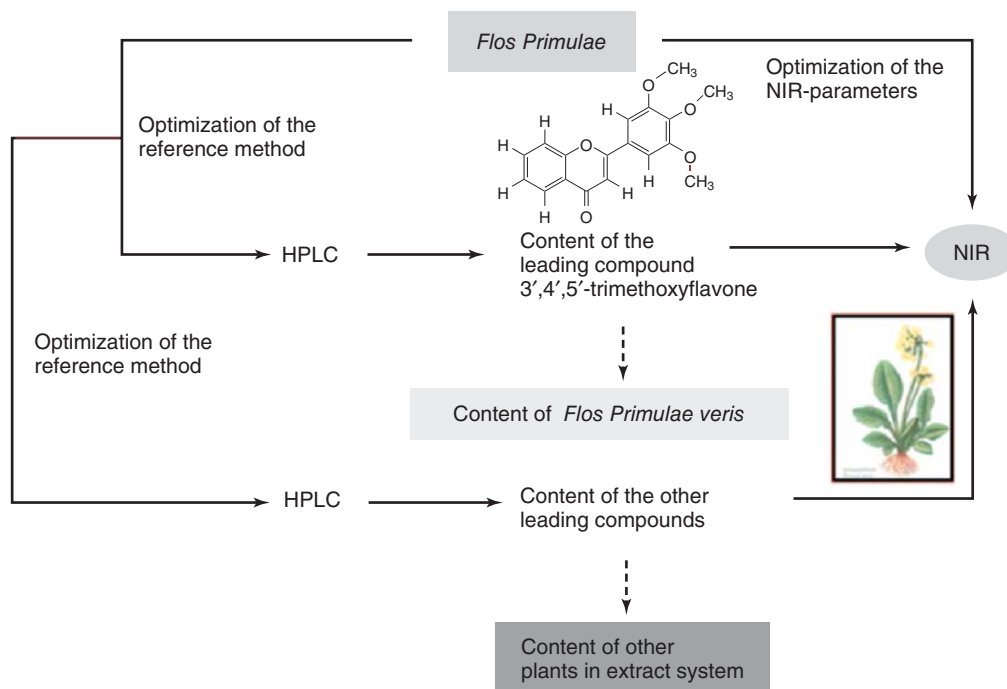
Figure 8   Strategy for calibrating the NIR spectroscopic system for the quantitative analysis of Flos *Primulae veris*.

quercetin in extracts of *Ginkgo biloba*. The calibration was calculated using 99 samples and 15 PLS factors. The RMSEP is 1.40% with a quercetin content range 4.62–7.33%.

Cannabis: Wilson and Heinrich (2006) reported about a NIRS method to discriminate between tetrahydrocannabinol (THC)-rich and hemp forms of *Cannabis*. Using PCA "high" and "low" THC content samples could be separated.

Primula: Huck *et al*. (1999) developed a method for controlling the flavonoid content in *Flos Primulae veris*, which is used as an expectorant related to its anti-inflammatory properties for the treatment of sinusitis. In Figure 8, a general strategy scheme for NIR-analysis of compounds in liquid plant extracts is shown. For the control of the *Primulae veris* Flos content, the leading compound 3′,4′,5′-trimethoxyflavone was determined by reversed-phase liquid chromatography (RP-LC) as a reference method. Recording the NIR-spectrum (Figure 9) and calculation of its first derivative spectrum allowed identifying characteristic absorption bands. The most intensive band in the spectrum belonged to the vibration of the second overtone of the carbonyl group ($5352 \, \text{cm}^{-1}$), followed by
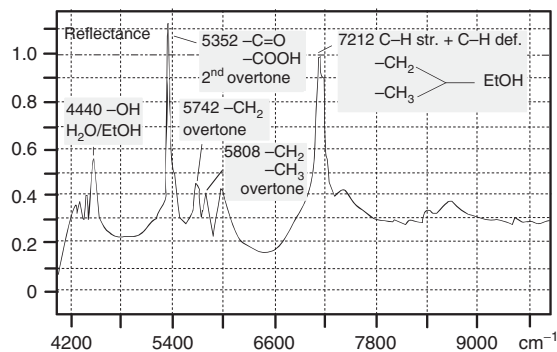
Figure 9   First derivative of the NIR spectrum with characteristic vibrations.

the C–H stretch and C–H deformation vibration of ethanol ($7212 \, \text{cm}^{-1}$), the-OH vibration of water and ethanol ($4440 \, \text{cm}^{-1}$), the $-CH_2$ overtone ($5742 \, \text{cm}^{-1}$), and the $-CH_2-/-CH_3$ overtone ($5808 \, \text{cm}^{-1}$). All recorded spectra were normalized to allow minimizing baseline shifts and transformed to their first derivative before calculating the PLS model. The high robustness of the NIRS model is demonstrated in the similarity of the results for SEE
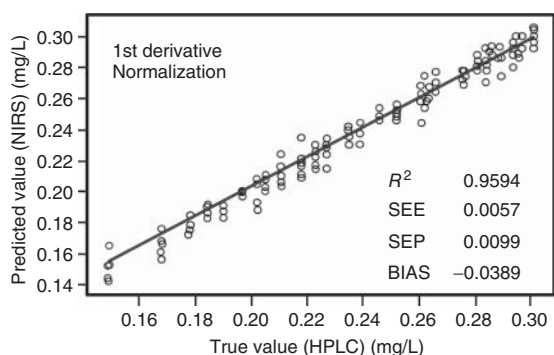
Figure 10 Calibration curve for 3′,4′,5′-trimethoxyflavone. Correlation between LC and NIRS.

Table 2 Validation and results of real samples for the quality control of the leading compound 3′,4′,5′-trimethoxyflavone in a plant extract.[a]

Ch Charge	HPLC value ($ng\,\mu L^{-1}$)	NIR value ($ng\,\mu L^{-1}$)	% H_2O	% EtOH
34	0.219	0.219	80.38	15.7
22	0.222	0.210	80.83	15.2
11	0.197	0.230	80.25	15.7
2	0.210	0.186	81.33	14.5
42	0.245	0.193	79.69	15.7

[a] *Source*: Mattle *et al*. (2010)

(0.0057 mg L^{-1}) and SEP (0.0099 mg L^{-1}). Accuracy is expressed in the bias. Its value is −3.89% with respect to the mean. Finally, a correlation coefficient of 0.95421 for the calibration curve of NIR values against LC values was established (Figure 10). The ethanol/water ratio was controlled simultaneously

with the same system with a correlation coefficient of 0.99530 for water (reference method: Karl-Fischer titration), and a coefficient of 0.99701 for ethanol (reference method: GC). Validation and results of real samples showed that the robustness and reproducibility of the NIRS model for the determination of the 3′,4′,5′-trimethoxyflavone, water, and ethanol content are high (Table 2). For the identification of *Flos Primulae veris* and quantitation of the
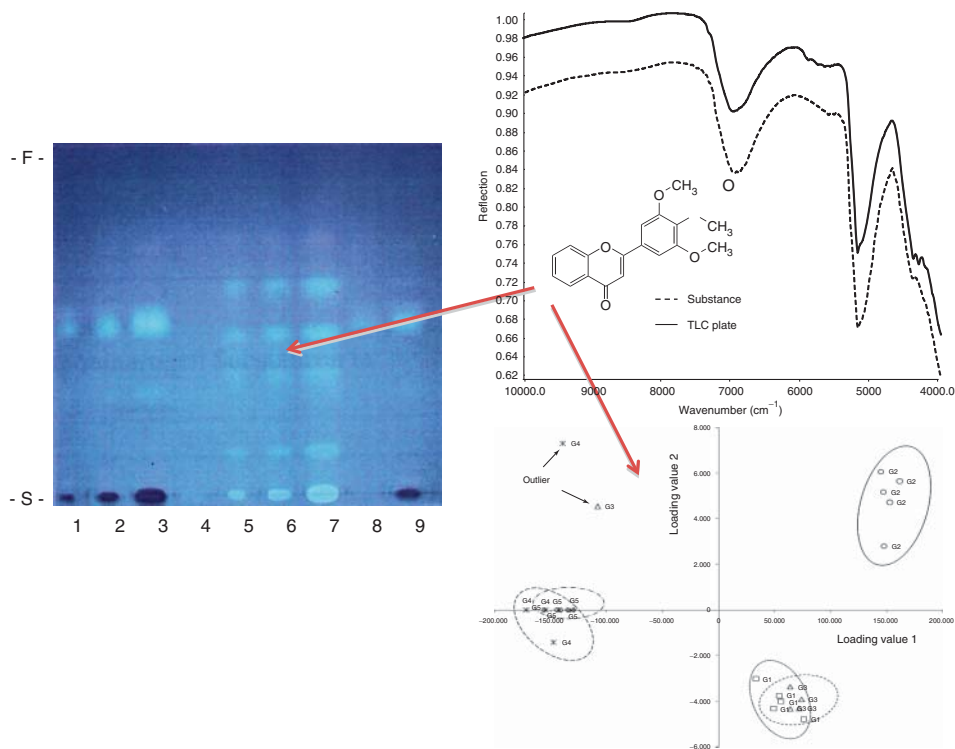
Figure 11 TLC of *Flos Primulae veris* and *Radix Primulae veris* hyphenated to near-infrared spectroscopy (TLC-NIR). Stationary phase: SiO$_2$ 60 F$_{254}$ 20×20, thin layer thickness: 1 mm; mobile phase: chloroform/acetone = 55 : 10; volume: 100 µL; detection: UV 365 nm 1–3 = *Flos Primulae* (butanol), 4 = DC-isolated substance, 5–7 = *Radix Primulae*, and 8–9 = purified extract of *Flos Primulae*.

leading compound, NIR spectroscopy was applied as a detector in TLC (Figure 11) (Mattle *et al.*, 2010).

St. John's wort: St. John's wort extract is used for the treatment of skin injuries, burns, and neuralgia, for its bacteriostatic and bactericide activities, and as a treatment for mild to moderate depression (Chiou *et al.*, 2001). It is still unclear how and why St. John's wort extract works as an antidepressant (even if it is known that the extract acts as a mild monoamine oxidase inhibitor and a strong serotonine reuptake inhibitor (Butterweck, 2003). Both hypericin and hyperforin play an important role as standards in the phytopharmaceutical industry (Barnes, Anderson, and Phillipson, 2001).

For the investigation of St. John's wort and its ingredients, different analytical procedures have been established including infrared imaging spectroscopy (Huck-Pezzei *et al.*, 2012 and 2013). UV-spectroscopy, fluorescence microscopy, TLC, LC, LC coupled to mass spectrometry (LC-MS), and CE (Huck *et al.*, 2006). For the analysis in human body fluids, tailored sample pretreatment procedures including liquid–liquid extraction (LLE) and solid-phase extraction (SPE) have been established and described in details in the literature (Pirker *et al.*, 2002). Nevertheless, all these methods are extremely time-consuming and peak-tailing effects in LC make quantitative determinations difficult. In contrast, NIR offers a fast alternative for the simultaneous quantitation. Before spectroscopic analysis via NIRS, a reference method based on LC, LC-MS, and CE was established (Huck *et al.*, 2006). In the following, 320 spectra of 80 extracts were recorded in transflection mode using light fiber optics over a wavelength range from 4008 to 9996 cm^{-1} with a resolution of 12 cm^{-1} at 23 °C and an optical pathway of 1 mm. Calculation of the first derivative spectra allowed identification of characteristic absorption bands. The most intensive band in the spectrum belonged to the vibration of the second overtone of the carbonyl group (5352 cm^{-1}),

followed by C–H stretch and C–H deformation vibration, the –OH vibration (4440 cm^{-1}), and the –CH_2 overtone (5742 cm^{-1}). Five primary factors were necessary to reach the best calibration equation. The robustness of the established NIRS model is high, which is demonstrated in similarity of results for SEE and SEP: 0.52 and 0.50 $\mu g\,mL^{-1}$ and 0.64 and 0.71 $\mu g\,mL^{-1}$ for hypericin and hyperforin, respectively (Table 3). Results show the possibility for phytopharmaceutical industry to replace LC method, usually applied to determine hypericin and hyperforine in the routine analysis, with NIR method, guaranteeing a high degree of robustness and reproducibility.

Scutellariae Radix: Scutellariae Radix, also known as *Huang-Qin*, is of high pharmacological interest because of the contained flavonoids. About 30 flavonoids were identified and quantified in *Radix Scutellariae*, using different analytical techniques such as CE (Liu and Sheu, 1994), GC, TLC, ion-pair high performance liquid chromatography (HPLC), HPLC with UV spectroscopy, high-speed counter-current chromatography (HSCCC), high performance liquid chromatography (HPLC), and high-performance liquid chromatography (HPLC) coupled to MS (Horvath, Martos, and Saxena, 2005). The main flavonoids contained in the plant are wogonin, baicalein, and baicalin; the latter found in higher amount. Several pharmacological effects of baicalin and its aglycone baicalein are reported in the literature (Motoo and Sawabu, 1994) including inhibition of cancer cell growth or induction of apoptosis in breast and prostatic cell lines, anti-inflammatory, antioxidant and free radical scavenging, antiviral (HIV), and anti-SARS coronavirus effects (Cushnie and Lamb, 2005).

Huang *et al.* (2009) determined baicalin and total flavonoids in Radix scutellariae measuring 61 samples in diffuse reflection mode with baicalin contents ranging from 12.24% to 21.34%, and

Table 3 Calibration and prediction results obtained for 80 St. John's wort extract samples by applying NIRS in the fiber optics transflectance mode.[a]

Compound	Reference		NIRS			
	Measurements		Calibration			
	Range ($\mu g\,mL^{-1}$)	Regression equation	SEE ($\mu g\,mL^{-1}$)	SEP ($\mu g\,mL^{-1}$)	BIAS	R^2
Hypericin	5.12–49.76	$y = 0.9861x + 0.3201$	0.54	0.68	1.8×10^{-14}	0.99
Hyperforin	0.10–145.41	$y = 0.9897x + 0.3648$	0.49	0.72	4.2×10^{-14}	0.99

[a] *Source*: Huck *et al.* (2006).

total flavonoid contents ranging from 16.08% to 26.52%. PLS calibration showed correlation coefficients of 0.902 for baicalin and 0.952 for total flavonoids.

6.2.2   Essential Oils

Pepper: The content and composition of the volatile fraction in various pepper samples were determined by applying NIRS and PLSR (Schulz *et al*., 2005). In the study, the determination of pungency, terpenoids, and flavor in white and black peppercorns was reported with the conclusion that both NIRS and Raman have the potential to substitute time-consuming traditional analytical procedures.

Cinnamon: Several African essential oils were analyzed to determine adulteration by Julianai *et al*. (2006). In cinnamon (*Cinnamomum zeylanicum*) and clove (*Syzygium aromaticum*) essential oils, which have similar compositions, 23 components (representing 97.8–99.9% of the oil) were quantified; in *Cinnamomum camphora*, 20 components; and in *Ravensara aromatica* and *Lippia multiflora*, 26 components. All squared correlation coefficients showed values above 0.985.

Honghua oil: Honghua oil, a traditional Chinese medicinal oil, is a mixture of several plant essential oils. Gas chromatographic (GC) investigation of 48 commercially available oils was carried out to establish PLS calibrations for the three marker components α-pinene, methyl salicylate, and eugenol with SEP values of 1.55%, 0.957%, and 0.389%, respectively (Wu *et al*., 2008).

6.2.3   Ginsenosides

Yap *et al*. performed simultaneous quantification of ginsenosides Rb1, Rb2, Rc, Re, Rd, Rg1, Ro, m-Rb1-, m-Rb2, m-Rd, and m-Rc in American ginseng. Among the calibration equations for the 11 individual ginsenosides, those of RB1, Re, and m-Rb1 showed the lowest relative standard deviation using HPLC as a reference method (Yap *et al*., 2005).

6.2.4   Volatile Compounds

The usability of NIRS for the prediction of volatile compounds employing GC-MS as a reference

method related to the aroma in Riesling wines was reported by Cozzolino *et al*. (2006).

6.2.5   Glycoside Compounds

Buckwheat: Tartary buckwheat (*Fagopyrum tartaricum*) was analyzed by Yang & Ren (2008) to quantify the rutin and D-chiro-inositol (DCI) content. PLS regression models delivered squared regression coefficients of 0.76 for rutin (8 PCs) and 0.86 for DCI (9 PCs).

6.2.6   Glucosinolates

Brassica: Several authors reported on the determination of glucosinolates in *Brassica* spp.. (Cozzolino, 2009) including *Brassica napus* (Bala and Singh, 2013), *B. pabularia, B. olearacea*, and *B. juncea* (Font *et al*., 2005a).

Cabbage: The potential of NIR spectroscopy for screening the total glucosinolates, gluconapin, gluconasturtin, and neoglucobrassacin contents of cabbage leaf cultivars from Portugal and Spain was reported (Font *et al*., 2005b).

Indian mustard: About 2700 winter Indian mustard seeds were analyzed using mPLS as regression method. The SEP values reported were $15.65 \, \text{mmol g}^{-1}$ of dry weight for glucosinolates (Font *et al*., 2004).

7   REGULATORY ISSUES

The European Medicine Agency published in 2012 guidelines on the use of NIRS by the pharmaceutical industry and the data requirements for new submissions and variations (EMEA/CHMP/CVMP/ QWP/17760/2009 Rev2; http://www.ema.europa.eu/ docs/en_GB/document_library/Scientific_guideline/ 2012/02/WC500122769.pdf). This guideline describes the regulatory requirements for marketing authorization applications and variation applications submitted for medicinal products for human or veterinary use, which include the use of NIRS. NIRS is described in the European Pharmacopoeia; however, a single reference to the Ph.Eur. General chapter on NIR spectroscopy (Ph.Eur. 2.2.40) as a

sole description for the NIRS procedure is insufficient to support the use of such a procedure in marketing authorization applications or variation submissions.

This guideline outlines the requirements for applications in which NIRS is used for qualitative and quantitative analyses or where it is used as a PAT for monitoring and controlling the synthesis of drug substances and the manufacturing processes of finished products. Approaches other than those described in this guidance may be used, if appropriately explained and justified. The chemometric principles described within this guideline may also be applicable to other analytical techniques.

In the European Pharmacopoeia 5, Chapter 2.2.40, a general description of NIRS is included along with several distinct recommendations on how to record corresponding spectra and establish qualitative and quantitative models. This includes more or less the fundamental principles, and in case of further interest, the reader is referred to *the literature* (http://lib.njutcm.edu.cn/yaodian/ep/EP5.0/02_ methods_of_analysis/2.2._physical_and_physico-chemical_methods/2.2.40.%20Near-infrared%20 spectrophotometry.pdf).

8 BENEFITS, LIMITATIONS, AND CONCLUSION

It was shown that NIRS offers a huge potential for the qualitative and quantitative analytical characterizations of different medicinal plants and their constituents deriving from manifold sources. Chemical parameters can be analyzed simultaneously with physical parameters. The main advantages of this technique over the traditional chemical and chromatographic methods are the rapidity and the ease of use in routine operations. Moreover, NIR is a non-destructive technique, requiring no or only minimal sample preparation. Nevertheless, for the quantitative analyses, a calibration model must be established with known analyte concentrations obtained by suitable reference methods. Although there have several instrumental advances been achieved, including, for example, miniaturization, the complexity of the spectrum in the NIR region and the chemometrical treatment afterward are still understood as a black box for many. Here, the formal education in both NIR spectroscopy and chemometrics will be of importance for the future.

Finally, there is no doubt that NIRS is a very attractive technique with a bright future in the arena of natural product analysis.

REFERENCES

Bala, M. and Singh, M. (2013) *Ind. Crops Prod.*, **42**, 357–362, DOI: 10.1016/j.indcrop.2012.06.014.

Barnes, J., Anderson, L. A. and Phillipson, D. J. (2001) *J. Pharm. Pharmacol.*, **53**(5), 583–600, DOI: 10.1211/0022357011775910.

Barton, F. (2002) *Spectrosc. Eur.*, **14**(1), 12–18 Retrieved from, http://karin.fq.uh.cu/~cnv1/qf/docencia/pregrado/estruc_2/curso_07_08/ir_07_08/IRcercano.pdf.

Bittner, L., Schönbichler, S., Bonn, G., *et al.* (2013) *Curr. Anal. Chem.*, **9**(3), 417–423.

Blanco, M. and Villarroya, I. (2002) *TrAC Trends Anal. Chem.*, **21**(4), 240–250, DOI: 10.1016/S0165-9936(02)00404-1.

Blanco, M., Coello, J., Iturriaga, H., *et al.* (1998) *Analyst*, **123**(8), 135R–150R, DOI: 10.1039/a802531b.

Butterweck, V. (2003) *CNS Drugs*, **17**(8), 539–562, DOI: 10.2165/00023210-200317080-00001.

Chen, Q., Zhao, J., Fang, C. H., *et al.* (2007) *Spectrochim. Acta A Mol. Biomol. Spectrosc.*, **66**(3), 568–574, DOI: 10.1016/j.saa.2006.03.038.

Chen, Y., Xie, M.-Y., Yan, Y., *et al.* (2008a) *Anal. Chim. Acta*, **618**(2), 121–130, DOI: 10.1016/j.aca.2008.04.055.

Chen, Q., Zhao, J., Liu, M., *et al.* (2008b) *J. Pharm. Biomed. Anal.*, **46**(3), 568–573.

Chen, Q., Zhao, J., Liu, M., *et al.* (2010a) *J. Pharmaceut. Biomed. Anal.*, **46**(3), 568–573 Retrieved from, http://cat.inist.fr/?aModele=afficheN&cpsidt=20066176.

Chen, Q., Jiang, P. and Zhao, J. (2010b) *Spectrochim. Acta A Mol. Biomol. Spectrosc.*, **76**(1), 50–55, DOI: 10.1016/j.saa.2010.02.045.

Chen, Q., Guo, Z., Zhao, J., *et al.* (2012) *J. Pharm. Biomed. Anal.*, **60**(23), 92–97.

Chiou, W., Jeong, H., Wu, T., *et al.* (2001) *Clin. Pharmacol. Therap.*, **70**(4), 305–310, DOI: 10.1016/S0009-9236(01)00127-8.

Cordella, C., Moussa, I., Martel, A.-C., *et al.* (2002) *J. Agric. Food Chem.*, **50**(7), 1751–1764, DOI: 10.1021/jf011096z.

Cozzolino, D. (2009) *Planta Med.*, **75**(7), 746–756, DOI: 10.1055/s-0028-1112220.

Cozzolino, D., Kwiatkowski, M., Parker, M., *et al*. (2004) *Anal. Chim. Acta*, **513**(1), 73–80, DOI: 10.1016/j.aca.2003.08.066.

Cozzolino, D., Smyth, H. E., Lattey, K. A., *et al*. (2006) *Anal. Chim. Acta*, **563**(1–2), 319–324, DOI: 10.1016/j.aca.2005.11.008.

Cushnie, T. P. T. and Lamb, A. J. (2005) *Int. J. Antimicrob. Agents*, **26**(5), 343–356, DOI: 10.1016/j.ijantimicag.2005.09.002.

Faix, O. (1992) in *Methods in Lignin Chemistry*, eds. S. Y. Lin and C. W. Dence, Springer Berlin Heidelberg, Berlin, Heidelberg, DOI: 10.1007/978-3-642-74065-7.

Ferrer-Gallego, R., Hernández-Hierro, J. M., Rivas-Gonzalo, J. C., *et al*. (2010) *Talanta*, **82**(5), 1778–1783, DOI: 10.1016/j.talanta.2010.07.063.

Ferrer-Gallego, R., Hernández-Hierro, J. M., Rivas-Gonzalo, J. C., *et al*. (2011) *LWT – Food Sci. Technol.*, **44**(4), 847–853, DOI: 10.1016/j.lwt.2010.12.001.

Font, R., Del Río, M., Fernández-Martínez, J. M., *et al*. (2004) *J. Agri. Food Chem.*, **52**(11), 3563–9, DOI: 10.1021/jf0307649.

Font, R., Del Río-Celestino, M., Cartea, E., *et al*. (2005a) *Phytochemistry*, **66**(2), 175–185, DOI: 10.1016/j.phytochem.2004.11.011.

Font, R., Del Rio-Clestino, M., Rosa, E., *et al*. (2005b) *J. Agric. Sci.*, **143**, 65–73 Retrieved from, http://journals.cambridge.org/download.php?file=%2FAGS%2FAGS143_01%2FS0021859605004806a.pdf&code=8b5633117fe45289a69d566e33aeaf5e.

Garde-Cerdán, T., Lorenzo, C., Alonso, G. L., *et al*. (2010) *Food Chem.*, **119**(2), 823–828, DOI: 10.1016/j.foodchem.2009.07.026.

Garde-Cerdán, T., Lorenzo, C., Zalacain, A., *et al*. (2012) *LWT – Food Sci. Technol.*, **46**(2), 401–405, DOI: 10.1016/j.lwt.2011.12.012.

Gautz, L. D., Kaufusi, P., Jackson, M. C., *et al*. (2006) *J. Agri. Food Chem.*, **54**(17), 6147–6152, DOI: 10.1021/jf060964v.

González-Martín, I., Hernández-Hierro, J. M., M. Bustamante-Rangel, *et al*. (2006) *Anal. Bioanal. Chem.*, **386**(5), 1553–1558, DOI: 10.1007/s00216-006-0666-0.

Guggenbichler, W., Huck, C., Kobler, A., *et al*. (2006) *J. Food, Agri. Environ.*, **4**(2), 98–106.

He, Y., Li, X. and Deng, X. (2007) *J. Food Eng.*, **79**(4), 1238–1242, DOI: 10.1016/j.jfoodeng.2006.04.042.

Herschel, W. (1800) *Philos. Trans. Royal Soc. London*, **90**, 284–292.

Horvath, C. R., Martos, P. A. and Saxena, P. K. (2005) *J. Chromatogr. A*, **1062**(2), 199–207, DOI: 10.1016/j.chroma.2004.11.030.

Huang, Q., Pan, R., Wei, J., *et al*. (2009) *Spectrosc. Spec. Anal.*, **29**(9), 2425.

Huck, C. W., Huber, C. G., Lagoja, I. M., *et al*. (1999) *Planta Med.*, **65**(5), 491 Retrieved from, http://cat.inist.fr/?aModele=afficheN&cpsidt=1833596.

Huck, C. W., Guggenbichler, W. and Bonn, G. K. (2005) *Anal. Chim. Acta*, **538**(1–2), 195–203, DOI: 10.1016/j.aca.2005.01.064.

Huck, C. W., Abel, G., Popp, M., *et al*. (2006) *Anal. Chim. Acta*, **580**(2), 223–230, DOI: 10.1016/j.aca.2006.07.062.

Huck-Pezzei, V. A., Pallua, J. D., Pezzei, C., *et al*. (2012) *Anal. Bioanal. Chem.*, **404**, 1771–1778.

Huck-Pezzei, V. A., Bittner, L. K., Pallua, J. D., *et al*. (2013) *Anal. Meth.*, **5**(3), 616, DOI: 10.1039/c2ay26030a.

Joubert, E., Manley, M. and Botha, M. (2006a) *Phytochem. Anal.*, **19**(2), 169–178, DOI: 10.1002/pca.1033.

Joubert, E., Manley, M. and Botha, M. (2006b) *J. Agri. Food Chem.*, **54**(15), 5279–5283, DOI: 10.1021/jf060617l.

Juliani, H. R., Kapteyn, J., Jones, D., *et al*. (2006) *Phytochem. Anal.*, **17**(2), 121–128, DOI: 10.1002/pca.895.

Krüger, H. and Schulz, H. (2007) *Stewart Postharvest Rev.*, **3**(4), 12, http://dx.doi.org/10.2212/spr.2007.4.4.

Lau, C.-C., Chan, C.-O., Chau, F.-T., *et al*. (2009) *J. Chromatogr. A*, **1216**(11), 2130–2135, DOI: 10.1016/j.chroma.2008.12.089.

Li, W., Xing, L., Fang, L., *et al*. (2010) *J. Pharmaceut. Biomed. Anal.*, **53**(3), 350–358, DOI: 10.1016/j.jpba.2010.04.011.

Liu, Y.-M. and Sheu, S.-J. (1994) *Anal. Chim. Acta*, **288**(3), 221–226, DOI: 10.1016/0003-2670(94)80134-7.

Lu, B. *et al*., (2011) *Afr. J. Biotechnol.*, **10**(42), 8448–8845 Retrieved from, http://www.ajol.info/index.php/ajb/article/view/95431.

Luypaert, J., Zhang, M. H. and Massart, D. L. (2003) *Anal. Chim. Acta*, **478**(2), 303–312, DOI: 10.1016/S0003-2670(02)01509-X.

Mattle, C., Heigl, N., Abel, G., *et al*. (2010) *J. Planar Chromatogr.*, **23**(5), 348–352 Retrieved from, http://www.akademiai.com/content/c30524vh26740661/.

McClure, W. F. (2003) *J. Near infrared Spectrosc.*, **11**(6), 487–518 Retrieved from, http://cat.inist.fr/?aModele=afficheN&cpsidt=15634296.

Motoo, Y. and Sawabu, N. (1994) *Cancer Lett.*, **86**(1), 91–95, DOI: 10.1016/0304-3835(94)90184-8.

Nicolaï, B. M., Beullens, K., Bobelyn, E., *et al*. (2007) *Postharvest Biol. Technol.*, **46**(2), 99–118, DOI: 10.1016/j.postharvbio.2007.06.024.

Osborne, B. G., Fearn, T. and Hindle, P. H. (1993) *Longman Scientific and Technical* Retrieved from, http://www.cabdirect.org/abstracts/19931464035.html;jsessionid=76ACF38E53B94494017C4E31252EF572.

Padilla, E., Ruiz, E., Redondo, S., *et al*. (2005) *Eur. J. Pharmacol.*, **517**(1–2), 84–91, DOI: 10.1016/j.ejphar.2005.04.044.

Pallua, J. D., Recheis, W., Pöder, R., *et al*. (2011) *Analyst*, DOI: 10.1039/c1an15615b.

Perez, R. M. (2008) http://informahealthcare.com/doi/abs/10.1076/phbi.41.2.107.14240?journalCode=phb Retrieved from .

Perron, N. R. and Brumaghim, J. L. (2009) *Cell Biochem. Biophys.*, **53**(2), 75–100, DOI: 10.1007/s12013-009-9043-x.

Pirker, R., Huck, C. and Bonn, G. (2002) *J. Chromatogr. B*, **777**(1–2), 147–153, DOI: 10.1016/S1570-0232(02)00080-6.

Rauha, J.-P., Remes, S., Heinonen, M., *et al*. (2000) *Int. J. Food Microbiol.*, **56**(1), 3–12, DOI: 10.1016/S0168-1605(00)00218-X.

Salzer, R. and Siesler, H. W. (2009) *Infrared and Raman Spectroscopic Imaging*, Wiley-VCH Retrieved from, http://www.amazon.com/Infrared-Spectroscopic-Imaging-Reiner-Salzer/dp/352731993X.

Schulz, H. (2004) Analysis of coffee, tea, cocoa, tobacco, spices, medicinal and aromatic plants, and related products, in *Near-Infrared Spectroscopy in Agriculture*, (ed. C. W. J. R. I. J. Roberts), American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, vol. **agronomymo**, pp. 345–376, DOI: 10.2134/agronmonogr44.c13.

Schulz, H., Baranska, M., *et al*. (2005) *J. Agri. Food Chem.*, **53**(9), 3358–3363, DOI: 10.1021/jf048137m.

Siebert, K. J. (2001) *J. Am. Soc. Brewing Chem.*, **59**(4), 147–156 Retrieved from, http://cat.inist.fr/?aModele=afficheN&cpsidt=14128454.

Sinelli, N., Spinardi, A., Di Egidio, V., *et al*. (2008) *Postharvest Biol. Technol.*, **50**(1), 31–36, DOI: 10.1016/j.postharvbio.2008.03.013.

Stecher, G., Huck, C. W., Stöggl, W. M., *et al*. (2003) *TrAC Trends Anal. Chem.*, **22**(1), 1–14, DOI: 10.1016/S0165-9936(03)00108-0.

Szłyk, E., Szydłowska-Czerniak, A. and Kowalczyk-Marzec, A. (2005) *J. Agri. Food Chem.*, **53**(18), 6980–6987, DOI: 10.1021/jf050672e.

Williams, P. and Norris, K. (1987) *Near-Infrared Technology in the Agricultural and Food Industries*, American Association of Cereal Chemists, Inc. Retrieved from, http://www.cabdirect.org/abstracts/19892442443.html.

Williams, P. C., Norris, K. H. and Sobering, D. C. (1985) *J. Agric. Food Chem.*, **33**(2), 239–244, DOI: 10.1021/jf00062a021.

Wilson, N. and Heinrich, M. (2006) *Planta Med.*, **72**(11), P_260, DOI: 10.1055/s-2006-950060.

Wu, Y.-W., Sun, S.-Q., Zhou, Q., *et al*. (2008) *J. Pharmaceut. Biomed. Anal.*, **46**(3), 498–504, DOI: 10.1016/j.jpba.2007.11.021.

Yang, N. and Ren, G. (2008) *J. Agri. Food Chem.*, **56**(3), 761–764, DOI: 10.1021/jf072453u.

Yap, K. Y.-L., Chan, S. Y., Weng Chan, Y., *et al*. (2005) *Assay Drug Develop. Technol.*, **3**(6), 683–699, DOI: 10.1089/adt.2005.3.683.

Yu, C.-Y., Tong, Z.-k. and Huang, H.-H. (2007) *J. Zhejiang Forest. College*, **05**, http://en.cnki.com.cn/Article_en/CJFDTOTAL-ZJLX200705007.htm

Zalacain, A., Ordoudi, S. A., Díaz-Plaza, E. M., *et al*. (2005) *J. Agri. Food Chem.*, **53**(24), 9337–9341, DOI: 10.1021/jf050846s.

Zhang, M. H., Luypaert, J., Fernández Pierna, J. A., *et al*. (2004) *Talanta*, **62**(1), 25–35, DOI: 10.1016/S0039-9140(03)00397-7.

Zhang, C., Shen, Y., Chen, J., *et al*. (2008) *J. Agri. Food Chem.*, **56**(18), 8268–8272, DOI: 10.1021/jf801830z.

Zhou, X., Xiang, B., Wang, Z., *et al*. (2007) *Anal. Lett.*, **40**(18), 3383–3391, DOI: 10.1080/00032710701689081.

Headspace Sampling and Gas Chromatography: A Successful Combination to Study the Composition of a Plant Volatile Fraction

Barbara Sgorbini, Cecilia Cagliero, Chiara Cordero, Erica Liberto, Patrizia Rubiolo and Carlo Bicchi

Dipartimento di Scienza e Tecnologia del Farmaco, Università di Torino, Torino, Italy

1 INTRODUCTION

GC is a fundamental analytical technique for analyzing plant metabolites, thanks to its versatility, efficiency, and sensitivity. GC is the technique of choice to analyze several classes of plant metabolites, containing compounds that are in the vapor phase or that can be vaporized at a suitable temperature.

In chemical terms, the *volatile fraction* of a matrix of vegetable origin can in general be defined as a mixture of volatiles that can be sampled because of their ability to vaporize spontaneously, and/or under suitable conditions, or by adopting appropriate techniques (Bicchi and Maffei, 2012; Rubiolo *et al*., 2010a,b). This definition, together with the fact that the volatile fraction of a plant mainly consists of homogeneous groups of medium-to-low polarity components, often with similar structures and physicochemical characteristics (e.g., monoterpenoids) leads to two techniques being the primary choices for the analysis of this fraction: (i) headspace (HS) for sample preparation, because it is easy to automate and to combine with

GC, and (ii) gas chromatography (GC) as such or combined with mass spectrometry (GC–MS) for analysis.

This chapter provides an overview of these two techniques in plant volatile analysis and discusses headspace sample preparation techniques that can be coupled directly to GC and fast GC, as such or in combination with mass spectrometry (GC–MS), qualitative and quantitative analysis, enantioselective GC with cyclodextrins derivatives as chiral selectors, the role of derivatization, and multidimensional (MD) GC.

2 HEADSPACE SAMPLING

2.1 Some Fundamental Definitions

HS is defined as an approach to sampling the gaseous or vapor phase in equilibrium with a solid or liquid matrix that aims to study its composition (Kolb and Ettre, 1997). This definition refers primarily to a *one-step gas extraction* technique, better known as

static headspace sampling (*S-HS*). S-HS composition mainly depends on analyte partition coefficients between matrix and vapor phase at equilibrium and is representative of the original matrix. An aliquot of headspace is in general online and automatically transferred to a GC or GC–MS system for analysis (Kolb and Ettre, 1997). Figure 1 reports a diagram of an equilibrium headspace system (Bicchi *et al.*, 2012).

Headspace sampling (HS) can also be achieved with a second approach, known as *dynamic headspace sampling* (*D-HS*). It usually consists of three main steps: (i) the continuous removal of the volatile components released from a matrix into the empty volume of the sealed vessel in which it is contained, by an inert gas flowing through or over it; (ii) the concentration of those components from the above-mentioned flow stream, through cryotrapping, or a solid adsorbent, a sorbent, a liquid stationary phase coated on a solid support, or a selective reagent for a specific class (or classes) of compounds; and (iii) the recovery of the collected analytes from the trap by solvent elution or thermal desorption online or off-line to a GC or GC–MS system for analysis. *D-HS* is a nonequilibrium *continuous gas extraction* process, better known as *dynamic headspace analysis* or the *purge-and-trap method* (Kolb and Ettre, 1997).
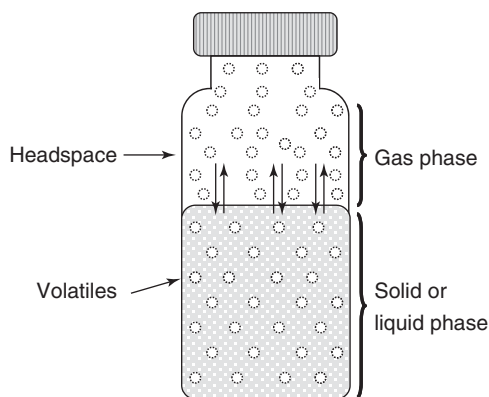
Figure 1 Diagram of an equilibrium static headspace system. (*Source*: Reprinted from Comprehensive Sampling and Sample Preparation: Analytical Techniques for Scientists, 4, Bicchi C., Cordero C., Liberto E., Sgorbini B., and Rubiolo P., Headspace Sampling in Flavor and Fragrance Field, 1–25, Copyright (2012), with permission from Elsevier.)

2.2 Headspace Sampling: A Short History and Evolution

HS is a sample preparation and clean-up technique that has "grown" in parallel to GC and that has been applied to plant volatile fraction analysis since its introduction at the end of the 1950s.

As reported by Ettre (2002) in an exhaustive review on the beginnings of headspace, the first description of HS was by Harger, Bridwell, and Raney (1939) at the Department of Biochemistry and Pharmacology of the School of Medicine of Indiana University in 1939 (of course, not in combination with GC); they applied this method to the rapid determination of alcohol in water and body fluids. However, the terms "headspace" and "headspace analysis" were first used by Stahl, Voelker, and Sullivan (1960), adapting an expression used in the food packaging industry. The first study in which HS sampling was combined with GC analysis was by Bovijn, Pirotte, and Berger (1958) who, in 1958, applied sampling of the "gaseous phase in equilibrium with a liquid phase" to monitoring hydrogen at the 1-ppb level, in the water of high pressure boilers.

D-HS analysis was introduced by Wahlroos (1963). One of its first applications concerned plant volatiles: Herout (1967) applied D-HS, using a short chromatographic precolumn filled with 8% Apiezon L coated on Chromosorb as trapping system, hydrodistillation (Clevenger apparatus), and solvent extraction (pentane) to study the scent composition of three flowers: hyacinth (*Hyacinthus orientalis*), violet (*Viola odorata*), and orchid (*Lycaste macrobulbon*). They later extended the analysis to white and violet-colored lilac flowers (*Syringa vulgaris* L. and *S. persica* L.) (Herout, Streibl, and Holasova, 1970); trapped volatiles were recovered by direct thermal desorption onto the GC column.

Fundamental contributions to the affirmation of D-HS were made by Teranishi's group in 1972, with their studies on volatile organic components present in breath and urine (Teranishi *et al.*, 1972), and by Zlatkis' group in 1973, with the introduction of Tenax as an adsorbent (Zlatkis, Lichtenstein, and Tishbee, 1973).

The success of HS was immediate, because of its simplicity, flexibility, and sensitivity, but its widespread affirmation, in particular for plant volatile routine analysis, dates only to the 1990s, when it became more widely adopted as a consequence both of the exponential increase in the demand for

controls and of the success of solvent-free sample preparation techniques, that is, those techniques in which analytes are isolated from a matrix without using a liquid solvent. The delay in the affirmation of HS was also due to the precision of the definition of the principles of S-HS and D-HS approaches that, paradoxically, slowed down the development of HS techniques. This hindrance was overcome in the early 1990s with the introduction of an additional approach, *high concentration capacity headspace techniques* (*HCC-HS*) (Bicchi, Cordero, and Rubiolo, 2004a). These techniques established a bridge between *S-HS* and *D-HS*, based on either the static or the dynamic accumulation of volatile(s) on polymers operating in sorption and/or adsorption, or, more seldom, on solvents.

2.3 Static Headspace

Since its introduction, S-HS in combination with GC has successfully been used to study the plant volatile fraction, because of its simplicity, speed, ease of automation, and versatility. It is in general applied to analyses in the low parts per million to percent range of concentrations, because of its limited sensitivity.

In S-HS sampling, the liquid or solid sample sealed in the headspace vial is equilibrated, at a given temperature for a suitable time; at equilibrium, an aliquot of the resulting vapor phase is transferred to the GC system for analysis. The vapor transfer to GC can be run by either manual or automatic suction with a (thermostated) gas-tight syringe, or with a number of systems, the best known of which are (i) the balanced-pressure sampling systems and (ii) the pressure/loop system (Kolb and Ettre, 1997).

Figure 2 reports a diagram of a S-HS-GC sampling system with sampled headspace transferred to GC by manual or automatic suction and conventional GC injection (i) or by pressurized sampling (ii) (Bicchi *et al.*, 2012).

2.4 Dynamic Headspace

D-HS is probably the vapor phase sampling approach most widely used to study the plant volatile fraction composition exhaustively, mainly because of its great flexibility in terms of both volume of "extracting" gas and number of possible trapping approaches and
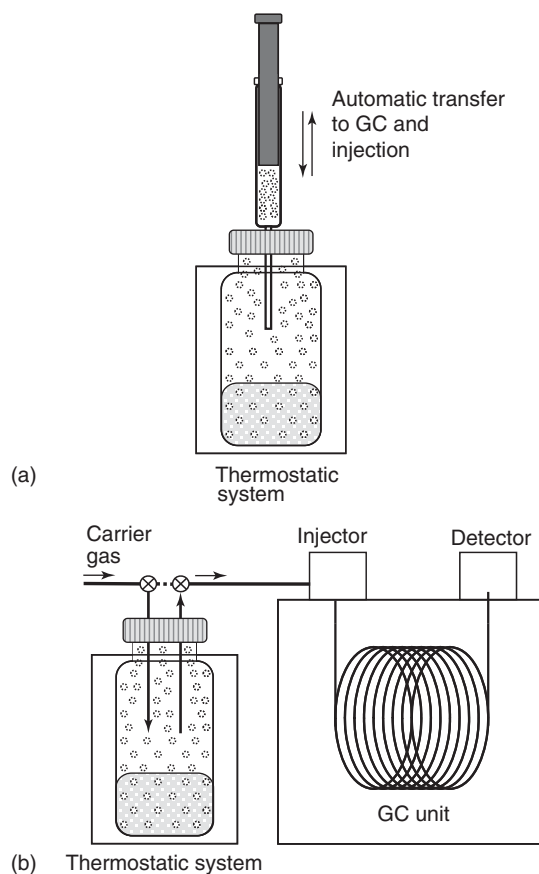
Figure 2 Diagram of an S-HS-GC sampling system with headspace sample transferred to GC by both manual or automatic suction and conventional GC injection (a) or by pressurized sampling and introduction (b). (*Source*: Reprinted from Comprehensive Sampling and Sample Preparation: Analytical Techniques for Scientists, 4, Bicchi C., Cordero C., Liberto E., Sgorbini B., and Rubiolo P., Headspace Sampling in Flavor and Fragrance Field, 1–25, Copyright (2012), with permission from Elsevier.)

materials. These factors enable recovery of the analytes under study to be optimized in function of their chemical nature, but D-HS requires more sophisticated instrumentation and sampling procedures, because of the large number of parameters to be standardized to obtain reliable sampling reproducibility and comparability. However, new technologies and collecting materials, together with increasing automation, have partially or entirely overcome these limits.

In D-HS sampling, the volatile fraction is recovered from the gaseous flow stream stripped *through* (purge-and-trap (P&T) approach) or *over* (dynamic

approach) the matrix onto a suitable trapping system (e.g., cold trap) or medium (a sorbent, an adsorbent, or a specific reagent or solvent for a specific class or classes of compounds). The sampled volatiles are usually released either online by thermal desorption or vaporization (after cryotrapping) directly to the GC or GC–MS system or, more seldom, off-line by solvent elution from the trap. Figure 3 reports a diagram of a D-HS-GC sampling and analysis (Bicchi *et al*., 2012).

A crucial step in D-HS is trapping, or better the choice of trapping mode and materials. The most widely used approach is adsorption on materials such as carbon in its various forms (carbon molecular sieves, activated charcoal, graphitized carbon black, etc.), tenax (poly-2,6-diphenyl-phenylene

(a)

(b)

Figure 3 Diagram of a D-HS-GC procedure: (a) D-HS sampling and (b) thermal desorption and GC analysis. (*Source*: Reprinted from Comprehensive Sampling and Sample Preparation: Analytical Techniques for Scientists, 4, Bicchi C., Cordero C., Liberto E., Sgorbini B., and Rubiolo P., Headspace Sampling in Flavor and Fragrance Field, 1–25, Copyright (2012), with permission from Elsevier.)

oxide), chromosorb 102 (styrene-divinylbenzene), chromosorb 104 (acrylonitrile-divinylbenzene), parapack Q (ethylvinylbenzene-divinylbenzene), or others. The advantages and limitations of adsorption are well known, the latter including analyte discrimination due to polarity, production of artifacts because of the catalytic activity of the adsorbent, and irreversible bonding onto the adsorbent surface (Baltussen, 2000). Some of these limits have more recently been overcome by sorption, an approach based on a group of highly inert and thermally stable polymeric materials (mainly polydimethylsiloxane, PDMS, and to a lesser extent polyacrylates) operating in partition. PDMS as trapping material was introduced by Grob and Habich (1985), with open tubular traps (OTT); others continued its development (Bicchi *et al*., 1989), but only in the second half of the 1990s did it become suitable for routine analysis, with the development of gum-phase extraction (GPE) thanks to the in-depth investigation on sorption fundamentals by Sandra's and Cramers' groups (Baltussen, Cramers, and Sandra, 2002). New techniques and approaches were developed on the basis of these studies, for example, stir bar sorptive extraction (SBSE) (Baltussen *et al*., 1999a), headspace sorptive extraction (HSSE) (see later) (Bicchi *et al*., 2000a; Tienpont *et al*., 2000), and equilibrium sorptive extraction (ESE) (Baltussen *et al*., 1999b).

2.5 Static and Trapped Headspace

Static and trapped headspace (S&T-HS) was introduced by Chaintreau's group to combine the advantages of S-HS and D-HS while overcoming their main limitations (Chaintreau, 2000). In S&T-HS, a moving piston evacuates the vapor phase in equilibrium with the matrix from the HS chamber by pressure and concentrates the vapor components onto an adsorbent in a cartridge through which the gas flows. The trapped analytes are then recovered by thermal desorption and online injected into a GC unit for analysis. The sensitivity depends closely on the sampled gas volume and, hence, on the cell volume.

Figure 4 reports a diagram of a cell for S&T-HS sampling (Chaintreau, 2000).

This technique has been shown to be very effective for theoretical studies, and for quantitation, either by combining the determination of the analyte air-to-liquid partition coefficients with multiple
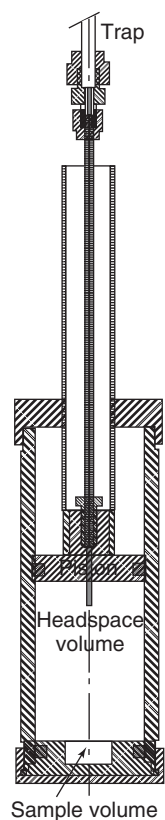
Figure 4 Diagram of a cell for S&T-HS sampling. (*Source*: A. Chaintreau (2000). Reproduced by permission of John Wiley & Sons, Ltd.)

headspace extraction (MHE) or by the standard addition (SA) procedure (see later) (Chaintreau, 2000).

2.6 High Concentration Capacity Headspace Techniques (HCC-HS)

The extension of HS sampling applications to routine use began in the early 1990s with the development of the HCC-HS approach, which was particularly favored by the success of solvent-free sample preparation techniques (Bicchi *et al.*, 2008a; Pawliszyn, 1997). HCC-HS techniques combine the advantages of S-HS and D-HS, and their success, in particular in the plant field, has mainly been related (i) to their ease of automation and combination online to the analytical instrumentation, which has favored the adoption of HS, whenever possible, instead of

time-consuming and/or polluting solvent-consuming extraction or distillation and (ii) to the new possibilities offered by sorption and adsorption, the main phenomena involved in analyte recovery onto a polymer from a liquid or vapor phases (Baltussen *et al.*, 1999b; Baltussen, 2000).

Several HCC-HS-based sampling techniques enabling the number of their possible applications to be extended have been developed as a complement to conventional S-HS and D-HS. The best known of these, together with their underlying principles, are described in brief, in particular, HS-solid-phase microextraction (HS-SPME) (Zhang and Pawliszyn, 1993), headspace liquid-phase microextraction (HS-LPME) (Tankeviciute, Kazlauskas, and Vickackaite, 2001; Theis *et al.*, 2001), in-tube sorptive extraction (INT, INCAT, and HS-SPDE) (Bicchi *et al.*, 2004b; McComb *et al.*, 1997; Musshoff *et al.*, 2002), HSSE (Bicchi *et al.*, 2000a; Tienpont *et al.*, 2000), solid-phase aroma concentrate extraction (SPACE) (Ishikawa *et al.*, 2004), and large surface area HCC-HS sampling (MESI, MME, and HS-STE) (Bicchi *et al.*, 2007a; Bruheim, Liu, and Pawliszyn, 2003; Segal *et al.*, 2000).

HCC-HS, HS-SPME, and D-HS techniques in the sample preparation of the plant volatile fraction, their advantages and limitations, have recently been reviewed by the authors' group (Bicchi and Maffei, 2012; Bicchi *et al.*, 2008b, 2012).

2.7 Headspace-Solid Phase Microextraction (HS-SPME)

The first, and today the most widely adopted, HCC-HS technique is HS-SPME, introduced by Zhang and Pawliszyn (1993) and derived from SPME, developed by Arthur and Pawliszyn (1990). In SPME, the fractions of interest in the liquid or vapor phase are accumulated onto a film of a sorbent and/or an adsorbent, coated on a retractable fused silica fiber, which is part of a stainless steel needle assembled onto a customized device. The sampled analytes are in general recovered by thermal desorption directly into the body of a conventional split/splitless GC injector.

Figure 5 reports a diagram of a HS-SPME holder and of a fused silica fiber (Bicchi *et al.*, 2012).

Figure 6 reports the HS-SPME-GC–MS pattern of a sample of American pepper fruits (*Schinus molle* L.).

SPME holder

Fused silica fiber

Polymer coating

Needle guide

Fiber

Thermostatic system
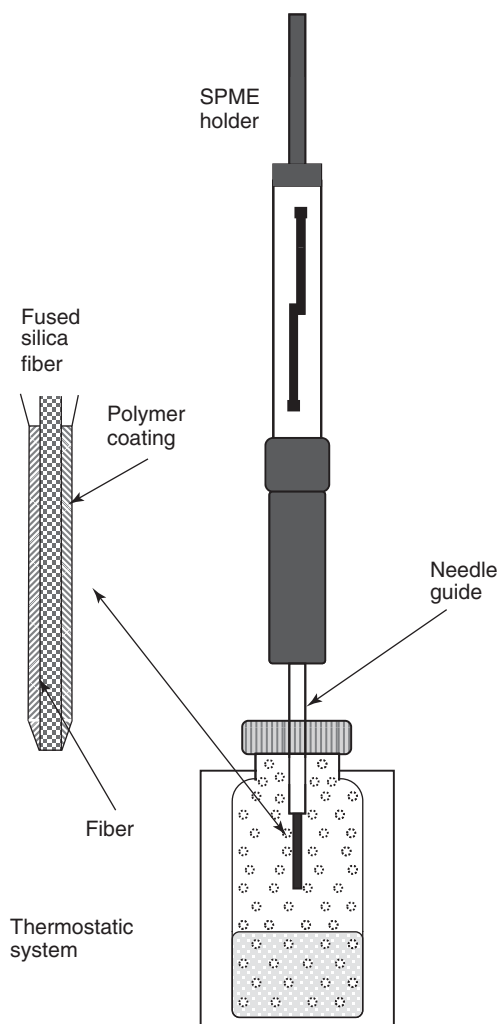
Figure 5   Diagram of a HS-SPME holder and of a fused silica fiber. (*Source*: Reprinted from Comprehensive Sampling and Sample Preparation: Analytical Techniques for Scientists, 4, Bicchi C., Cordero C., Liberto E., Sgorbini B., and Rubiolo P., Headspace Sampling in Flavor and Fragrance Field, 1–25, Copyright (2012), with permission from Elsevier.)

The same group also advanced a theory for SPME applied to HS sampling (Gorecki and Pawliszyn, 1995; Zhang and Pawliszyn, 1993), which can easily be extended to all other HCC-HS techniques. This theory hypothesizes that analyte recovery from the headspace of a matrix onto a fiber depends on two closely related but distinct equilibria: (i) the sample/HS equilibrium, responsible for the HS formation and composition (measured by its distribution coefficient, K_{hs}), and (ii) the HS/fiber

equilibrium, conditioning the analyte accumulation onto the SPME fiber (expressed by its distribution coefficient, K_{fh}). The theoretical model shows that a direct proportion is established at equilibrium between the amount of analyte accumulated by the fiber and its initial concentration in the sample matrix. However, as the time to reach equilibrium can vary from a few minutes to several hours, nonequilibrium sampling under rigorously standardized conditions of the key parameters (i.e., nature of polymeric fiber coating; coating volume, extraction temperature, sample agitation, pH and ionic strength, phase ratio (β), and extraction time) (Ai, 1997) is usually applied to assure comparable and significant results. The most used coatings are polydimethylsiloxane (PDMS), polyacrylate (PA), and composite materials such as PDMS/divinylbenzene (DVB), carbowax/PDMS (CW)/PDMS), carboxen (CAR)/PDMS, and CAR/DVB/PDMS. SPME theory, technology, evolution, and applications have periodically been updated by Pawliszyn and coworkers (Pawliszyn, 1997, 2002), whereas HS-SPME in the plant field was reviewed by Belliardo *et al.* (2006).

SPME operating in the D-HS mode was also developed and patented by Silva, Aquiar, and Augusto (2004).

2.8   Headspace Liquid-Phase Microextraction (HS-LPME)

LPME was first introduced for sampling in the liquid phase by Jeannot and Cantwell (1996) at the beginning of the 1990s; its use was later extended to S-HS by Theis *et al.* (2001) and Tankeviciute, Kazlauskas, and Vickackaite (2001). In HS-LPME, the volatile fraction is accumulated into a drop of a solvent generated at the tip of a syringe needle suspended into the HS of the investigated matrix in a sealed vial; after sampling, the drop is retracted into the needle and directly (and if possible automatically) injected into the GC injector. The solvent must produce stable drops, be compatible with GC analysis, and have a vapor pressure low enough to avoid its evaporation during sampling. The wide range of adoptable nonpolar, polar, and water-miscible solvents makes HS-LPME very flexible and it can easily be modified for specific applications and/or online or off-line combined with other extraction and distillation techniques (Ouyang, Zhao, and Pawliszyn,
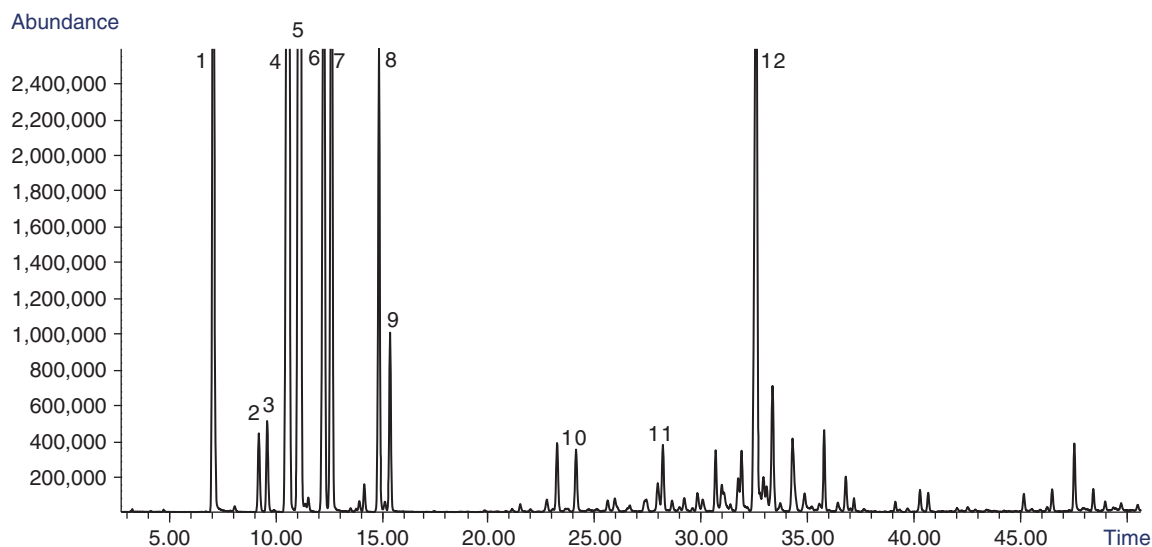
Figure 6 HS-SPME-GC–MS profile of an American pepper sample (*Schinus molle* L.). Analysis system: Agilent 6890 GC-5973 MSD system (Agilent, Little Falls, DE, USA). Sample preparation: SPME fiber: CAR/PDMS/DVB, 2 cm; sample amount: 20 mg; vial volume: 20 mL; sampling time: 30 min; sampling temperature: 60 °C. GC–MS conditions: column: MEGAWAX 20M (50 m, d_c: 0.20 mm, d_f: 0.20 μm); oven temperature: from 50 (1 min) to 230 °C (5 min) at 3 °C min^{-1}. Desorption temperature: 230 °C; desorption time: 5 min; injection mode: split, ratio 1 : 20; carrier gas: He; flow rate: 1 mL min^{-1}. List of identified compounds: (1) α-pinene, (2) β-pinene, (3) sabinene, (4) Δ-3-carene, (5) α-phellandrene, (6) limonene, (7) β-phellandrene, (8) *p*-cymene, (9) α-terpinolene, (10) α-copaene, (11) β-caryophyllene, and (12) germacrene D.

2007). HS-LPME is widely used in the plant field, in particular with 1-octanol and hydrocarbons of different volatilities as solvents (Bicchi *et al.*, 2008b).

Figure 7 shows a diagram of a HS-LPME sampling system (Bicchi *et al.*, 2012).

Figure 8 reports the hydrodistillation-LPME-GC–FID pattern of the *Lavandula angustifolia* L. volatile fraction (Fakhari *et al.*, 2005).

2.9 In-Tube Microextraction

In-tube microextraction is the basic approach of a group of techniques operating in D-HS mode, principally developed to increase the concentration capability of the S-HS-based HCC-HS techniques (e.g., HS-SPME). With in-tube microextraction techniques, analytes are recovered from a vapor or liquid phase in a needle whose inside is coated or packed with a sorbent or an adsorbent, through which a fixed volume of a liquid or a gas can be pulled-in and pushed-out by a gas-tight syringe or a pump, for an appropriate number of times within a fixed interval of time. Alternatively, they are recovered passively

by diffusion. After sampling, the analytes are online thermally desorbed and introduced by means of a fixed volume of carrier gas into the GC injector body for GC or GC–MS analysis or recovered by solvent elution and transferred online or off-line into an analytical instrument.

Several approaches have been proposed, three of which are mentioned here. The inside needle capillary adsorption trap (INCAT) (McComb *et al.*, 1997) consists of a hollow needle, the inside of which is coated with either carbon or a short piece of GC capillary column. Sampling is as described earlier and analyte recovery is by thermal desorption.

In solid phase dynamic extraction (SPDE), also known as "*the magic needle*" (Lipinski, 2001), analytes from a vapor phase (HS-SPDE) (Bicchi *et al.*, 2004b; Musshoff *et al.*, 2002) are accumulated on a thick film (50 μm, about 4.5 μL) of a polymer, coated onto the inside wall of the stainless steel needle (5.5 or 7.5 cm long) of a gas-tight syringe (2.5 mL). Again, sampling is carried out as described earlier and analytes are recovered by thermal desorption. Several polymeric coatings are available: PDMS, PDMS/activated charcoal,
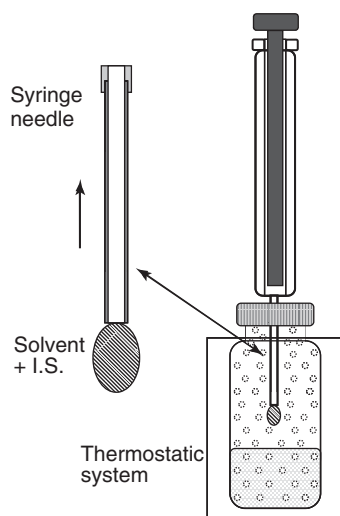
Figure 7 Diagram of a HS-LPME sampling system. (*Source*: Reprinted from Comprehensive Sampling and Sample Preparation: Analytical Techniques for Scientists, 4, Bicchi C., Cordero C., Liberto E., Sgorbini B., and Rubiolo P., Headspace Sampling in Flavor and Fragrance Field, 1–25, Copyright (2012), with permission from Elsevier.)

Figure 9 Diagram of a HS-SPDE sampling system. (*Source*: Reprinted from Comprehensive Sampling and Sample Preparation: Analytical Techniques for Scientists, 4, Bicchi C., Cordero C., Liberto E., Sgorbini B., and Rubiolo P., Headspace Sampling in Flavor and Fragrance Field, 1–25, Copyright (2012), with permission from Elsevier.)

PDMS/cyanopropyl-methyl phenyl-methylsilicone (OV) 225, PDMS/phenyl-methylpolysiloxane, polyethylene glycol (PEG), and polydimethyl siloxane, 7% phenyl-, and 7% cyanopropyl (OV 1701).

The HS-SPDE sampling system is illustrated in Figure 9 (Bicchi *et al*., 2012).

Figure 10 reports the HS-SPDE-GC-MS pattern of a dried leaf sample of rosemary (*Rosmarinus officinalis* L.) (Bicchi *et al*., 2004b).

In needle trap extraction, the sampling devices (needle trap devices, NTDs) are packed with a solid
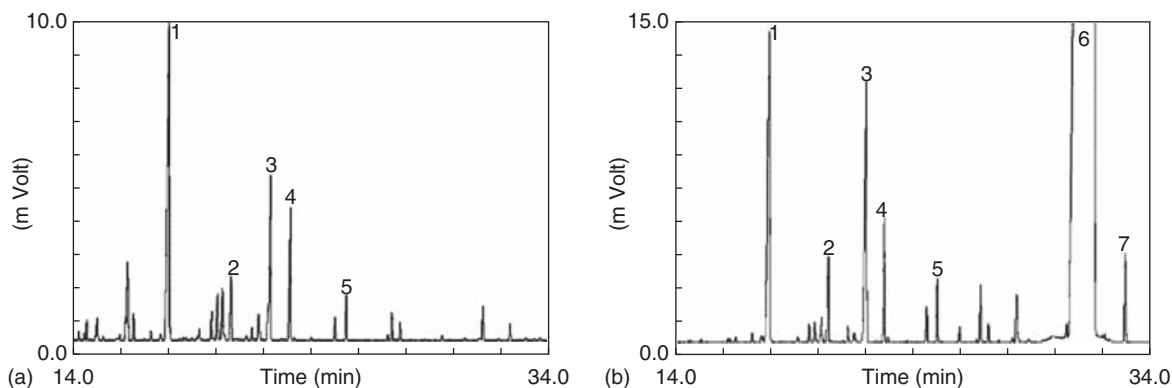
Figure 8 (a) *Lavandula angustifolia* essential oil GC-FID profile . Analysis system: Thermoquest-Finnigan Trace instrument (Thermo electron – Milano, Italy). GC–MS conditions: column: DB-1 (60 m, d_c: 0.25 mm, d_f: 0.20 μm); oven temperature: from 60 to 250 °C at a rate of 5 °C min^{-1}; injection mode: split, ratio 1 : 20; carrier gas: N$_2$; flow rate: 1.1 mL min^{-1}; injector temperature: 250 °C, detector temperature: 280 °C. (b) *L. angustifolia* HD-HSME-GC-FID profile. Sampling conditions: sample amount, 2 g; extraction time, 4 min; drop volume, 3 mL. List of identified compounds: (1) linalool, (2) α-terpineol, (3) linalyl acetate, (4) lavandulyl acetate, (5) geranyl acetate, (6) solvent (*n*-hexadecane), and (7) internal standard (*n*-heptadecane). (*Source*: A.R. Fakhari, *et al*. (2005). Reproduced from Elsevier.)
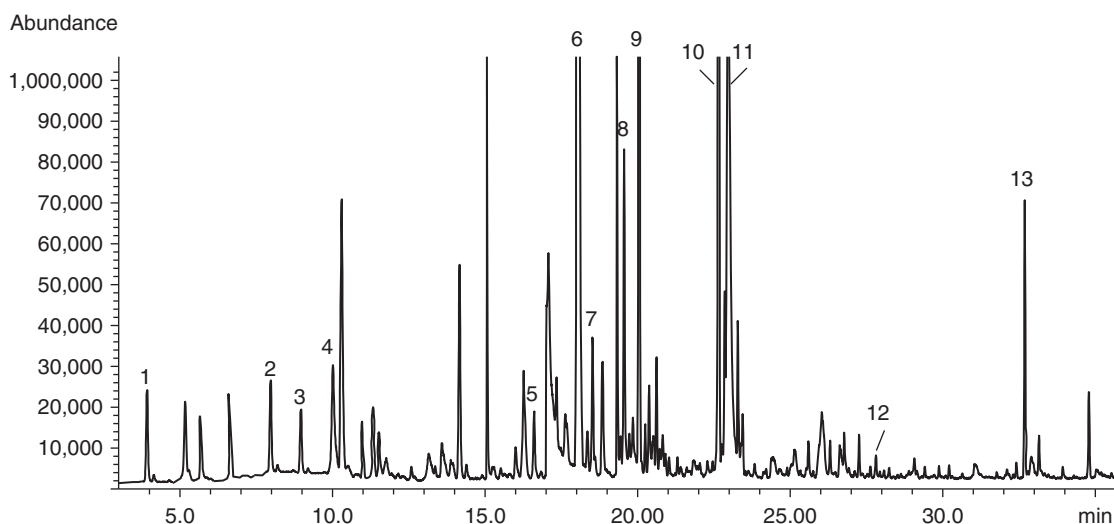
Figure 10 Rosemary (*Rosmarinus officinalis* L.) HS-SPDE–GC–MS profile; analysis system: CTC-Combi-PAL-Autosampler (Bender and Hobein, Zurich, Switzerland) assembled on a GC–MS system consisting of an Agilent model 6890 Series Plus/5973 N. Sample preparation: SPDE needle: PDMS containing 10% of activated carbon; sample volume: 2 mL; vial volume: 21.2 mL; equilibration time: 15 min; sampling temperature: 50 °C. Sampling conditions: agitator (sampling) temperature: 50 °C; headspace syringe temperature: 55 °C; number of filling cycles per extraction: 50; plunger speed for extraction: 50 μL s^{-1} (each aspiration taking 40.5 s); helium volume for desorption: 1 mL; plunger speed for desorption: 15 μL s^{-1}; predesorption time in the GC injection port: 30 s; desorption temperature: 230 °C. GC–MS conditions: column: HTS-FSOT capillary column (PEG 20M, 25 m, d_c: 0.25 mm, d_f: 0.25 μm, Mega (Legnano, Italy)); oven temperature from 20 (2 min) to 220 °C (5 min) at 5 °C min; injector temperature: 230 °C, mode splitless; carrier gas: He; flow rate: 1 mL min^{-1}. MSD conditions – ionization mode: EI (70 eV), temperature: MS source: 230 °C, quadrupole: 150 °C, transfer line: 280 °C; mass range: 35–350 amu. List of identified compounds: (1) α-pinene, (2) limonene, (3) 1,8-cineole, (4) isoamyl alcohol, (5) linalool oxide, (6) camphor, (7) 3,5-octadien-2-one, (8) linalool, (9) bornyl acetate, (10) verbenone, (11) borneol, (12) β-ionone, and (13) thymol. (*Source*: Reprinted from *J. Chromatogr., A*, 1024, Bicchi C., Cordero C., Liberto E., Rubiolo P., and Sgorbini B., Automated headspace solid-phase dynamic extraction to analyse the volatile fraction of food matrices, 217–226, Copyright 2004, with permission from Elsevier.)

material as such or in combination with others of different natures (silica, tenax, carbon, polydivinyl-benzene, etc.); NTDs have recently been the object of in-depth studies, in particular for environmental applications. The main theoretical and practical aspects and applications of in-tube microextraction techniques (INCAT, SPDE, MEPS, NTD, and ITEX) have been the object of a group of reviews, most of them by Pawliszyn's group (Eom, Tugulea, and Pawliszyn, 2008; Lord, Zhan, and Pawliszyn, 2010; Wang, Fang, and Pawliszyn, 2005a).

Other approaches based on the same principles have more recently been described, in particular inside needle dynamic extraction (INDEX) (Ampuero, Bogdanov, and Bosset, 2004; Bosset *et al*., 2004) and in-tube extraction (ITEX) (www.ctc.ch/misc/documents/Itex.pdf). To the best of the authors' knowledge, these techniques have very seldom been reported for use in the field of vegetable matrices.

2.10 Headspace Sorptive Extraction (HSSE)

HSSE was introduced in 2000 by Bicchi *et al*. (2000a) and Tienpont *et al*. (2000) as an extension of SBSE (Baltussen *et al*., 1999a; Baltussen, 2000) for HS in the vapor phase. In HSSE, the analytes are accumulated on a device (known as a *Twister*®) consisting of a thick film of PDMS (25–250 μL) coated onto a glass-coated magnetic stir bar and suspended in the matrix headspace. After sampling, the PDMS stir bar is placed in a glass tube, from which the recovered analytes are thermally desorbed and online transferred to a GC or GC–MS. Some applications in *D-HS* mode, with the twister suspended in a gaseous flow stream for a fixed time, have also been described (Splivallo *et al*., 2007).

Figure 11 reports a diagram of a HSSE sampling system together with conventional (a), dual phase (b), and SE-HSSE twisters (c) (Bicchi *et al*., 2012).

(a) PDMS twisters

PDMS Magnet Glass

(b) Dual phase twisters

Inner phase

Magnets   PDMS
              tubing

(c) SE-HSSE
twisters

PDMS
tubing

Solvent

Glass
stopper

Thermostatic
system
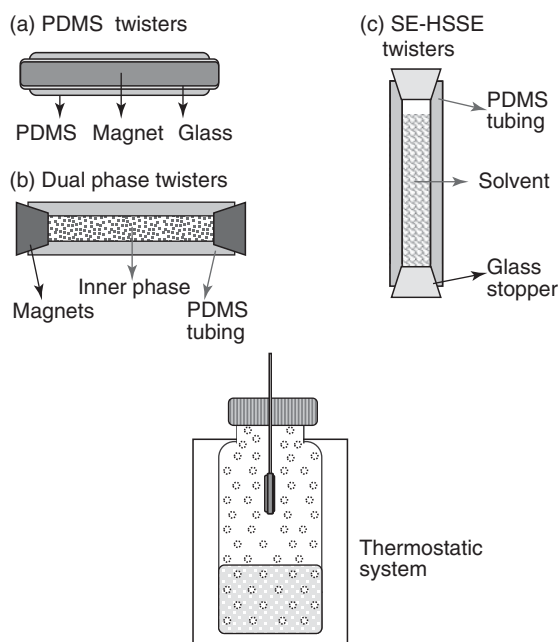
Figure 11 Diagram of a HSSE sampling system together with (a) conventional, (b) dual phase, and (c) SE-HSSE twisters. (*Source*: Reprinted from Comprehensive Sampling and Sample Preparation: Analytical Techniques for Scientists, 4, Bicchi C., Cordero C., Liberto E., Sgorbini B., and Rubiolo P., Headspace Sampling in Flavor and Fragrance Field, 1–25, Copyright (2012), with permission from Elsevier.)

Figure 12 reports the HSSE-GC-MS pattern of a dried leaf sample of rosemary (*R. officinalis* L.).

Several studies have concerned PDMS, because it is an apolar polymer that can discriminate recovery of polar analytes from complex or multi-ingredient matrices. Two approaches have principally been followed: (i) replacing PDMS with a polar polymeric sorbent having similar operative characteristics and performance (Nie and Kleine-Benne, 2011; Sgorbini *et al.*, 2012) or (ii) modifying PDMS polarity, for example, dual-phase twisters (DP-twisters) (Bicchi *et al.*, 2005) or solvent enhanced-HSSE (SE-HSSE) (Sgorbini *et al.*, 2010).

2.11   Solid-Phase Aroma Concentrate Extraction (SPACE)

SPACE is an S-HS technique, introduced by Ishikawa *et al.* (2004), in which analytes are accumulated on a mixture of adsorbents, mainly graphite carbon coated onto a stainless steel rod suspended in the vapor phase; the sampled analytes are then recovered by thermal desorption online to the GC or GC–MS system.

2.12   Large Surface Area High Concentration Capacity HS Sampling

The dimension of the active contact surface of the accumulating material has been shown to influence extraction efficiency and sensitivity over time, for both in-solution and S-HS sampling, because of the larger surface area/extraction phase volume ratio (Bruheim, Liu, and Pawliszyn, 2003). A number of HCC-HS techniques have been developed with the aim of maximizing analyte recovery by exploiting this property (Bruheim, Liu, and Pawliszyn, 2003). Membrane extraction sorbent interface (MESI) was introduced by Yang *et al.* (1994) for in-solution sampling and applied to HS sampling by Segal *et al.* (2000). In MESI, the analytes are dynamically and selectively recovered by a multistep process (Mulder, 1991; Yang *et al.*, 1994) involving (i) vaporization of the components from the investigated matrix, (ii) their transfer by an inert gas flow to a thin-film PDMS membrane, (iii) their diffusion through the membrane, (iv) their stripping from the opposite side of the membrane into a flowing gas, and (v) their concentration onto a sorption (adsorption) trap. The sampled analytes are then recovered by thermal desorption and online analyzed by GC or GC–MS. A portable MESI-GC microsystem was used for on-site in-field monitoring of biogenic emissions from plants (Liu *et al.*, 2004) and a conventional system was used to sample the volatiles emitted from plants into indoor air (Wang *et al.*, 2002).

Sorptive tape extraction (STE) was developed by Sisalli *et al.* (2006) for sampling directly at the surface of the matrix investigated by means of a thin flexible PDMS tape; this has in particular been used for *in vivo* sampling at the human skin surface to study sebum composition. The sampled analytes are then recovered by either thermal or solvent desorption and analyzed by GC or GC–MS. Bicchi *et al.* (2007a) developed STE by direct contact (DC-STE) or for HS sampling (HS-STE) to increase recovery over time, to be used (i) to monitor chemical messages emitted by plants (or animals) as a consequence of a variation in their metabolism because of stress and (ii) to sample aromatic plants and fruits.
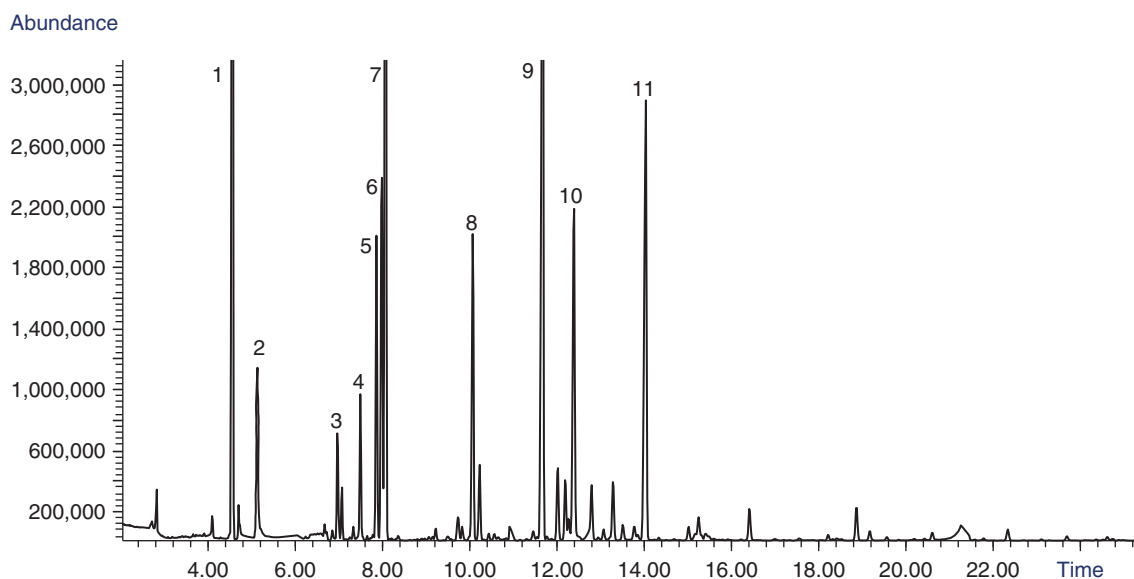
Figure 12 *Rosmarinus officinalis* L. HSSE-GC–MS profile. Analysis conditions – Analysis system: Agilent 6890 GC-5973N MSD system (Agilent, Little Falls, DE, USA) provided with a Gerstel TDU unit installed on an MPS-2 multipurpose sampler (Gerstel, Mülheim a/d Ruhr, Germany). Sample preparation: Twisters: polymeric coating: PDMS, 1-cm long, 0.5-mm thick; sample amount: 20 mg; vial volume: 20 mL; sampling time: 30 min; sampling temperature: 60 °C. GC–MS conditions: column: Mega5 capillary column (30 m, d_c: 0.25 mm, d_f: 0.25 μm, Mega (Legnano, Italy)); oven temperature: from 0 (1 min) to 50 °C (0 min) at 50 °C min^{-1}, then to 250 °C (5 min) at 3 °C min^{-1}; desorption program: CIS: from −50 to 250 °C (5 min) at 12 °C s^{-1}, TDU: from 30 to 250 °C (5 min) at 60 °C min^{-1}; injection mode: split, ratio 1 : 20; carrier gas: He; flow rate: 1 mL min^{-1}. MSD conditions – ionization mode: EI (70 eV), temperature: MS source: 230 °C, quadrupole: 150 °C, transfer line: 280 °C; mass range: 35–350 amu. List of identified compounds: (1) *Z*-3-hexen-1-ol, (2) α-pinene, (3) β-myrcene, (4) Δ-3-carene, (5) *p*-cymene, (6) limonene, (7) 1,8-cineole, (8) linalool, (9) camphor, (10) borneol, (11) verbenone, and (12) germacrene D.

Figure 13 is a scheme of a HS-STE sampling system together with that of a PDMS tape (Bicchi *et al.*, 2012).

Figure 14 reports the HS-STE-GC–MS pattern of licorice (*Glycyrrhiza glabra* L.) leaves.

3 GAS CHROMATOGRAPHY AND VOLATILE FRACTION

3.1 Gas Chromatography: Definition and Short History

The volatile fraction of a plant is, in general, a complex but homogeneous mixture that can contain several hundred components with quite similar structures and physicochemical characteristics, having relatively different polarities and medium-to-high volatility. A high efficiency technique is, therefore, required for their separation: capillary GC is the technique of election for their analysis (Bicchi and Maffei, 2012). The importance of GC in this field has grown steadily, in particular after the introduction of metabolite profiling to characterize plant species and behavior (and thence also metabolomics). The technological improvement of analysis techniques, the dramatic increase in the number of chemical analyses a routine laboratory is required to run (Rubiolo *et al.*, 2010b), and the ever-more-important role played by data processing to meet increasing demands for a "higher" level of information, among others, have radically changed analysis strategies over the past 15–20 years. In developing a modern method for analyzing the plant volatile fraction, a sequence of complementary but closely connected successive steps must be kept in mind, that is, component separation, identification, quantitation (when required), and further data processing.

The following sections of this chapter examine the most recent advances for routine work in GC and fast-GC combined with FID and MS, automatic component identification, quantitation approaches, enantioselective-GC combined with FID (ES-GC) and/or MS (ES-GC–MS), and multi dimensional GC
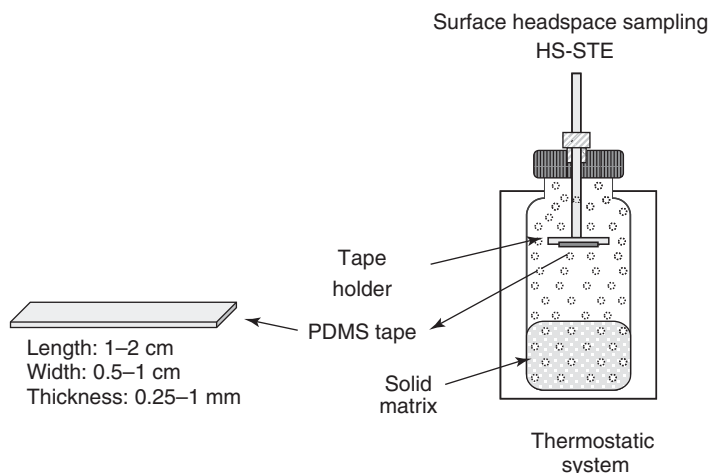
Figure 13 Scheme of a HS-STE sampling system and of a PDMS tape. (*Source*: Reprinted from Comprehensive Sampling and Sample Preparation: Analytical Techniques for Scientists, 4, Bicchi C., Cordero C., Liberto E., Sgorbini B., and Rubiolo P., Headspace Sampling in Flavor and Fragrance Field, 1–25, Copyright (2012), with permission from Elsevier.)
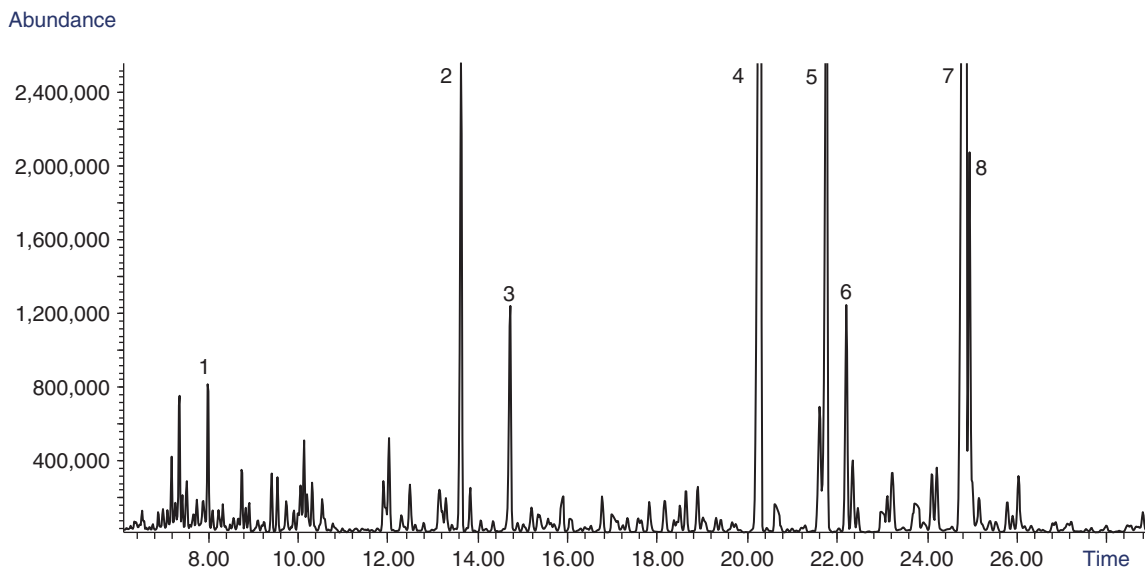
Figure 14 Licorice leaves (*Glycyrrhiza glabra* L.) HS-STE-GC–MS profile. Analysis conditions – Analysis system: Agilent 6890 GC-5973N MSD system (Agilent, Little Falls, DE, USA) provided with a Gerstel TDU unit installed on an MPS-2 multipurpose sampler (Gerstel, Mülheim a/d Ruhr, Germany). Sample preparation: Twisters: polymeric coating: PDMS, 1-cm long, 0.5-mm thick; sample amount: 25 mg; vial volume: 20 mL; sampling time: 60 min; sampling temperature: 30 °C. GC–MS conditions: columns: Mega5 capillary column (30 m, d_c: 0.25 mm, d_f: 0.25 μm, Mega (Legnano, Italy)); oven temperature: from 0 (1 min) to 70 °C (0 min) at 80 °C min^{-1}, then to 250 °C (5 min) at 3 °C min^{-1}; desorption program: CIS: from −50 to 250 °C (5 min) at 12 °C s^{-1}, TDU: from 30 to 250 °C (5 min) at 60 °C min^{-1}; injection mode: split, ratio 1 : 20; carrier gas: He; flow rate: 1 mL min^{-1}. MSD conditions – ionization mode: EI (70 eV), temperature: MS source: 230 °C, quadrupole: 150 °C, transfer line: 280 °C; mass range: 35–350 amu. List of identified compounds: (1) limonene, (2) dodecane, (3) nerol, (4) neryl acetate, (5) italicene, (6) α-bergamotene, (7) γ-curcumene, and (8) α-curcumene.

(GC–GC and GC $\times$ GC) applied to the plant volatile fraction analysis.

Chromatography in general comprises a process that separates the components (solutes) of a mixture, on the basis of a series of consecutive equilibria resulting from their differential interaction between two phases, the first fixed or stationary and having a large specific surface and the second (a liquid or a gas) moving in contact with or through the stationary phase.

In GC, the mobile phase is a gas; if the stationary phase is a high viscosity, low volatility liquid and the separation is based on partition, it is defined as "gas-liquid chromatography" (GLC), whereas if the stationary phase is a solid and mainly operates in adsorption, it is known as "gas–solid chromatography" (GSC) (Jennings, Mittlefehldt, and Stremple, 1997). The former is more widely used, whereas the latter is mainly adopted to separate highly volatile compounds, including fixed gases.

The history of GC was concisely but effectively described by Bartle and Myers (2002). James and Martin (1952) are considered to be the inventors of GC; their 1952 study reported the separation of volatile fatty acids by partition chromatography with nitrogen as mobile phase and silicone oil/stearic acid supported on diatomaceous earth as stationary phase. However, the origins of GC lie in a 1941 publication in which Martin and Synge (1941) first described liquid-phase partition chromatography; in their words, "*Very refined separations of volatile substances should be possible in a column in which permanent gas is made to flow over gel impregnated with a non-volatile solvent . . . .*" Martin and coworkers focused their studies on partition as a separation principle, although other researchers (including Hesse, Cramer, and Phillips (Smolkova-Keulemansova, 2000)) in the same period were working to develop gas adsorption chromatography. The petroleum industry, in particular Shell (e.g., Keulemans and Adlard) and British Petroleum (Desty), immediately adopted GC for compositional analysis, making a big contribution to its development (Smolkova-Keulemansova, 2000). GC rapidly extended its fields of application from theoretical studies (e.g., investigation of reaction kinetics) to biochemistry (amino acid analysis), natural products (e.g., essential oils and steroids), and food and flavors.

Early GC was carried out on packed columns, typically 1- to 5-m long and 1- to 5-mm i.d., filled with particles each coated with a liquid or elastomeric stationary phase. The resolution of packed columns was limited by their length, mainly because of the pressure drop consequent on the resistance to gas flow. This limit was overcome with the introduction of capillary columns: the idea was again suggested by Martin (1957) in 1956 but was independently achieved in 1957 by Golay (1957), who also introduced the theoretical bases in 1958 (Golay, 1958).

In a capillary column, the stationary phase is coated on the inner wall, either as a thin film (wall-coated open tubular) or impregnated into a porous layer on the inner wall (porous layer or support-coated open tubular) with the mobile phase flowing within a single channel. Capillary columns offer dramatically increased separation efficiency versus packed columns, as well as lower operative temperatures, with much better separation in equal times or comparable separation in shorter times. The much smaller amounts of stationary phase employed in a capillary column require specific sample introduction methods and sensitive detectors.

3.1.1 Conventional GC Analysis

The usual strategy in plant volatile analysis in general requires two different-polarity stationary phases to be adopted to obtain complementary standardized chromatographic data because of the high structural similarity of sample components (Mondello *et al*., 1995; Rubiolo *et al*., 2010b; d'Acampora Zellner *et al*., 2008). The apolar stationary phases most widely used in routine analysis are those based on methyl polysiloxanes (SE-30, OV-1, OV-101, DB-1, HP-1, PS-347.5, etc.) and methyl-phenyl-polysiloxanes (SE-52, SE-54, DB-5, HP-5, PS-086, etc.), whereas the most widely employed polar phases are polyethylene glycols (PEG-20M, CW-20M, DB-Wax, etc.). Medium polarity phases based on cyanopropyl-phenyl polysiloxane (i.e., OV-1701 and DB-1701) have also recently gained ground.

The choice of a correct set of column (i.e., coated with orthogonal stationary phases) is fundamental for component identification, which is usually carried out by GC–MS through a synergistic combination of chromatographic data (Kovats indices (Is) (Kovats, 1958), linear retention indices (I^Ts) (Van den Dool and Kratz, 1963; Van den Dool, 1974), relative

retention times, locked retention times (Blumberg and Klee, 1998a; Giarrocco, Quimby, and Klee, 1997), and mass spectra and dedicated libraries (see later).

3.2 New Trends in Stationary Phases

The most recent achievements in this field include stationary phases based on (i) sol–gel technology and (ii) ionic liquids (ILs).

A sol–gel material is a synthetic glass with "ceramic-like" properties, which is generally produced by hydrolysis and condensation of a metal alkoxide (tetraethoxysilane) to form a "glass-like" material at room temperature; this is then coated with suitable polymers (stationary phase) to obtain capillary columns with better thermal stability and less column bleeding (de Zeeuv and Luong, 2002; Sidelnikov, Patrushev, and Belov, 2006). Sol–gel columns coated with dimethylpolysiloxane and polyethylene glycol are currently available from SGE.

ILs are among the most widely investigated new materials for use as GC stationary phases (Dorman *et al.*, 2010). The "room-temperature ionic liquids" (RTILs) are a group of low melting point, nonmolecular solvents with differing solvation properties, whose physical–chemical properties mainly depend on the specific cation–anion combination. ILs are monocationic or geminal dicationic and polyionic ILs, often consisting of an N- or P-containing organic cation (e.g., alkyl imidazolium and phosphonium) and an organic or inorganic anion (Anderson and Armstrong, 2003, 2005; Armstrong, He, and Liu, 1999; Berthod, Ruiz-Angel, and Carda-Broch, 2008). These materials have already demonstrated unique selectivity, while their reliability in terms of efficiency and temperature stability is constantly increasing and is now almost comparable to that of conventional stationary phases. RTILs have "dual behavior" acting as nonpolar stationary phases when used with nonpolar or slightly polar solutes, while being highly interactive and retentive when used with analytes with acidic or basic functional groups (Anderson and Armstrong, 2003, 2005). Commercially available IL capillary columns have been ranked on the basis of their polarity, estimated with McReynolds and Rohrschneider probes (McReynolds, 1970; Rohrschneider, 1966), and have

been found to be more polar than classical PEG 20M columns and more similar in polarity to 25% methyl–75% cyanopropyl siloxane (e.g., SP-2330), 100% cyanopropyl siloxane (e.g., SP-2580), and 1,2,3-tris(2-cyanoethoxy) propane (TCEP). Their properties have also led to interesting applications in flavors and fragrances (Payagala *et al.*, 2009) and essential oil analysis (Cagliero *et al.*, 2012b).

Figure 15 reports chromatographic profiles of chamomile essential oil (*Matricaria chamomilla* L.) obtained with (a) an SE52 capillary column (25 m L $\times$ 0.25 mm i.d. $\times$ 0.25 mm d_f Mega, Legnano (Milan), Italy) and (b) an SLB-IL76 IL stationary phase (30 m L $\times$ 0.25 mm i.d. $\times$ 0.20 mm d_f, Supelco, Bellafonte, PA, USA). See figure caption for chromatographic conditions and analyte identification.

3.3 GC-FID and GC–MS and Derivatization

GC is generally used for volatiles or molecules that can be vaporized by applying temperatures up to 400–450 °C. However, medium-to-high boiling and medium polarity analytes can also be analyzed by GC-FID or GC–MS, after modifying their structure and, as a consequence, their volatility through suitable derivatization reactions. Quite recently, Halket *et al.* (2005) and Fiehn (2008) critically discussed the role of GC–MS after sample derivatization in plant metabolite profiling. Derivatization reagents are of course chosen depending on the analyte structure, or better the function(s) to be derivatized: for example, sugar and phenolic compound hydroxyls are generally derivatized to the corresponding trimethylsilyl derivatives (TMS) with pyridine and *N*,*O*-bis-(trimethylsilyl)trifluoroacetamide (BSTFA) with 1% trimethylchlorosilane (Isidorov and Szczepaniak, 2009), whereas free fatty acids are transformed to the corresponding methyl or ethyl esters. The main advantages of GC–MS with derivatization include (i) a well-repeatable FID response to analytes, depending on their amounts, at least within a homogeneous group of compounds, (ii) the possibility to apply I^T_S to locate (and identify) analytes in the total chromatogram (Isidorov and Szczepaniak, 2009), (iii) mass spectra with diagnostic and reproducible fragmentation patterns, and (iv) better peak shape of components interacting with the stationary phase, which otherwise produce leading or
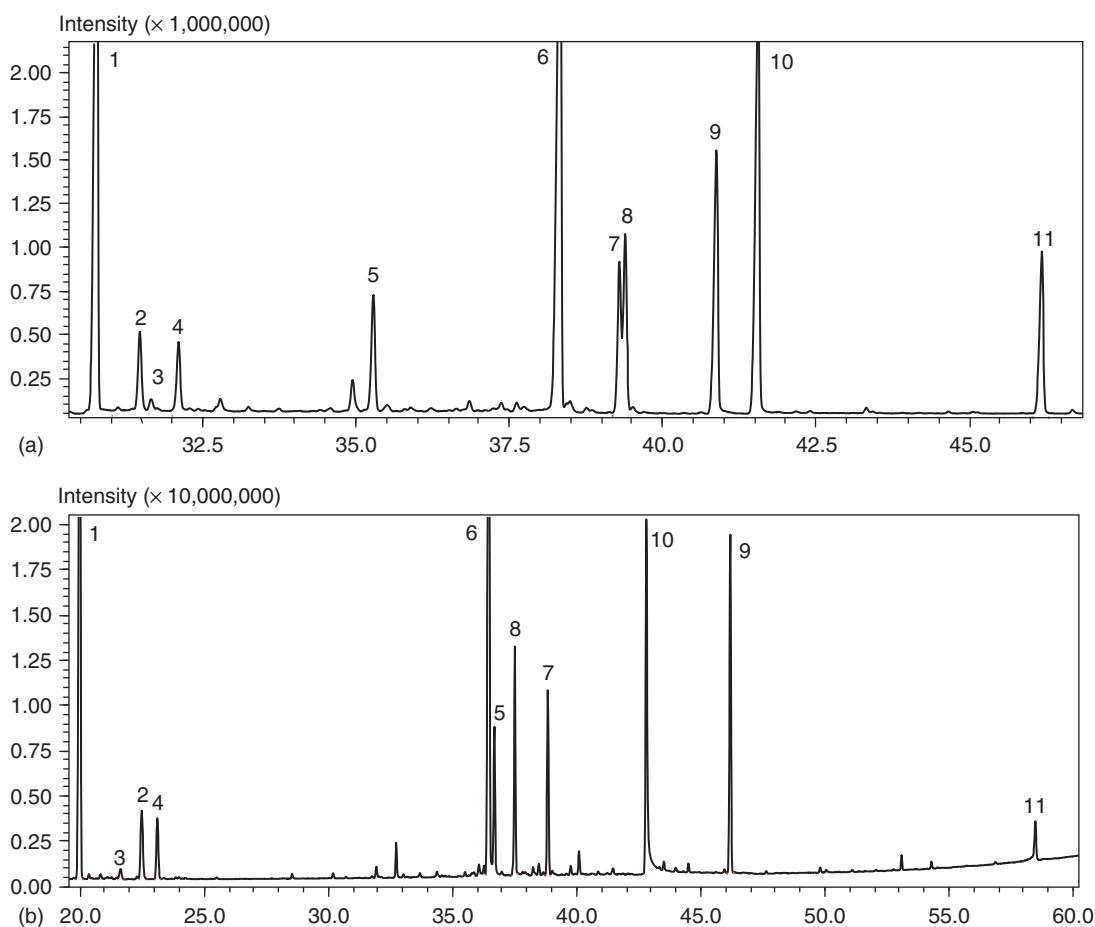
Figure 15 *Matricaria chamomilla* L. essential oil GC–MS profile obtained with (a) a Mega 5 capillary column (25 m × 0.25 mm d_c × 0.25 mm d_f, Mega, Legnano (Milan), Italy) and (b) a SLB-IL76 ionic liquid stationary phase (30 m × 0.25 mm d_c × 0.20 mm d_f), Supelco, Bellafonte, PA, USA. Legend: (1) (*E*)-β-farnesene, (2) germacrene D, (3) β-selinene, (4) bicyclogermacrene, (5) spathulenol, (6) α-bisabolol oxide B, (7) α-bisabolone oxide A, (8) α-bisabolol, (9) chamazulene, (10) α-bisabolol oxide A, and (11) spiroether.

tailing peak distortion, thus improving the reliability of quantitative data. Conversely, derivatization of low volatility components requires (i) the correct application of percent normalization using an internal standard and FID response factors, (ii) the availability of reference standards to build a dedicated library $I^T{}_S$ and mass spectra and of TMS derivatives for correct identification and quantitation, (iii) careful tuning of derivatization conditions to achieve high yield and to avoid multiple products because of incomplete reactions in molecules containing a number of hydroxyls with different reactivities, which can make the chromatogram increasingly complex, and (iv) the possibility of altering the gas chromatographic and MS performance (injector, liner activity, column efficiency, and ion source pollution), so as to avoid any decomposition of reagents and derivatized analyte(s).

3.4 Fast-GC and Fast-GC–qMS Analysis

High speed GC for routine analysis in the plant volatile field has only become popular over the past 10–15 years, although it was first applied by Proot and Sandra (1986) to essential oils in the mid-1980s. Fast chromatographic techniques, and in particular fast-GC, aims to reduce analysis time to a minimum, while keeping separation and qualitative and

quantitative results comparable. Two approaches are currently adopted to speeding up a GC analysis; both achieve a reduction in analysis time of up to a factor of 10 compared to conventional methods. They are both based on shortening columns: the most popular method keeps the efficiency of the GC system constant by reducing the inner diameter (d_c) of the column (0.1 mm or less) (F-GC) (Mondello *et al.*, 2004); the second is based on the rational decrease in excessive efficiency, not necessary for a given separation, by shortening a conventional capillary column (e.g., from $25\,m \times 0.25\,mm\ d_c$ to $5\,m \times 0.25\,mm\ d_c$) (SCC-GC) (Bicchi *et al.*, 2001). SCC-GC is particularly useful for routine quali-quantitative analysis of medium complexity samples (up to about 30–40 components) with 5-m columns (vs 30 m) combined with a suitable temperature program (Bicchi *et al.*, 2001); if the efficiency of the short column is not sufficient for the required separation, a stationary phase with a suitable selectivity for the application can also be chosen (Sandra, Proot, and Dirick, 1987). Figure 16 reports the GC profiles of peppermint essential oil (*Mentha* x *piperita* L.) analyzed by (a) conventional-GC, (b) short conventional column-GC, and (c) F-GC with narrow-bore columns (for analysis conditions and peak identifications, see the figure caption).

The theoretical and practical aspects involved in F-GC have been discussed in several review articles, for instance, by Cramers *et al.* (1999) and David *et al.* (1999). In everyday work, a "fast" GC analysis of a complex mixture in general takes less than 10–15 min and is carried out with a column with d_c between 0.25 and 0.1 mm, length 5–15 m, and temperature programs from 20 to $60\,°C\,min^{-1}$, resulting in peak widths in the 0.5- to 5-s range (Blumberg and Klee, 1998b; Magni, Facchetti, and Cavagnino, 2002).

Klee and Blumberg (2002) introduced a procedure to develop an optimized F-GC method, starting from the original conventional method. This approach (commonly known as *method translation*) implies defining the best speed/separation trade-off with a conventional d_c column, and then automatically translating the resulting method to narrow-bore columns via dedicated software (http://www.chem.agilent.com). It is based on distinguishing the parameters influencing the speed of a GC analysis into translatable and nontranslatable: (i) translatable parameters are column length, inner diameter and film thickness, carrier gas and flow rate, proportional changes in heating rates, duration of

temperature plateau, and detector working at reduced pressure (MS); (ii) nontranslatable parameters are stationary phase, phase ratio, and initial and plateau temperatures. The method translation procedure operates on four fundamental parameters: (i) void time (t_M), that is, a time unit adopted to express time-related components in all chromatographic metrics, (ii) the normalized temperature program, that is, duration of each temperature plateau and heating rates expressed in t_M units measured at the same temperature, (iii) efficiency optimal flow (EOF), that is, the flow corresponding to the column minimum plate height (H_{min}) F_H, and (iv) speed optimal flow (SOF) that, for thin film columns, can be calculated from $F_H \times \sqrt{2}$. Two methods are mutually translatable if they have identical nontranslatable parameters and the same normalized temperature program.

Improvements in instrumentation have also contributed considerably to making routine F-GC reliable: in particular (i) electronic pressure control of the mobile phase, (ii) highly effective oven temperature regulation and control, and (iii) high speed detectors, that is, high frequency FID, high speed quadrupole (qMS), and time-of-flight (TOF) mass spectrometers. Mass spectrometric detection is an indispensable tool for reliable and officially accepted GC analyses; quadrupole analyzers are the most popular MS detectors for GC in routine analysis, because of their reliability and acceptable cost. Their performance with F-GC has been exhaustively and critically reviewed, both in general (Mastovska and Lehotay, 2003) and in essential oil analysis (Rubiolo *et al.*, 2008). The latter study used peppermint essential oil as model sample and dealt with separation, identification, and quantitation of 10 components characteristic of this essential oil, with F-GC-qMS with narrow-bore columns. The results were compared to those obtained by conventional GC-qMS; results with F-GC-qMS were fully comparable or better, while a reduction in analysis time by a factor of about 10 was achieved (from about 35 to 3–4 min).

3.5 Qualitative Analysis

As already mentioned, plant volatile fractions often consist of complex groups of components with very similar structures (e.g., mono- and sesqui-terpenoids and essential oils). They often have indistinguishable
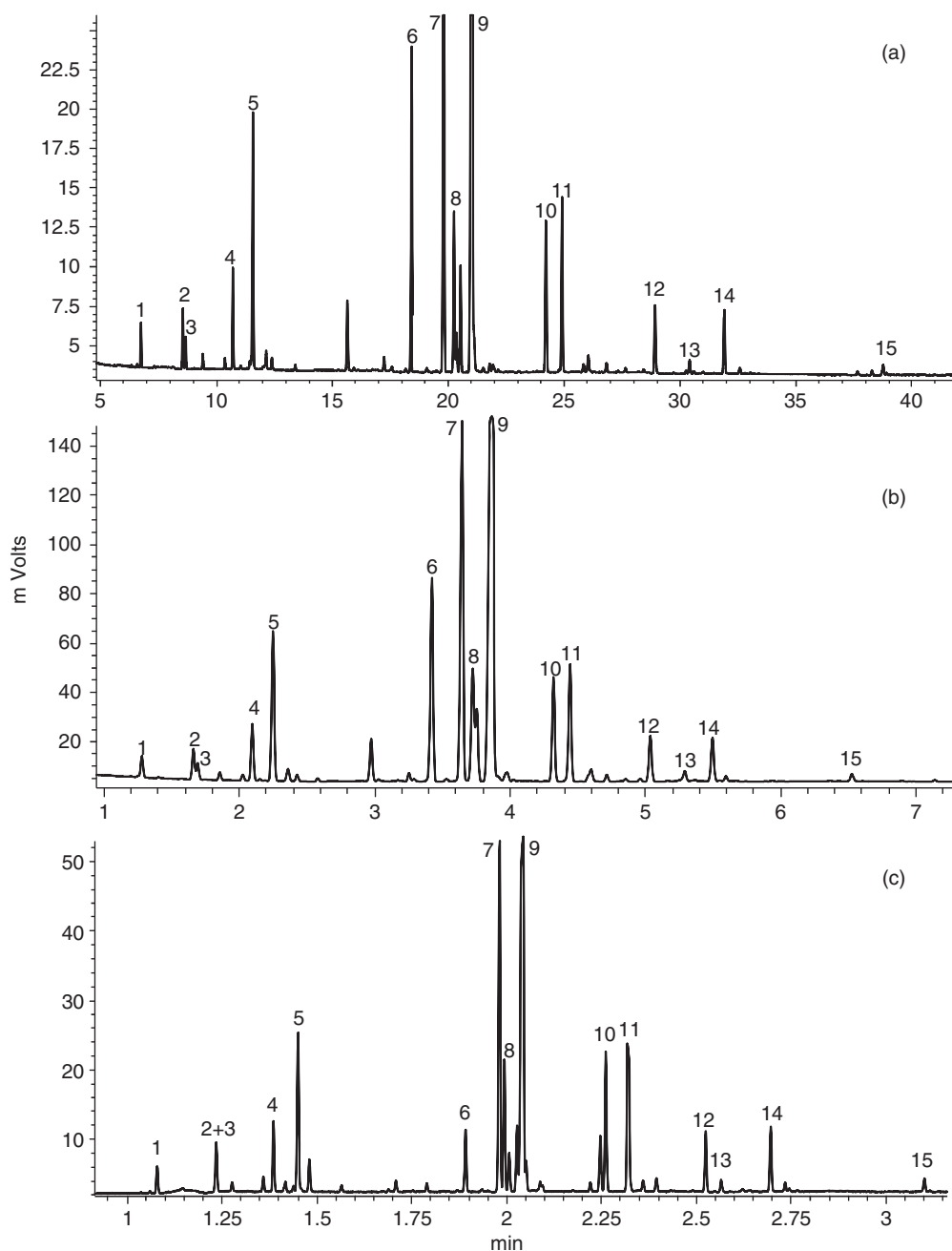
Figure 16 Peppermint (*Mentha x piperita* L.) essential oil GC-FID profiles: (a) conventional GC, (b) short conventional GC column, (c) F-GC with narrow-bore column. Analysis system: Trace GC Ultra (Thermo, Rodano, Milano, Italia). Sample amount: 1 mL of essential oil diluted 1 : 200 in cyclohexane. GC-FID conditions: columns: OV1701 (25 m, d_c: 0.25 mm, d_f: 0.25 μm); OV1701 (5 m, d_c: 0.25 mm, d_f: 0.25 μm); OV1701 (5 m, d_c: 0.10 mm, d_f: 0.10 μm). Oven temperature: from 50 (1 min) to 250 °C (5 min) at 5 °C min^{-1}; from 50 (1 min) to 250 °C (5 min) at 15 °C min^{-1}; from 50 (1 min) to 250 °C (5 min) at 50 °C min^{-1}. Injector temperature: 230 °C; injection mode: split, ratio 1 : 20; carrier gas: H$_2$; flow rate: 1 mL min^{-1}. Detector temperature: 250 °C. List of identified compounds: (1) α-pinene, (2) β-pinene, (3) sabinene, (4) limonene, (5) 1,8-cineole, (6) menthofurane, (7) menthone, (8) menthyl acetate, (9) menthol, (10) pulegone, (11) neomenthol, (12) β-caryophyllene, (13) piperitone, (14) germacrene D, and (15) viridiflorol.

mass spectra, making chromatographic data indispensable for their reliable identification. Retention indices on the same stationary phases are, therefore, widely used in this field to enable different laboratories to share data and successfully exploit the available linear retention index databases. Linear retention indices (I^Ts) were first introduced by Kovats (1958) for isothermal analysis (Is); their use was later extended to temperature programmed analysis (linear retention indices (I^Ts)) by Van den Dool *et al*. (Van den Dool and Kratz, 1963; Van den Dool, 1974). I^T markedly improves the robustness of the identification of an analyte obtained with its mass spectrum, because it supplies a numerical value indicating the entity of its interaction with the applied stationary phase, relative to a reference standard mixture (homologous series of hydrocarbons or fatty acid methyl or ethyl esters) and provides its unequivocal position in the chromatogram. The effectiveness of the I^T/mass spectrum combination is reinforced by their orthogonality, as they are based on completely different principles. In spite of their advantages, only a few commercially available mass spectral library software packages include retention index information to help component identification (Mondello *et al*., 1995; Shellie *et al*., 2003; Van Asten, 2002; Adams, 2007; FFNSC library; NIST 0-NIST/EPA/NIH Mass Spectral Library, 2005; AMDIS, 2007) and even fewer such packages provide for these two parameters' operating "interactively." The role and performance of I^Ts in the analysis of essential oils were quite recently reviewed by d'Acampora Zellner *et al*. (2008). Costa *et al*. (2007) developed an interactive I^T/mass spectrum system (FFNSC MS Library, 2011) to characterize and identify plant volatile fraction components, where I^Ts are automatically calculated and incorporated as an active part of the matching criteria together with mass spectra. The correct identification of an analyte is assured by the range within which its I^T must fall (retention index allowance (RIA)), which must be determined preliminarily. Automatic mass spectral deconvolution (AMDIS) is another software package developed by the National Institute of Standards and Technology (USA) that actively uses I^Ts, often in combination with NIST Mass Spectral Libraries.

Retention time locking (RTL) (Blumberg and Klee, 1998a; Giarrocco, Quimby, and Klee, 1997) is a different approach for reliable analyte identification from its GC retention data in programmed temperature analysis. The principle underlying RTL involves determining the adjustment of inlet pressure necessary to achieve the desired match in retention time(s) of an analyte or analytes with similarly configured GC systems. The RTL-based software package (Flavfid) operates without index determination and can be combined with an additional mass spectrum library in the identification process (AMDIS).

3.6 Enantiomer Separation with Conventional and Fast Enatioselective GC (Es-GC, Es-Fast-GC, and Es-Fast-GC-qMS)

Routine GC chiral recognition, to determine the enantiomeric excess (ee) or enantiomeric ratio (er) of one or more markers of a sample, is an additional and crucial step for an exhaustive analysis of plant volatile fractions, because it enables the stereochemistry of an enantiomer (i.e., its absolute configuration) to be correlated with its biological activity (e.g., its odor), as well as provides the opportunity to define its biosynthesis correctly, to confirm its geographical origin, to evaluate plant sample authenticity and discover possible frauds, and to check and reveal any technological treatments to which a vegetable matrix has been subjected.

Derivatized cyclodextrins (CDs) are nowadays the chiral selectors of election for enantioselective GC (Es-GC) of components present in the plant volatile fraction, without derivatization. Cyclodextrins, also known as cycloamiloses, cycloglucanes, and cyclomaltoligoses, are a homologous series of nonreducing cyclic oligosaccharides, comprising 6–12 (β)-D-glucopyranose units linked by an α-1-4-glycoside bond. CDs are derived from the enzymatic degradation of starch polysaccharide by cyclodextrin glycosiltransferase, from either *Klebsiella pneumonia* or *Bacillus macerans* (Schurig and Nowotny, 1990).

CDs were first introduced as GC chiral selectors for packed columns by Sybilska and Koscielski (1983), while at almost the same time, in 1987, Juvancz, Alexander, and Szejtli (1987) and Schurig and Nowotny (1988) applied them to capillary GC. One of the crucial advances whereby CD became established in routine Es-GC was due to Nowotny *et al*. (1989), who proposed diluting CD derivatives in moderately polar polysiloxane (OV-1701) to improve their chromatographic properties and

extend their operative temperatures range. Three years later, Dietrich *et al*. (1992) introduced a new generation of derivatives, with a substitution pattern giving the resulting CDs high enantioselectivity and stability and good chromatographic properties; these derivatives (still today the most widely used) mainly consist of β-cyclodextrin substituted in position 6 (i.e., the CD narrow side) with a bulky group (*tert*-butyldimethylsilyl-(TBDMS) or *tert*-hexyldimethylsilyl-(THDMS)) and with alkyl or acyl groups (mainly methyl, ethyl, and acetyl) in positions 2 and 3 of its wider side. Enantiomer discrimination with CDs is due to the relatively small difference in the energy of the host–guest interactions that each enantiomer of a racemate establishes with the cyclodextrin selector (Jung, Schmalzing, and Schurig, 1991; Schurig and Nowotny, 1988).

Chiral recognition of marker components in complex real-world samples, such as plant volatile fractions, often requires a two-dimensional approach, because the number of peaks of chiral analytes may be doubled by Es-GC, thus increasing the complexity of the chromatogram and the probability of peak overlap and, consequently, of an incorrect ee and/or er determination. This drawback can be overcome by two complementary but distinct approaches:

1. The first, and most used, approach comprises using a second dimension in separation, that is, MD GC systems. In general, conventional heart-cut 2DGC (GC–GC) is used when a limited number of components have to be submitted to chiral recognition (Dugo, Dugo, and Mondello, 2002; Schomburg *et al*., 1984), whereas comprehensive GC (GC × GC) is applied when very complex samples have to be analyzed and/or when a large number of components must be investigated simultaneously (Dugo, Dugo, and Mondello, 2002; Shellie *et al*., 2004).
2. The second approach entails a second dimension in detection, that is, exploiting the selectivity of mass spectrometry (MS) as a detector. MS is not a selective chiral probe; therefore, enantiomers cannot be discriminated by their mass spectra. Es-GC–MS can provide an unequivocal identification of a given enantiomer in a complex mixture, by reversing the conventional GC–MS approach, that is, the mass spectra (or diagnostic ions) are used to locate the enantiomers in the chromatogram and retention indices (I^Ts) to identify them.

Both of these approaches require the availability of a standard of at least one of the two enantiomers, or of a plant sample unequivocally known to contain one of them, so that their correct elution order in the chromatogram may be assigned (Konig and Hochmuth, 2004). Liberto *et al*. (2008) built an Es-GC–MS library to identify enantiomers automatically, consisting of about 140 chiral standards of plant volatiles; identification is based on the interactive use of MS spectra, RIA (see earlier) and I^Ts, determined on four CD derivatives, that is, $6^{I–VII}$-*O*-pentyl-$2^{I–VII}$-$3^{I–VII}$-*O*-methyl-β-CD, $6^{I–VII}$-*O*-TBDMS-$2^{I–VII}$ - $3^{I–VII}$ - *O*-methyl - β-CD-β-CD, $6^{I–VII}$- *O*-TBDMS-$2^{I–VII}$-$3^{I–VII}$-*O*-ethyl-β-CD-β-CD, and $6^{I–VII}$-*O*-TBDMS-$2^{I–VII}$-$3^{I–VII}$-*O*-acetyl-β-CD.

The earlier considerations are also determinant to speed-up Es-GC analysis (Bicchi *et al*., 2008a). A conventional chiral GC recognition with CDs as chiral stationary phases (CSPs) in general takes quite a long time, because of the high chromatographic efficiency needed, and the slow temperature programs necessary to obtain enantiomer baseline separation, because of the small energy difference between the enantiomer and the CSP host–guest interactions. Es-GC analysis can successfully be speeded up by adopting: (i) short conventional internal diameter (d_c) or narrow-bore columns and (ii) mass spectrometry as selective detector, combined interactively (or not) with I^Ts. Short columns also enable enantiomer elution temperatures to be reduced, with a gain of enantioselectivity that can (at least partially) recoup the loss of efficiency (Bicchi *et al*., 2008a; Cagliero *et al*., 2012a).

This approach was investigated in depth on a series of real-world samples in the essential oil field, analyzed on four narrow-bore columns of different lengths (1-, 2-, 5-, and 10-m long, 0.10-mm d_c, 0.10-μm d_f) coated with different CD chiral selectors; the results were compared to those of a corresponding conventional d_c column (25-m long, 0.25-mm d_c, 0.15- and 0.25-μm d_f) (Bicchi *et al*., 2008a). Figure 17 reports the ES-GC–MS profiles of bergamot essential oil (*Citrus bergamia* Risso et Poiteau) on a 5 m × 0.10 mm d_c × 0.10 μm d_f coated with 30% $6^{I–VII}$-*O*-TBDMS-$2^{I–VII}$-$3^{I–VII}$-*O*-acetyl-β-CD in PS086 at 5 °C min^{-1} (a) and 10 °C min^{-1} (b) together with the corresponding linalyl acetate extract ion profiles at *m/z* 80 (for analysis conditions, see the figure caption). The analysis time compared to the corresponding conventional GC analysis is reduced by a factor of about 5 at 5 °C min^{-1} and 8 at 10 °C min^{-1}.
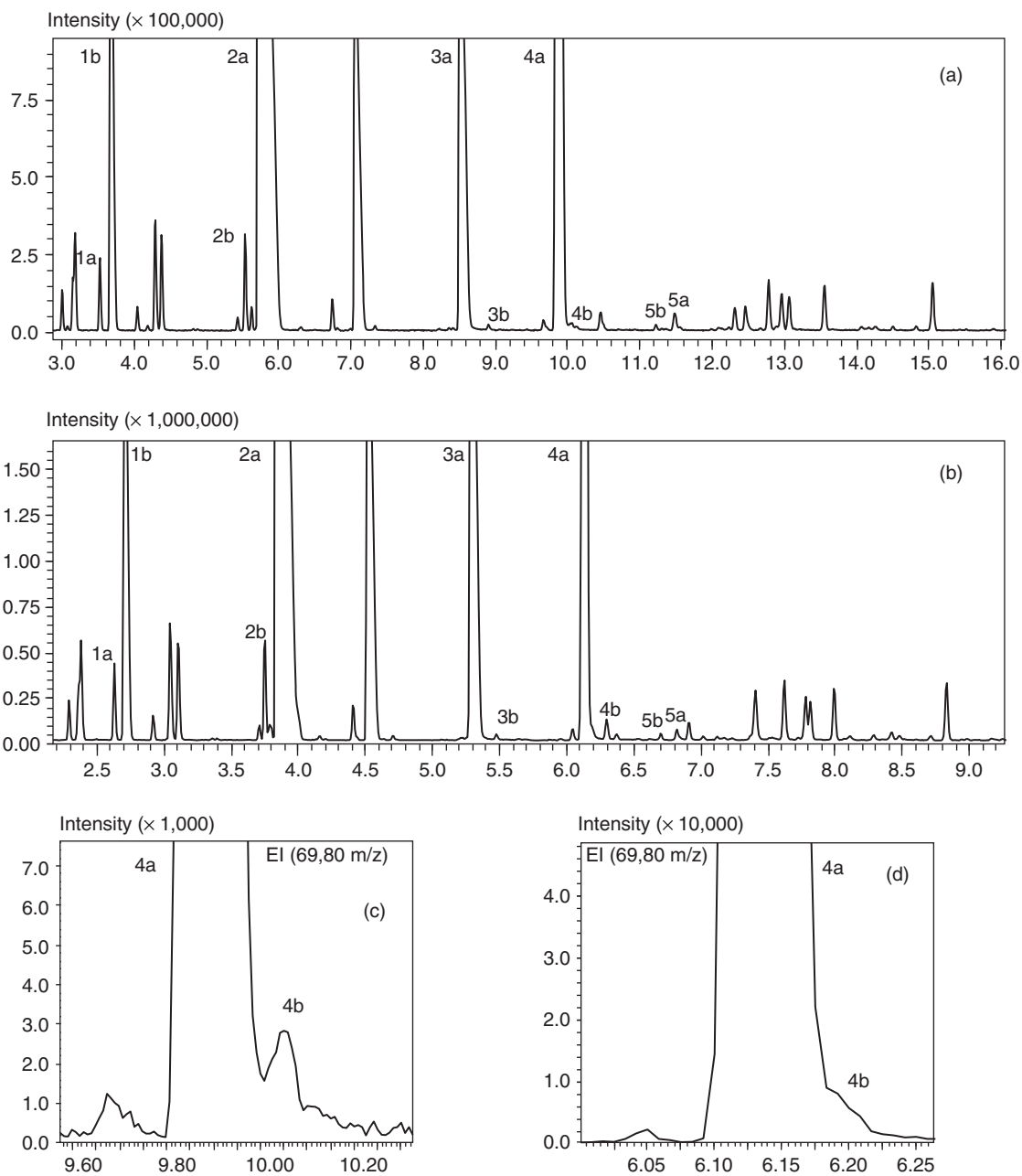
Figure 17 ES-GC–MS profiles of *Citrus bergamia* Risso et Poiteau essential oil on a 10 m × 0.10 mm d_c × 0.10 mm d_f column coated with 30% 6[I–VII]-*O*-TBDMS-2[I–VII]-3[I–VII]-*O*-ethyl-β-CD in OV1701 at 5 °C min^{-1} (a) and 10 °C min^{-1} (b). Extract ion profiles of linalyl acetate (69, 80 *m/z*) at 5 (c) and 10 °C min^{-1} (d). Peak identification: (1) β-pinene, (2) limonene, (3) linalool, (4) linalyl acetate, and (5) α-terpineol; a, (*R*)-enantiomer; b, (*S*)-enantiomer.

A second and complementary approach consists of seeking the best trade-off, between separation of the most critical peak pairs in a sample and analysis time. As reported by Klee and Blumberg (Blumberg, 1997; Blumberg and Klee, 1998a), this strategy implies four main steps (Section 3.4): (i) determination of the critical parameters required for the GC method translation approach (i.e., t_M) from the initial routine analysis conditions on the conventional column, (ii) determination of the optimal normalized multirate temperature program (r) for a predetermined fixed column flow-rate, (iii) determination of efficiency-optimized flow (EOF) and speed-optimized flow (SOF) for the normalized optimal multirate temperature program, and (iv) translation of the method to the corresponding narrow-bore columns (Bicchi *et al.*, 2010; Blumberg, 1997). This strategy was applied to a lavender essential oil (*L. angustifolia* P. Mill.) sample, first optimizing the analysis conditions on a conventional 25 m × 0.25 mm d_c starting from routine conditions and then transferring the method to shorter narrow-bore columns (10 and 5 m × 0.10 mm) using method-translation software, so that peak elution order and separation were maintained. Figure 18 reports the ES-GC–MS profiles of the lavender essential oil sample, analyzed on a 30% $6^{I–VII}$-O-TBDMS-$2^{I–VII}$-O-ethyl-$3^{I–VII}$-O-methyl-β-CD in PS086 chiral selector: (a) routine conditions on conventional column (25 m × 0.25 mm × 0.25 μm), (b) SOF conditions on conventional column (25 m × 0.25 mm × 0.25 μm), and (c) SOF conditions on 10-m NB column (10 m × 0.10 mm × 0.10 μm) (for analysis conditions, see the figure caption). The analysis time was reduced from about 40 min, with the routine method, to 25 min under optimal conditions with the conventional d_c column and to 13 min with the narrow-bore column, while keeping the same degree of separation.

3.7 Multidimensional Gas Chromatography

An MD separation was defined Giddings (1987) as: " … *an orthogonal two column separation, with complete transfer of solute from the separation system 1 (column 1) to the separation system 2 (column 2), such that the separation performance from each system (column) is preserved*." Two techniques, commonly known as MD GC, are at present used for

volatile fraction analysis: (i) heart-cut GC–GC where *one or a very few components* (peaks) eluting from the first column (1st dimension – ^{1}D) are online and automatically transferred to a second column (2nd dimension – ^{2}D) coated with a different stationary phase, through a dedicated time-programmable interface, for a given time fraction of the whole chromatographic run (Deans, 1968; Mondello *et al.*, 1998; Schomburg *et al.*, 1982), and (ii) two-dimensional comprehensive GC (GC × GC) introduced by Liu and Phillips (1991), where *each component* eluting from the first column is online and automatically trapped, refocused, and reinjected into the second column, in a fixed time (4–8 s) by a thermal or valve-based focusing device (*modulator*). A GC × GC system consists of a column of conventional inner diameter and length, connected to a very short narrow-bore column, the latter enabling a fast analysis in the second dimension, taking the time of a modulation. GC × GC is the most recent and powerful separation technique available at present.

Both heart-cut GC–GC and comprehensive GC have been used for several applications in analysis of the plant volatile fraction (Adahchour *et al.*, 2003; Cordero *et al.*, 2007; Debonneville, Thome, and Chaintreau, 2004; Dugo, Dugo, and Mondello, 2002; Shellie and Marriott, 2003; Shellie, Marriott, and Morrison, 2001). As discussed earlier (Section 3.6), heart-cut GC–GC is used in chiral recognition in complex real-world samples to avoid peak overlap because of the doubling of peaks corresponding to chiral components, thus obtaining correct "ee" or "er" determinations. Figure 19 reports heart-cut GC–GC analysis of bergamot essential oil (*C. bergamia* Risso et Poiteau) with a HP-5 (25 m × 0.25 mm d_c × 0.25 mm d_f) as a first column and 30% $6^{I–VII}$-O-TBDMS-$2^{I–VII}$-$3^{I–VII}$-O-ethyl-β-CD in PS086 (25 m × 0.25 mm d_c × 0.25 mm d_f) as a second column; for analysis conditions and peak identification, see the figure caption.

However, with heart-cut GC–GC, only a small number of peaks (or fractions) (possibly eluting at relatively different temperatures) can be transferred from the first to the second column in a single GC run; different eluting temperatures are required in particular when heart-cut GC–GC is run in a single oven. GC × GC does not suffer from this drawback, because each peak eluting from the first dimension is online and fully transferred to the second column. Unlike heart-cut GC–GC, in GC × GC chiral
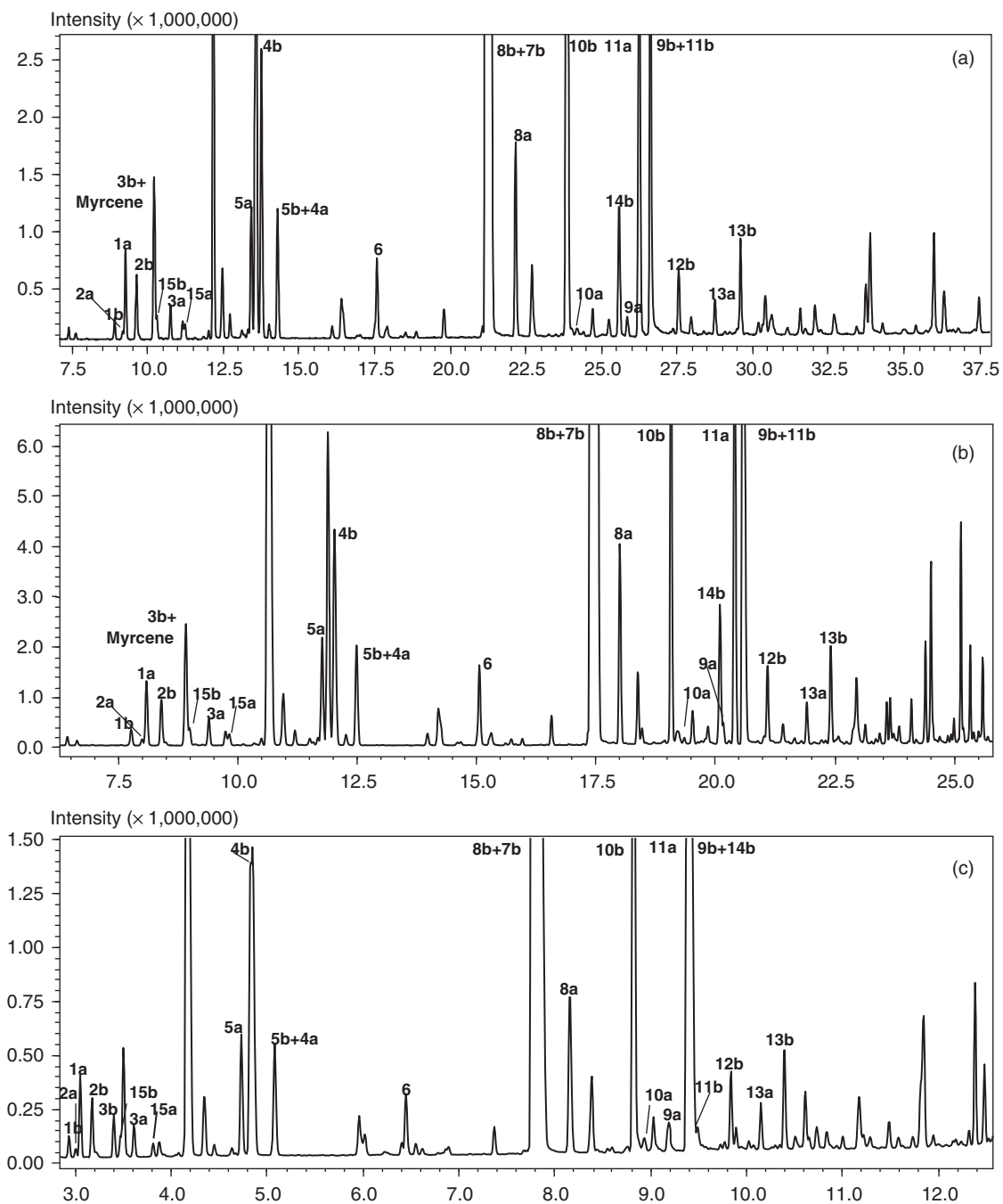
Figure 18 ES-GC–MS profile of *Lavandula angustifolia* P. Mill. essential oil analyzed on a 30% 6^{I-VII}-*O*-TBDMS-2^{I-VII}-*O*-ethyl-3^{I-VII}-*O*-methyl-β-CD in PS086 column: (a) routine conditions on a conventional column (25 m × 0.25 mm d_c × 0.25 mm d_f), (b) SOF conditions on a conventional column, (c) SOF conditions on a 10-m NB column (10 m × 0.10 mm d_c × 0.10 mm d_f). Peak identification: (1) α-pinene, (2) camphene, (3) β-pinene, (4) α-phellandrene, (5) limonene, (6) 1-octen-3-ol, (7) camphor, (8) linalool, (9) borneol, (10) linalyl acetate, (11) terpinen-4-ol, (12) lavandulol, (13) α-terpineol, (14) lavandulyl acetate, and (15) sabinene; a, (*S*)-enantiomer; b, (*R*)-enantiomer.

GC–GC bergamotto
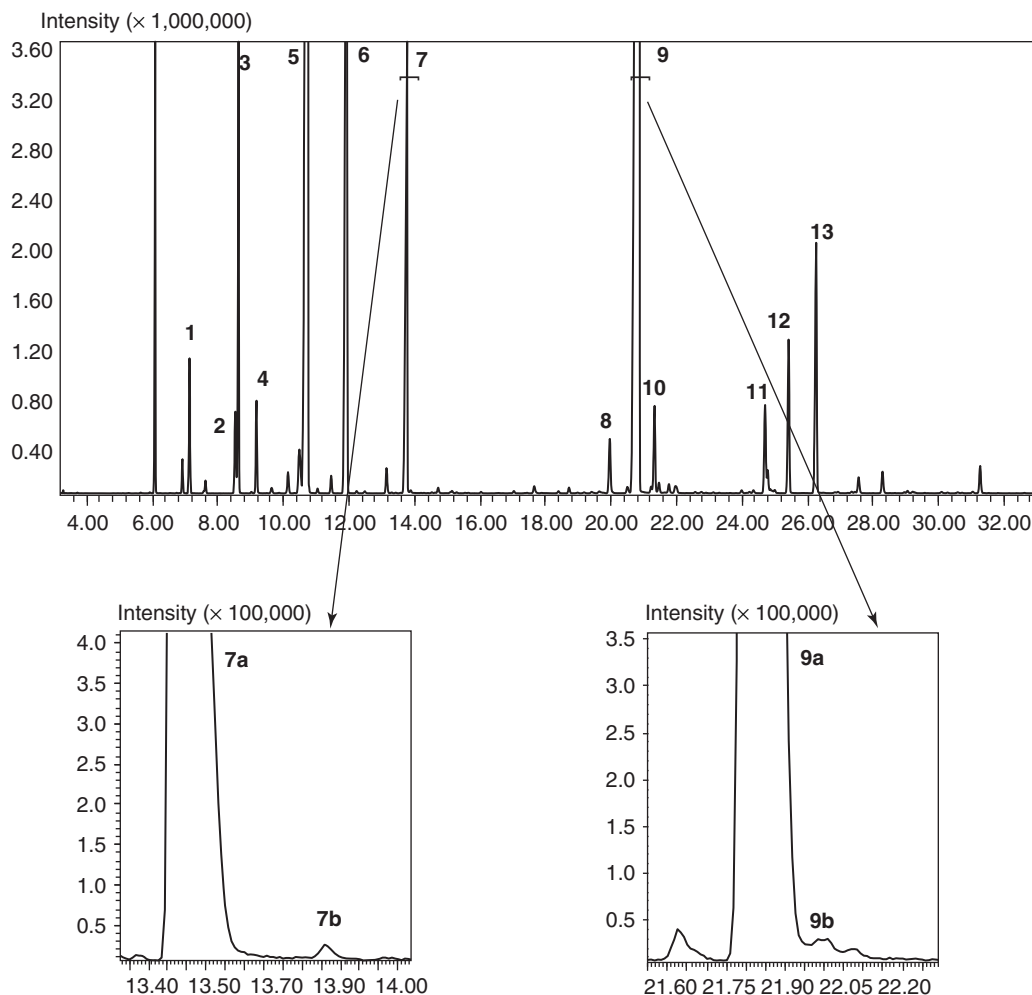
Figure 19 Heart-cut GC–GC analysis of *Citrus bergamia* Risso et Poiteau 1st column: HP-5 (25 m × 0.25 mm d_c × 0.25 mm d_f), 2nd column: 30% 6$^{I–VII}$-O-TBDMS-2$^{I–VII}$-3$^{I–VII}$-O-ethyl-β-CD in PS086 (25 m × 0.25 mm d_c × 0.25 mm d_f). Analysis conditions: temperature program, 1st column 50 °C/3 °C/min/220 °C; 2nd column, 60 °C/2 °C/min/180 °C. Peak identification: (1) α-pinene; (2) sabinene; (3) β-pinene; (4) β-myrcene; (5) limonene; (6) γ-terpinene; (7) linalool; (8) neral; (9) linalyl acetate; (10) geranial; (11) α-terpinyl acetate; (12) neryl acetate; and (13) geranyl acetate; a, (*R*)-enantiomer; b, (*S*)-enantiomer.

recognition, the chiral column must be in the first dimension because of the high efficiency required for effective enantiomer separations (Shellie *et al.*, 2004; Shellie and Marriott, 2002).

GC × GC was immediately successful in the mineral oil industry, where samples consist of thousands of components (Adahchour, Beens, and Brinkman, 2008), but it is now also extensively applied to plant volatile fraction analysis. This is particularly the case for samples containing hundreds of components (e.g., vetiver essential and oils) or components that

may be very similar in structure and difficult to separate (e.g., sandalwood essential oils), but that are decisive to define product quality and authenticity, and also when the profile is used as a parameter characterizing a species.

Figure 20 reports the GC × GC contour plots of the Vetiver (*Chrysopogon zizanioides* (L.) Roberty) essential oil (a) from Haiti and of the volatile fraction of dried roots (b) from Java sampled by HS-SPME with more than 300 separated components (for analysis conditions, see the figure caption).
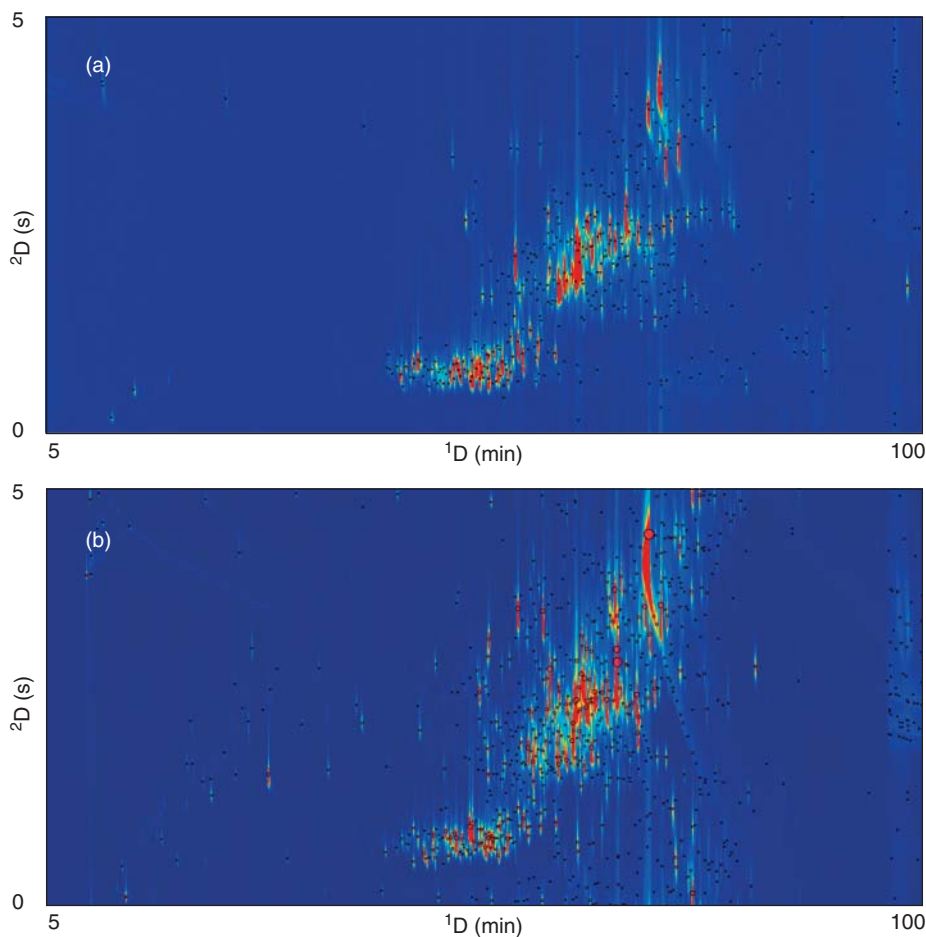
Figure 20   GC×GC contour plots of the Vetiver (*Chrysopogon zizanioides* (L.) Roberty) essential oil from Haiti (a) and of the volatile fraction of dried roots (b) from Java. Black bubbles indicate 2D separated peaks. Analysis conditions: Vetiver essential oil (a) was diluted in cyclohexane ($10 \, mg \, L^{-1}$) and directly injected ($2 \, \mu L$) in a split–splitless injector; dried roots from Java were submitted to HS-SPME sampling as follows: 2-cm Stableflex $50/30 \, \mu m$ DVB-Carboxen-PDMS (Supelco, Bellafonte, USA) – sample amount: 500 mg, vial volume: 20 mL; sampling time: 60 min; temperature: 80 °C. GC×GC analyses: GC×GC–MS system: Agilent 6890 GC–Agilent 5975 MSD ionization mode: EI 70 eV (Agilent, Little Falls, DE, USA); transfer line temperature: 280 °C, scan range: *m/z* 35–250 in fast scanning mode ($10,000 \, amu \, s^{-1}$). GC×GC interface: KT 2004 loop modulator (Zoex Corporation, Houston, TX, USA), modulation time: 5 s. Column set: 1D: SE54 column (30 m, d_c: 0.25 mm, d_f: $0.25 \, \mu m$), 2D OV1701 column (1 m, d_c: 0.1 mm, d_f: $0.10 \, \mu m$) (Mega – Legnano (Milan), Italy). Analysis conditions: injection mode: split, ratio: 1/20, temperature: 270 °C; carrier gas: helium; temperature program: 60 °C (1 min)/2.5 °C/min/260 °C (10 min).

3.8   Headspace, GC (GC–MS) and Volatile Quantitation

The need for data concerning the relative abundance of components, and/or the quantitation of diagnostic markers of plant volatile fractions, is continually increasing not only because of the ever growing demand for quality and safety controls but also because of the crucial role they play in metabolite profiling (Bicchi and Maffei, 2012). A technique or a method cannot be fully accepted as reference technique if it cannot provide reliable quantitation. It is now generally accepted that quantitative analysis involves the entire analytical procedure (i.e., sample preparation, analysis, and data processing). Unfortunately, though, it is often approached ambiguously because it may be quite complex and time-consuming to run correctly, as was quite recently discussed in

depth by Bicchi *et al.* (2008c) and Cicchetti, Merle, and Chaintreau (2008) for the flavor, fragrance, and essential oil fields. This section summarizes the basic approaches to quantitation with HS and GC. These two techniques are highly integrated and complementary; in general, all approaches developed for HS (or HCC-HS) are based on data resulting from a GC analysis. The first part of this section is a general discussion about the component quantitation of the plant volatile fraction in GC, whereas the second part includes a rather detailed description of quantitation approaches in HS or HCC-HS sampling.

3.8.1 Quantitative Analysis in GC

The different chemical compositions and physical states of matrices to be investigated, and the varying nature of the data required, means that there is no "one size fits all" for quantitative analysis: this section discusses some of the most evident aspects and fields of application of the three most popular quantitation approaches: (i) relative % abundance, (ii) internal standard normalized % abundance, and (iii) true quantitation of one or more components (target analysis) with or without a validated method.

Relative % abundance is the most popular approach and is used when the relative percent ratio among the components of extracts, distillates, essential oils, or headspaces are to be determined, and in particular when the resulting data will be used to compare samples of the same matrix within a group. In spite of its clear definition and relative simplicity, this approach is often applied incorrectly: two conditions are mandatory for the correct comparison of the composition of a group of samples but are often "forgotten": (i) the raw data must be normalized *versus* an internal standard (or at least vs an external standard, if an automatic injector is available) and (ii) the % abundance versus the sum of the normalized areas of a fixed number of markers must be determined; for preference, they should be selected from among components common to all samples investigated. From these considerations, the relative % abundance as such can correctly be used only when what is required is the relative ratios of components for the investigated sample. A further parameter influencing the correct quantitation is that detector response varies with the chemical structures (analyte response factor). The most common GC detectors for this field,

that is, MS and FID, perform differently. MS as detector *cannot* be used in total ion current (TIC) mode for quantitation in normalized % abundance, because ion formation and intensity depend on the analyte structure, and the resulting data would not be representative of the true component ratios in the sample. FID gives response factors that differ widely depending on the analyte structure, for example, some esters can respond up to 1.6 times as strongly, compared to a linear hydrocarbon (*n*-nonane) (Costa *et al.*, 2008). To determine the response factor of each component experimentally, using FID, would be very time consuming, due to the high complexity of plant volatile fraction; the result would also be limited by the lack of pure standards. Two solutions have been proposed to overcome this difficulty: (i) the first, introduced by Costa *et al.* (2008), involves the determination of a response factor for each class or subclass of compounds (hydrocarbons, aldehydes, alcohols, esters, etc.) in the investigated sample, calculated versus the internal standard, taking one component representative of each class; (ii) the second approach, proposed by Cicchetti, Merle, and Chaintreau (2008) and Chaintreau *et al.* (2003), involves building up of a response factor database, calculating them predictively from combustion enthalpies and molecular structures.

Normalized % abundance, however, is not sufficient to quantify the "true" amount of an analyte in a plant volatile fraction ("true" quantitation). When one or more diagnostic markers must be quantified, for example, because the amount present is restricted by the legislation in force, then true quantitation is necessary. Analyte concentration or absolute amount in a real-world sample must be determined from its chromatographic area (measured in SIM mode with MS) in the suitably diluted sample, normalized versus the internal (or external) standard, and quantified via a calibration curve built up from amounts of marker standard in the linearity concentration range (Cordero *et al.*, 2007; Schieberle and Grosch, 1987). Quantitation with GC–MS can also be carried out by the stable isotope dilution assay (SIDA), introduced by Schieberle and Grosch (Schieberle and Grosch, 1987; Steinhaus, Fritsch, and Schieberle, 2003), described in greater detail later. Isotope-labeled components (in general 2D or ^{13}C-labeled compounds) overcome most of the problems mentioned earlier, giving response factors and recoveries that are equal or very close to those of the native compounds and the same fragmentation

pattern, but with a known increase in molecular weight, that is, with diagnostic or molecular ions that is easy to detect by MS.

3.8.2   Quantitative Analysis in HS

Quantitative analysis in HS depends on the analyte distribution between the matrix and the vapor phase, in equilibrium or not with it. These considerations can quite easily be extended to HCC-HS techniques, bearing in mind that in this case the analyte or analytes are distributed in a three-phase system (again in equilibrium or not): that is, sample, headspace, and accumulating polymer. Several parameters influence HS recovery and quantitation:

1. the experimental conditions, namely temperature, sampling time, stirring, sample and headspace volumes, and, as a consequence, phase ratio (β);
2. the matrix effect, that is, the influence of other sample components on the release of the investigated analyte(s) into the headspace (partition coefficient);
3. the difficulty to "build" a reliable reference matrix of a known composition for correct analyte calibration;
4. *salting out*, that is, the addition of a significant amount of a strong electrolyte (in general a salt such as sodium chloride, or potassium sulfate or chloride) to an aqueous sample to reduce the sample/headspace partition coefficient of the investigated analyte(s) and, as a consequence, their solubility, while increasing their concentration in the vapor phase. Salting out is, therefore, especially effective for quantitation of weakly polar analytes in aqueous samples.

These factors are in general met by applying sampling conditions as standardized as possible and by adopting suitable calibrations and internal and/or external standards. Three calibration approaches are the most widely used, although they all need to be carefully managed. They are SA, SIDA, and MHE (see later) (Bicchi *et al*., 2012; Kolb and Ettre, 1997).

Correct quantitation is conditioned by three main factors:

1. the physical state of samples. In the plant field, solid samples have very often to be analyzed.

Two approaches are possible with solid matrices: quantitation on the matrix as such or on the matrix suspended in a nonvolatile liquid. When possible, sample suspension in liquid phase (in particular water) is recommended because (i) it enables the resulting suspension to be spiked reliably with an internal standard and (ii) sensitivity can be increased when analyte solubility in the dispersion liquid is low or null (Bicchi *et al*., 2012).

2. correct use of the quantitation approaches (SA, SIDA, and MHE);
3. standardization and/or normalization of the accumulating polymer(s) in HCC-HS techniques. This is a fundamental aspect with HCC-HS techniques, as shown by Bicchi *et al*. (2007b) for fibers in HS-SPME when used in routine HS analysis of the headspace of aromatic and medicinal plants. The considerations given here for HS-SPME fibers can easily be extended to other HCC-HS devices.

The performance of the polymer coating(s) can be monitored through different but complementary approaches, including

1. analysis of a standard mixture or a reference sample (better if it contains the analyte(s) of interest) to evaluate the performance of the accumulating polymer(s) over time. A number of objective approaches have been developed in this respect, including (i) analyte relative concentration factors (CFs) (Bicchi, Drigo, and Rubiolo, 2000b) and (ii) the criterion function, F_{ij}, introduced by Zuba, Parczewski, and Reichenbacher (2002) and simplified by Hamm *et al*. (2003) with which fibers are regularly tested with a reference sample or standard solution to verify and compare their performance
2. adopting an internal standard inside the sample. This approach enables data to be normalized, and recoveries of the accumulating polymer(s) compared, in particular when the vegetable matrix is suspended in a dispersion liquid.
3. normalizing polymer coating performances through equilibrium in-fiber internal standardization, as proposed by Pawliszyn's group for HS-SPME (Wang *et al*., 2005b), which entails preloading a standard into the polymer coating as an approach to standardize sampling internally.

3.9 Quantitation Approaches: Standard Addition (SA), Stable Isotope Dilution Assay (SIDA), and Multiple Headspace Extraction (MHE)

In addition to the internal and external standard methods usually adopted for liquid sampling, three approaches are widely used in headspace quantitation of volatiles, in particular from solid matrices: SA, SIDA, and MHE (Bicchi *et al.*, 2012).

3.9.1 Standard Addition (SA)

SA was the first approach introduced for quantitation of headspace components that overcame the matrix effect, and it is still the most widely used (Kolb and Ettre, 1997). In addition to the analysis of the sample as such, SA requires that other determinations be carried out on the same sample, after known amounts of standard of the investigated analyte(s) have been added.

SA can be carried out as a "single-point" or a "multiple-point" calibration. *Single-point calibration* requires both the sample as such and the sample spiked with a known amount of analyte standard to be analyzed. The amount of spiked standard must provide a signal at least 20% more intense than that of the analyte in the sample as such. The result with single-point calibration can be affected by deviations in sampling performance or in linearity. These deviations can be avoided by *multiple-point calibration*. The latter method requires two steps: (i) the sample is analyzed as such, and after having been spiked with increasing and known amounts of standard analyte, (ii) the analyte amount initially present in the sample is calculated from the standard-added linear calibration function by extrapolation.

The single-point calibration procedure should mainly be used in routine analyses, after the linear range of the investigated analyte(s) over the concentration range of interest has been determined by the multiple-point calibration method. The main limitations of SA are (i) the need for specific calibration for each analyte, entailing a large number of measures and making the method time consuming, (ii) the possible nonavailability of analyte standard, and (iii) the nonhomogeneous distribution of the added standard within the solid matrix, making it necessary to spike the standard directly into the gas phase of the vial ("gas phase addition," GPA) to obtain reliable results (Kolb and Ettre, 1997).

3.9.2 Stable Isotope Dilution Assay (SIDA)

The SIDA was introduced by Schieberle's and Grosch's groups (Schieberle and Grosch, 1987; Steinhaus, Fritsch, and Schieberle, 2003) for reliable quantitation of analytes in liquid samples, in particular for trace analysis (Rychlik and Asam, 2008). SIDA adopts a labeled stable isotope of the target analyte as internal standard, thus possessing almost identical chemical and physical properties. The analyte labeled with an appropriate isotope is, therefore, "conceptually" the most effective internal standard, provided that labeling is stable throughout the analytical procedure. The analyte and its isotopic analog can then be easily differentiated by GC–MS because of their different molecular weights: this makes MS detection indispensable. With SIDA, any loss of analyte is, therefore, fully compensated by an identical loss of the isotopolog, unlike what occurs with conventional internal standards, where the ratio between standard and analyte may be altered during sample clean-up procedures, which often include several steps.

The content of the target analyte can easily be calculated by spiking the sample with a known amount of labeled standard, after having established a relationship between isotopolog ratio and the intensities of analyte and standard. SIDA is less time consuming than SA and affords highly specific quantitation because it is generally based on ions diagnostic of the analyte(s) investigated. However, it requires (i) labeled standards to be available and (ii) labeling to remain very stable throughout the analytical procedure to keep the isotopic ratio unvaried. In general, labeling with $[^{13}C]$ or $[^{15}N]$ is very stable, whereas $[^{18}O]$ or $[^{2}H$ or D$]$ can be the object of losses, because of possible "exchange," (iii) a further requirement is that the equilibrium between native analyte and isotopic analog should maintain the same concentration ratio in all sample compartments. The application of SIDA to complex solid matrices, such as in the analysis of the plant volatile fraction, can be difficult, because uniform distribution of the labeled standard throughout the matrix is only possible if the sample is liquid or suspended in a suitable liquid.

3.9.3 Multiple Headspace Extraction (MHE)

MHE is a dynamic gas extraction carried out stepwise at regular time intervals to quantify volatiles in solid or complex liquid samples (Kolb and Ettre, 1997). MHE was introduced by Suzuki, Tsuge, and Takeuki (1970) and McAuliffe (1971) and developed with the fundamental contribution of Kolb and Ettre (1997). Kolb and Ettre (1997) compared MHE " … *to a repeated liquid extraction of a sample in a separatory funnel: in each step, part of the analyte present is removed until no analyte is left in the original sample* …." After the first HS, a portion of the headspace is removed, so that a new equilibrium has to be established and a smaller peak results in HS-GC analysis. With a suitable number of "extractions" and a suitable matrix amount, it is, therefore, possible to strip off all the volatile(s) from the sample to obtain an exponential decay of the peak areas and to calculate the total area of an analyte in the sample by summing the areas of the target peak in each "extraction." MHE, thus, excludes the influence of the sample matrix on the final results, that is, the matrix effect.

In routine work, a small number of extractions (generally three or four) are sufficient to predict the total area of the analyte in the sample because MHE follows a logarithmic path, making it unnecessary to continue until the analyte has been entirely removed from the sample matrix. The total area of the analyte(s) under investigation can be calculated with the following equation:

$$A_T = \sum_{i=1}^{i=\infty} A_i = \frac{A_1}{(1 - e^{-q})} = \frac{A_1}{(1 - Q)} \qquad (1)$$

where A_1 is the analyte area after the first analysis, A_T the total area of the investigated analyte, A_i the peak area obtained in the ith extraction, and $-q$ ($Q = e^{-q}$) a constant representing the exponential decline during the stepwise MHE procedure, which can be calculated from the linear regression analysis equation:

$$\ln A_i = -q(i - 1) + \ln A_1 \qquad (2)$$

The total analyte amount in the sample can then be obtained from the total area, A_T, with a calibration curve or more simply with an external standard procedure. Equation 2 shows that the total area of the analyte in the sample can be calculated from the area of the first extraction A_1 and the constant Q (calculated from the linear regression equation); conversely, a Q value for each sample should ideally be determined, making the procedure extremely time consuming for routine laboratories.

MHE can easily be extended to HCC-HS techniques (Bicchi *et al.*, 2011; Ezquerro, Pons, and Tena, 2003). The main difference, compared to conventional MHE applied to S-HS, is that the analyte is partitioned among sample, headspace, and fiber. MHS-SPME can be carried out under nonequilibrium conditions, significantly reducing the analysis time (Ezquerro, Pons, and Tena, 2003), provided that all SPME parameters (i.e., extraction time and agitation) are kept constant in each individual extraction (Martinez-Urunuela, Gonzalez-Saiz, and Pizarro, 2005). The MHE procedure can be simplified when analyte amounts fall within a reasonably limited range in relatively homogeneous samples. Under these conditions, Q and correlation coefficient (r) values are very similar, thus enabling an average Q value to be adopted for the routine determination of subsequent samples. This hypothesis was demonstrated to be correct in the authors' laboratory, with the quantitation by HS-SPME of furan and 2-methyl-furan in 150 samples of roasted coffee of different varieties, origins, and of several spices (Bicchi *et al.*, 2011). The possibility of HS quantitation with a single-area determination makes the MHE approach very rapid and highly time competitive, not least because the concentration can be calculated from the total area very quickly, through an external standard determination (Kolb and Ettre, 1997).

4 CONCLUSIONS

HS and GC used together are ideal techniques to study the complexity of the plant volatile fraction and to obtain reliable information on sample composition. Thanks to the high recovery effectiveness, reliability, and flexibility of HCC-HS techniques, the most suitable approach may be applied depending on the nature of the sample to be analyzed. The high GC separative power, advanced technology, and combination with MS have meant that these two techniques have contributed synergistically to the affirmation of the role of the volatile fraction as a fundamental parameter to study plant metabolism.

Conversely, the importance of the information provided by the volatile fraction in the quest to understand plant biochemistry and behavior and the ever increasing demand for that information have also stimulated HS and GC technique development, in particular favoring their online combination. This thrust has been aimed at obtaining reliable and representative profiles and improving component identification through effective coupling with mass spectrometry and automatic data handling and processing.

ACKNOWLEDGMENTS

REFERENCES

d'Acampora Zellner, B. D., Bicchi, C., Dugo, P., *et al.* (2008) *Flavour Fragr. J.*, **23**(5), 297–314.

Adahchour, M., van Stee, L. L. P., Beens, J., *et al.* (2003) *J. Chromatogr., A*, **1019**(1–2), 157–172.

Adahchour, M., Beens, J. and Brinkman, U. A. T. (2008) *J. Chromatogr., A*, **1186**(1–2), 67–108.

Adams, R. P. (2007) *Identification of Essential Oil Components by Gas Chromatography/Mass Spectrometry*, Allured Publishing, Illinois, USA.

Ai, J. (1997) *Anal. Chem.*, **69**(6), 1230–1236.

AMDIS (2007) Last accessed January 2014.

Ampuero, S., Bogdanov, S. and Bosset, J. O. (2004) *Eur. Food Res. Technol.*, **218**(2), 198–207.

Anderson, J. L. and Armstrong, D. W. (2003) *Anal. Chem.*, **75**(18), 4851–4858.

Anderson, J. L. and Armstrong, D. W. (2005) *Anal. Chem.*, **77**(19), 6453–6462.

Armstrong, D. W., He, L. F. and Liu, Y. S. (1999) *Anal. Chem.*, **71**(17), 3873–3876.

Arthur, C. L. and Pawliszyn, J. (1990) *Anal. Chem.*, **62**(19), 2145–2148.

Baltussen, E. (2000) New concepts in sorption based sample preparation for chromatography. PhD Thesis, Technische Universiteit Eindhoven: Eindhoven (The Netherlands).

Baltussen, E., Sandra, P., David, F., *et al.* (1999a) *J. Microcol. Sep.*, **11**(10), 737–747.

Baltussen, E., David, F., Sandra, P., *et al.* (1999b) *Anal. Chem.*, **71**(22), 5193–5198.

Baltussen, E., Cramers, C. A. and Sandra, P. J. F. (2002) *Anal. Bioanal. Chem.*, **373**(1–2), 3–22.

Bartle, K. D. and Myers, P. (2002) *TRAC-Trends Anal. Chem.*, **21**(9–10), 547–557.

Belliardo, F., Bicchi, C., Cordero, C., *et al.* (2006) *J. Chromatogr. Sci.*, **44**(7), 416–429.

Berthod, A., Ruiz-Angel, M. and Carda-Broch, S. (2008) *J. Chromatogr., A*, **1184**(1–2), 6–18.

Bicchi, C. and Maffei, M. (2012) The plant volatilome: methods of analysis, in *High-Throughput Phenotyping in Plants: Methods and Protocols – Series: Methods in Molecular Biology*, Humana Press, New York, vol. **918**, pp. 289–310.

Bicchi, C., Damato, A., David, F., *et al.* (1989) *HRC-J. High. Resolut. Chromatogr.*, **12**(5), 316–321.

Bicchi, C., Cordero, C., Iori, C., *et al.* (2000a) *HRC-J. High Resolut. Chromatogr.*, **23**(9), 539–546.

Bicchi, C., Drigo, S. and Rubiolo, P. (2000b) *J. Chromatogr., A*, **892**(1–2), 469–485.

Bicchi, C., Brunelli, C., Galli, M., *et al.* (2001) *J. Chromatogr., A*, **931**(1–2), 129–140.

Bicchi, C., Cordero, C. and Rubiolo, P. (2004a) *J. Chromatogr. Sci.*, **42**(8), 402–409.

Bicchi, C., Cordero, C., Liberto, E., *et al.* (2004b) *J. Chromatogr., A*, **1024**(1–2), 217–226.

Bicchi, C., Cordero, C., Liberto, E., *et al.* (2005) *J. Chromatogr., A*, **1094**(1–2), 9–16.

Bicchi, C., Cordero, C., Liberto, E., *et al.* (2007a) *J. Chromatogr., A*, **1148**(2), 137–144.

Bicchi, C., Cordero, C., Liberto, E., *et al.* (2007b) *J. Chromatogr., A*, **1152**(1–2), 138–149.

Bicchi, C., Liberto, E., Cagliero, C., *et al.* (2008a) *J. Chromatogr., A*, **1212**(1–2), 114–123.

Bicchi, C., Cordero, C., Liberto, E., *et al.* (2008b) *J. Chromatogr., A*, **1184**(1–2), 220–233.

Bicchi, C., Liberto, E., Matteodo, M., *et al.* (2008c) *Flavour Fragr. J.*, **23**(6), 382–391.

Bicchi, C., Blumberg, L., Cagliero, C., *et al.* (2010) *J. Chromatogr., A*, **1217**(9), 1530–1536.

Bicchi, C., Ruosi, M. R., Cagliero, C., *et al.* (2011) *J. Chromatogr., A*, **1218**(6), 753–762.

Bicchi, C., Cordero, C., Liberto, E., *et al.* (2012) Headspace Sampling in Flavor and Fragrance Field, in *Comprehensive Sampling and Sample Preparation: Analytical Techniques for Scientists*, eds. J. Pawliszyn, L. Mondello and P. Dugo, Elsevier, Academic Press, Oxford, UK, pp. 1–25.

Blumberg, L. M. (1997) *HRC-J. High Resolut. Chromatogr.*, **20**(12), 679–687.

Blumberg, L. M. and Klee, M. S. (1998a) *Anal. Chem.*, **70**(18), 3828–3839.

Blumberg, L. M. and Klee, M. S. (1998b) Theory and Practice of Fast Capillary GC. Efficiency and Speed of Analysis, in *27th International Symposium on Capillary Chromatography*, book of abstract, eds. P. Sandra and A. J. Rackstaw, Riva del Garda, Italy.

Bosset, J. O., Pillonel, L., Altieri, D., *et al.* (2004) *Mitt. Lebensmitt. unters. Hyg.*, **95**, 85.

Bovijn, L., Pirotte, J. and Berger, A. (1958) in *Gas Chromatography*, ed. D. H. Desty, London, pp. 310–320.

Bruheim, I., Liu, X. C. and Pawliszyn, J. (2003) *Anal. Chem.*, **75**(4), 1002–1010.

Cagliero, C., Bicchi, C., Cordero, C., *et al.* (2012a) *Food Chem.*, **132**(2), 1071–1079.

Cagliero, C., Bicchi, C., Cordero, C., *et al.* (2012b) *J. Chromatogr., A*, **1268**, 130–138.

Chaintreau, A. (2000) Sample preparation, Headspace techniques, in *Encyclopedia of Analytical Chemistry*, ed. R. A. Meyers, John Wiley & Sons Ltd, Chichester, UK, pp. 4229–4246.

Chaintreau, A., Joulain, D., Marin, C., *et al.* (2003) *J. Agric. Food Chem.*, **51**(22), 6398–6403.

Cicchetti, E., Merle, P. and Chaintreau, A. (2008) *Flavour Fragr. J.*, **23**(6), 450–459.

Cordero, C., Bicchi, C., Joulain, D., *et al.* (2007) *J. Chromatogr., A*, **1150**(1–2), 37–49.

Costa, R., De Fina, M. R., Valentino, M. R., *et al.* (2007) *Nat. Prod. Commun.*, **2**(4), 413–418.

Costa, R., Zellner, B. D., Crupi, M. L., *et al.* (2008) *Flavour Fragr. J.*, **23**(1), 40–48.

Cramers, C. A., Janssen, H. G., van Deursen, M. M., *et al.* (1999) *J. Chromatogr., A*, **856**(1–2), 315–329.

David, F., Gere, D. R., Scanlan, F., *et al.* (1999) *J. Chromatogr., A*, **842**(1–2), 309–319.

Deans, D. R. (1968) *Chromatographia*, **1**, 18–21.

Debonneville, C., Thome, M. A. and Chaintreau, A. (2004) *J. Chromatogr. Sci.*, **42**(8), 450–455.

Dietrich, A., Maas, B., Karl, V., *et al.* (1992) *HRC-J. High Resolut. Chromatogr.*, **15**(3), 176–179.

Dorman, F. L., Whiting, J. J., Cochran, J. W., *et al.* (2010) *Anal. Chem.*, **82**(12), 4775–4785.

Dugo, G., Dugo, P. and Mondello, L. (2002) Multidimensional Chromatography: foods, flavours and fragrances applications, in *Multidimensional Chromatography*, eds. L. Mondello, A. C. Lewis and K. D. Bartle, John Wiley & Sons Ltd., Chichester.

Eom, I. Y., Tugulea, A. M. and Pawliszyn, J. (2008) *J. Chromatogr., A*, **1196**, 3–9.

Ettre, L. S. (2002) *LCGC North America*, **20**(12), 1120–1129.

Ezquerro, O., Pons, B. and Tena, M. T. (2003) *J. Chromatogr., A*, **999**(1–2), 155–164.

Fakhari, A. R., Salehi, P., Heydari, R., *et al.* (2005) *J. Chromatogr., A*, **1098**(1–2), 14–18.

FFNSC MS Library, (2011) *FFNSC 2.0 - Flavors and Fragrances of Natural and Synthetic Compounds - Mass Spectral Database*, Chromaleont, Messina, Italy.

Fiehn, O. (2008) *TRAC-Trends Anal. Chem.*, **27**(3), 261–269.

Giarrocco, V., Quimby, B. and Klee, M. (1997) *Retention Time Locking: Concepts and Application; Gas chromatography application note*, Agilent Technologies, Wilmington, USA.

Giddings, J. C. (1987) *J. High Resolut. Chromatogr.*, **10**(5), 319–323.

Golay, M. J. E. (1957) in *Gas Chromatography*, ed. V. J. Coates, Academic Press, New York, p. 1.

Golay, M. J. E. (1958) in *Gas Chromatography*, ed. D. H. Desty, Butterworth, London, p. 36.

Gorecki, T. and Pawliszyn, J. (1995) *Anal. Chem.*, **67**(18), 3265–3274.

Grob, K. and Habich, A. (1985) *J. Chromatogr.*, **321**(1), 45–58.

Halket, J. M., Waterman, D., Przyborowska, A. M., *et al.* (2005) *J. Exp. Bot.*, **56**(410), 219–243.

Hamm, S., Lesellier, E., Bleton, J., *et al.* (2003) *J. Chromatogr., A*, **1018**(1), 73–83.

Harger, R. N., Bridwell, E. G. and Raney, B. B. (1939) *J. Biol. Chem.*, **128**, xxxviii–xxxix.

Herout, V. (1967) *Planta Med.*, **15**, 68–78.

Herout, V., Streibl, M. and Holasova, M. (1970) *Flavor Ind.*, **1**, 673–676.

Ishikawa, M., Ito, O., Ishizaki, S., *et al.* (2004) *Flavour Fragr. J.*, **19**(3), 183–187.

Isidorov, V. A. and Szczepaniak, L. (2009) *J. Chromatogr., A*, **1216**(51), 8998–9007.

James, A. T. and Martin, A. J. P. (1952) *Biochem. J.*, **50**(5), 679–690.

Jeannot, M. A. and Cantwell, F. F. (1996) *Anal. Chem.*, **68**(13), 2236–2240.

Jennings, W., Mittlefehldt, E. and Stremple, P. (1997) *Analytical Gas Chromatography*, Academic Press, San Diego, CA, USA.

Jung, M., Schmalzing, D. and Schurig, V. (1991) *J. Chromatogr.*, **552**(1–2), 43–57.

Juvancz, Z., Alexander, G. and Szejtli, J. (1987) *J. High Resolut. Chromatogr.*, **10**(2), 105–107.

Klee, M. S. and Blumberg, L. M. (2002) *J. Chromatogr. Sci.*, **40**(5), 234–247.

Kolb, B. and Ettre, L. S. (1997) *Static Headspace-Gas Chromatography, Theory and Practice*, Wiley-VCH, New York.

Konig, W. A. and Hochmuth, D. H. (2004) *J. Chromatogr. Sci.*, **42**(8), 423–439.

Kovats, E. (1958) *Helv. Chim. Acta*, **41**, 1915–1932.

Liberto, E., Cagliero, C., Sgorbini, B., *et al.* (2008) *J. Chromatogr., A*, **1195**(1–2), 117–126.

Lipinski, J. (2001) *Fresenius J. Anal. Chem.*, **369**(1), 57–62.

Liu, Z. and Phillips, J. B. (1991) *J. Chromatogr. Sci.*, **29**, 227–231.

Liu, X. Y., Pawliszyn, R., Wang, L. M., *et al.* (2004) *Analyst*, **129**(1), 55–62.

Lord, H. L., Zhan, W. Q. and Pawliszyn, J. (2010) *Anal. Chim. Acta*, **677**(1), 3–18.

Magni, P., Facchetti, R. and Cavagnino, D. (2002) Ultra fast GC with conventional instruments using direct resistively heated capillary columns, in *29th International Symposium on Capillary Chromatography*, book of abstract, ed. Sandra P, Riva del Garda (Italy).

Martin, A. J. P. (1957) , in *Vapour Phase Chromatography*, ed. D. H. Desty, Butterworths, London.

Martin, A. J. P. and Synge, R. L. M. (1941) *Biochem. J.*, **35**(12), 1358–1368.

Martinez-Urunuela, A., Gonzalez-Saiz, J. M. and Pizarro, C. (2005) *J. Chromatogr., A*, **1089**(1–2), 31–38.

Mastovska, K. and Lehotay, S. J. (2003) *J. Chromatogr., A*, **1000**(1–2), 153–180.

McAuliffe, C. (1971) *Chem. Technol.*, **1**, 46–51.

McComb, M. E., Oleschuk, R. D., Giller, E., *et al.* (1997) *Talanta*, **44**(11), 2137–2143.

McReynolds, W. O. (1970) *J. Chromatogr. Sci.*, **8**(12), 685–691.

Mondello, L., Dugo, P., Basile, A., *et al.* (1995) *J. Microcol. Sep.*, **7**(6), 581–591.

Mondello, L., Catalfamo, M., Dugo, C., *et al.* (1998) *J. Chromatogr. Sci.*, **36**(4), 201–209.

Mondello, L., Casilli, A., Tranchida, P. Q., *et al.* (2004) *J. Chromatogr. Sci.*, **42**(8), 410–416.

Mulder, M. (1991) *Basic Principles of Membrane Technology*, Kluwer, Dordrecht.

Musshoff, F., Lachenmeier, D. W., Kroener, L., *et al*. (2002) *J. Chromatogr., A*, **958**(1–2), 231–238.

Nie Y and Kleine-Benne E. 2011. Gerstel Application Note 3 (http://www.gerstel.com/pdf/p-gc-an-2011-03.pdf).

NIST 05 – NIST/EPA/NIH. (2005) Mass Spectral Library (National Institute of Standards and Technology, Gaithersburg, MD, USA.

Nowotny, H. P., Schmalzing, D., Wistuba, D., *et al*. (1989) *HRC-J. High Resolut. Chromatogr.*, **12**(6), 383–393.

Ouyang, G., Zhao, W. N. and Pawliszyn, J. (2007) *J. Chromatogr., A*, **1138**(1–2), 47–54.

Pawliszyn, J. (1997) *Solid Phase Microextraction – Theory and Practice*, Wiley-VCH, New York.

Pawliszyn, J. (2002) Solid Phase Microextraction, in *Sampling and Sample Preparation for Field and Laboratory*, ed. J. Pawliszyn, Elsevier, Amsterdam, pp. 389–478.

Payagala, T., Zhang, Y., Wanigasekara, E., *et al*. (2009) *Anal. Chem.*, **81**(1), 160–173.

Proot, M. and Sandra, P. (1986) *J. High Resolut. Chromatogr.*, **9**(11), 618–623.

Rohrschneider, L. (1966) *J. Chromatogr., A*, **39**(C), 383–397.

Rubiolo, P., Liberto, E., Sgorbini, B., *et al*. (2008) *J. Sep. Sci.*, **31**(6–7), 1074–1084.

Rubiolo, P., Sgorbini, B., Liberto, E., *et al*. (2010a) *Flavour Fragr. J.*, **25**(5), 282–290.

Rubiolo, P., Sgorbini, B., Liberto, E., *et al*. (2010b) Analysis of the plant volatile fraction, in *The Chemistry and Biology of Volatiles*, ed. A. Herrmann, Wiley, Chichester, UK.

Rychlik, M. and Asam, S. (2008) *Anal. Bioanal. Chem.*, **390**(2), 617–628.

Sandra, P., Proot, M. and Dirick, G. (1987) Considerations on selection of capillary columns for essential oil analysis, in *Capillary Gas Chromatography in Essential Oil Analysis*, eds. P. Sandra and C. Bicchi, Huetig Verlag, Heidelberg.

Schieberle, P. and Grosch, W. (1987) *J. Agric. Food Chem.*, **35**(2), 252–257.

Schomburg, G., Weeke, F., Muller, F., *et al*. (1982) *Chromatographia*, **16**, 87–91.

Schomburg, G., Husamann, H., Hubinger, E., *et al*. (1984) *J. High Resolut. Chromatogr.*, **7**, 404–410.

Schurig, V. and Nowotny, H. P. (1988) *J. Chromatogr.*, **441**(1), 155–163.

Schurig, V. and Nowotny, H. P. (1990) *Angew. Chem.-Int. Ed. Eng.*, **29**(9), 939–957.

Segal, A., Gorecki, T., Mussche, P., *et al*. (2000) *J. Chromatogr., A*, **873**(1), 13–27.

Sgorbini, B., Budziak, D., Cordero, C., *et al*. (2010) *J. Sep. Sci.*, **33**(14), 2191–2199.

Sgorbini, B., Cagliero, C., Cordero, C., *et al*. (2012) *J. Chromatogr., A*, **1265**, 39–45.

Shellie, R. and Marriott, P. J. (2002) *Anal. Chem.*, **74**(20), 5426–5430.

Shellie, R. and Marriott, P. (2003) *Flavour Fragr J*, **18**(3), 179–191.

Shellie, R., Marriott, P. and Morrison, P. (2001) *Anal. Chem.*, **73**(6), 1336–1344.

Shellie, R., Marriott, P., Zappia, G., *et al*. (2003) *J. Ess. Oil Res.*, **15**(5), 305–312.

Shellie, R., Mondello, L., Dugo, G., *et al*. (2004) *Flavour Fragr. J.*, **19**(6), 582–585.

Sidelnikov, V. N., Patrushev, Y. V. and Belov, Y. P. (2006) *J. Chromatogr., A*, **1101**(1–2), 315–318.

Silva, R. C., Aquiar, P. M. S. and Augusto, F. (2004) *Chromatographia*, **60**(11–12), 687–691.

Sisalli, S., Adao, A., Lebel, M., *et al*. (2006) *LCGC Eur.*, **19**(1), 33–39.

Smolkova-Keulemansova, E. (2000) *HRC-J. High Resolut. Chromatogr.*, **23**(7–8), 497–501.

Splivallo, R., Bossi, S., Maffei, M., *et al*. (2007) *Phytochemistry*, **68**(20), 2584–2598.

Stahl, W. H., Voelker, W. A. and Sullivan, J. H. (1960) *Food Technol.*, **14**, 14–16.

Steinhaus, M., Fritsch, H. T. and Schieberle, P. (2003) *J. Agric. Food Chem.*, **51**(24), 7100–7105.

Suzuki, M., Tsuge, S. and Takeuki, T. (1970) *Anal. Chem.*, **42**, 1705–1708.

Sybilska, D. and Koscielski, T. (1983) *J. Chromatogr.*, **261**, 357–362.

Tankeviciute, A., Kazlauskas, R. and Vickackaite, V. (2001) *Analyst*, **126**(10), 1674–1677.

Teranishi, R., Mon, T. R., Robinson, A. B., *et al*. (1972) *Anal. Chem.*, **44**, 18–20.

Theis, A. L., Waldack, A. J., Hansen, S. M., *et al*. (2001) *Anal. Chem.*, **73**(23), 5651–5654.

Tienpont, B., David, F., Bicchi, C., *et al*. (2000) *J. Microcol. Sep.*, **12**(11), 577–584.

Van Asten, A. (2002) *Trends Anal. Chem.*, **21**, 698–708.

Van den Dool H. (1974) Standardisation of gas chromatographic analysis of essential oils. PhD Thesis, University of Groningen: Groningen (The Netherlands).

Van den Dool, H. and Kratz, P. D. (1963) *J. Chromatogr.*, **11**, 463–471.

Wahlroos, O. (1963) *Ann. Acad. Sci. Fenn. Ser A. II, Chem.*, **122**, 1.

Wang, L. M., Lord, H., Morehead, R., *et al*. (2002) *J. Agric. Food Chem.*, **50**(22), 6281–6286.

Wang, A. P., Fang, F. and Pawliszyn, J. (2005a) *J. Chromatogr., A*, **1072**(1), 127–135.

Wang, Y. X., O'Reilly, J., Chen, Y., *et al*. (2005b) *J. Chromatogr., A*, **1072**(1), 13–17.

Yang, M. J., Harms, S., Luo, Y. Z., *et al*. (1994) *Anal. Chem.*, **66**(8), 1339–1346.

de Zeeuv, J. and Luong, J. (2002) *Trac-Trends Anal. Chem.*, **21**(9–10), 594–607.

Zhang, Z. Y. and Pawliszyn, J. (1993) *Anal. Chem.*, **65**(14), 1843–1852.

Zlatkis, A., Lichtenstein, H. A. and Tishbee, A. (1973) *Chromatographia*, **6**, 67–70.

Zuba, D., Parczewski, A. and Reichenbacher, M. (2002) *J. Chromatogr. B*, **773**(1), 75–82.

11

Analysis of Natural Products by Capillary Electrophoresis and Related Techniques

Markus Ganzera and Anja Krüger

Institute of Pharmacy, Pharmacognosy, University of Innsbruck, Innsbruck, Austria

1 INTRODUCTION

In the United States Pharmacopeia, electrophoresis is defined as the migration of charged electrical species, suspended or dissolved in an electrolyte, through which an electric current is passed. Accordingly, the German physicist Ferdinand Friedrich von Reuss was the first to describe such a phenomenon in 1807, when he noticed that clay particles dispersed in water start to migrate in an electric field (Zaitseva, 2001). Classical electrophoresis as reported by Tiselius in the first half of the twentieth century utilizes gels or impregnated paper strips for separation, an approach with several major disadvantages (e.g., sometimes problematic *in situ* quantification or extended analysis time). In order to overcome these limitations, separations in open tubes were attempted as detection, high electric field strengths, and cooling are easier to achieve. Hjerten (1967) was the first one to report on successful separations in 3-mm glass tubes, but in order to control thermal convection they needed to be under rotation during analysis. The years to follow brought major improvements by the studies of Virtenen [advantages of columns with small internal diameter (ID)], Everaerts (on-column UV detection in 200 μm Teflon capillaries), and Jorgenson and Lukacs, who utilized small diameter (<100 μm ID) fused-silica capillaries for the first time (Jorgenson and Lukacs, 1981). Their pioneering work on theory and application can be considered

the beginning of modern capillary electrophoresis (CE or HPCE, high performance capillary electrophoresis), a starting point from which the technique has impressively evolved. Just to name the invention of micellar electrokinetic chromatography (MEKC or MECC) by Terabe *et al*. (1984), the first report on capillary electrophoresis-mass spectrometry (CE-MS) in the year 1987 (Olivares *et al*., 1987), or the availability of commercial instruments in 1989 as further milestones. Currently, the following techniques are compiled under the term CE, capillary zone electrophoresis (CZE), MEKC, microemulsion electrokinetic chromatography (MEEKC), capillary electrochromatography (CEC), capillary gel electrophoresis (CGE), capillary isotachophoresis (CITP), and capillary isoelectric focusing (CIEF) (Natishan, 2005). Owing to many desirable features, some of them are already used in routine analysis, others are still in developmental stage and mainly of academic interest. Yet, for a relatively young separation technique, CE plays an important and indispensable role in analytical sciences already. After a brief introduction to separation theory and instrumentation, the relevance of CE for natural products analysis is described and discussed in this contribution.

2 SEPARATION THEORY

In contrast to many other separation techniques, the fundamentals of CE can be described in rather

simple terms. As stated at the beginning, ions or charged solutes will move in an applied electric field, anions toward the anode and cations to the cathode. Their migration direction and speed (migration velocity) are depending on electric field strength and the analytes own electrophoretic mobility. The latter is influenced by charge of the ion, ion radius, and solution viscosity; higher charge and smaller size will result in greater mobility and vice versa (Weinberger, 2000). When glass capillaries are used, this mobility is superimposed by another unique driving force, the electroosmotic flow (EOF). Above a pH value of 3.0, the silanol groups of uncoated silica capillaries are increasingly deprotonated (Karger and Foret, 1993). Accordingly, when a buffer (background electrolyte, BGE) is present, buffer cations will form a double layer at the negatively charged inner wall. The double layer is typically very thin (a few hundred nanometers), composed of a rigid (Stern layer) and diffuse layer, and the corresponding potential difference is called *zeta potential* (Karniadakis, Beskok, and Aluru, 2005). When voltage is applied, cations of the diffuse layer are free to move, carrying the whole bulk of buffer to the cathode. In contrast to pressure-driven separation techniques such as HPLC, the resulting flow profile is flat and not parabolic. This is one of the reasons for superiority of CE in terms of separation efficiency (Altria, 2010). Another one is additionally diminished band broadening, which is only caused by longitudinal diffusion; other factors mentioned in the Van Deemter equation, for example, mass transfer, are minimized because of the use of open capillaries with small diameter (Li, 1992).

For natural products analysis, the most common CE-techniques are CZE, MEEK, MEEKC, and CEC; they are described in detail in Chapter 3 of this article. Here, the generation of an EOF is desired, as it speeds up migration, which is the result of EOF and compound-specific migration. In case detection is achieved at the cathodic side, cations (separated based on their mobility) will elute first, followed by neutral, not separated species (only for CZE). Anions can be detected as long as the EOF is stronger than their native migration. If required, the EOF can also be reverted by the addition of cationic surfactants such as cetyltrimethylammonium bromide (CTAB). For other approaches such as CGE, CITP, or CIEF, an EOF is usually disadvantageous and suppressed using coated capillaries (e.g., with methylcellulose) or adding special agents (e.g., *s*-benzylthiouronium bromide) (Weinberger, 2000).

Unfortunately, CE separations and particularly those of natural products in plant extracts are not as facile as they possibly seem. Even if the principles are simple, CE method development is usually tedious. It involves the selection of separation and detection mode, optimization of separation conditions (type of buffer, its pH value, buffer molarity, temperature, voltage, and need of modifiers such as organic solvents or cyclodextrins), and sample preparation and capillary treatment within analyses (Wätzig, Degenhardt, and Kunkel, 1998). Each parameter has to be optimized and controlled; for example, the pH value of the buffer will have an influence on charge of the ion, its solubility in the BGE and EOF/zeta potential, or the separation temperature is relevant as the viscosity of many solvents changes by 2% per degree Celsius. Not to forget the possible electrostatic adsorption of constituents on the capillary surface, causing problems in reproducibility and peak shape (Karger and Foret, 1993). In such a case, sample preparation and capillary preconditioning need to be optimized for obtaining reproducible results.

3 RELEVANT SEPARATION MODES

3.1 Capillary Zone Electrophoresis

CZE, sometimes also termed *free solution capillary electrophoresis* (*FSCE*), is the simplest and most common CE mode. Its separation mechanism is, as mentioned in the previous chapter of this article, based on different effective mobilities of the analytes at a given pH value (Gotti, 2011). It only will permit the separation of cations and/or anions, and therefore pH control of the buffer solution is usually the most crucial parameter to adjust. Taking an acidic compound (e.g., phenolic acid) as an example, at low pH values (<pK −2), the compound remains basically noncharged and cannot be separated by CZE from other neutral analytes, at its pK value half of it is ionized, and only at a pH $\geq$ pK +2 all molecules are deprotonated. Accordingly, such a buffer pH should be selected for analysis, which, in this case, also assures solubility in aqueous buffers and sufficient EOF. Concerning the type of buffer, the analyst can choose from a long list, varying in many factors such as pH range, buffer capacity, solubility, or conductivity. Common electrolyte systems include phosphate (pH 1.1–3.1 and 6.2–8.2), acetate (pH

3.8–5.8), borate (pH 8.1–10.1), or Tris-buffers (pH 7.3–9.3) (Li, 1992). Their influence on the separation can further be modified by buffer molarity or the addition of modifiers and chiral selectors. Typical examples for the latter are cyclodextrins, compounds that are able to form stereoselective analyte-selector complexes, and thereby enable the separation of enantiomers (Scriba, 2008). Adding an organic solvent to the aqueous BGE serves two purposes, to modify selectivity and enhance solubility of the analytes. It even can end up with the full replacement of water by an organic solvent, an approach that is called *nonaqueous capillary electrophoresis* (*NACE*). In the early days of CE, there was a lack of interest concerning this technique, possibly because of instrumental problems (evaporation of solvents during analysis). Today, these problems are solved and NACE a recognized alternative to aqueous CZE, especially as respective buffers usually generate low electric current (Cherkaoui, Geiser, and Veuthey, 2000). Joule heating, resulting in air bubble formation in the buffer and thus instable electric current, is less limiting, and higher voltages can be applied for faster and more efficient separations. The choice of solvent, which is to be used, depends on factors such as solvent properties, chemical stability, instrumental demands (e.g., UV transparency), viscosity, and volatility (Geiser and Veuthey, 2009).

3.2 Micellar Electrokinetic Chromatography

MEKC is an elegant way to overcome a major limitation of CZE, which is only being able to resolve charged analytes. It can be described as micellar solubilization by means of a surfactant (El Deeb, Iriban, and Gust, 2011). The latter is composed of a hydrophobic, nonpolar hydrocarbon tail and a hydrophilic, polar head with anionic, cationic, nonionic, or zwitterionic character. Once a certain concentration is reached in the solution (CMC, critical micelle concentration), the monomers aggregate to form micelles, structures that exist in dynamic equilibrium with the monomeric tensides (Unger, 2009). These micelles act as pseudo-stationary phase and, except those of nonionic surfactants, display external charge and electrophoretic mobility. Depending on their distribution coefficient, neutral analytes will now be allocated between micelle and solution and therefore migrate differently depending on their chemical structure (Terabe, 2008). Charged

species can be separated too as they move in the electric field anyway.

The most commonly used tenside in MEKC is sodium dodecylsulfate (SDS), an anionic surfactant with a CMC of 8.1 mM in water. Respective micelles move toward the anode and not to the cathode, where detection is usually achieved. Anyway, as long as the net movement is in direction toward the detector, neutral analytes will be detectable. Other previously utilized detergents include sodium octane sulfonate (anionic), cetyltrimethyl ammonium chloride (cationic), cocamidopropyl betaine (zwitterionic), polyoxyethylene sorbitan monolaureate (Tween 20, nonionic; they only can be used in combination with others), or bile salts such as sodium cholate (El Deeb, Iriban, and Gust, 2011). More recently, the use of polymers, liposomes, or nanoparticles (gold and carbon) as pseudo-stationary phases in MEKC has been described as well (Liu, 2009; Shpak, Pirogov, and Shpigun, 2004). Regardless what actually is used, the molecular structure of the surfactant has a pronounced effect on the analyte–micelle interaction and therefore on separation selectivity and efficiency. Other parameters that also have to be considered are separation temperature, pH value of buffer, and the possible need of modifiers such as organic solvents, chiral additives, or ion pairing reagents. Sample preconcentration techniques such as sweeping or stacking (Section 4) are feasible with MEKC buffers as well; on the other hand, their use for CE-MS is limited to a large extent as most surfactants are nonvolatile. There have been attempts to solve respective problems by partial filling technique, instrumental modifications (e.g., APCI instead of ESI interface), or the use of specific micelle polymers; yet, user-friendly solutions for routine analysis are still in demand (Silva, 2007).

3.3 Microemulsion Electrokinetic Chromatography

At first sight, MEEKC seems very similar to MEKC, yet there are distinct differences between these two CE modes. Microemulsions are defined as stable, isotopically clear solutions of oil (e.g., octane) and water stabilized by a surfactant and cosurfactant (Ryan *et al.*, 2011). Accordingly, oil-in-water (O/W) and water-in-oil (W/O) systems have to be differentiated. For O/W systems, SDS can be used as surfactant at concentrations well above the

CMC, and a straight alcohol such as butanol serves as cosurfactant. The surfactant lowers interfacial tension, and the cosurfactant reduces electrostatic repulsion between the surfactants head groups, so that stability of the emulsion is strongly enhanced (Ryan *et al*., 2009). As a result, small droplets of oil (size ≤ 10 nm, pseudo-stationary phase) are formed, which, depending on the type of surfactant used, are charged. The analytes (neutral and charged) are selectively partitioned between oil and aqueous phases; for example, the more hydrophobic they are, the more they will be in the oil droplet. Thus, the separation principle resembles that of MEKC to a large extent, but one major advantage of MEEKC is a greater solubilizing power for polar and nonpolar compounds at the same time. It has also been suggested that compounds can partition more effectively in a microemulsion droplet than a more rigid micelle. This leads to higher rates of mass transfer and more efficient separations (Ryan *et al*., 2011).

Compared to MEKC, this method offers more options for fine-tuning, thus, it is an interesting but challenging approach (Ha, Hoogmartens, and Van Schepdael, 2006); extensive preformulation studies are usually required, investigating amount and hydrophobicity of the drug(s) on one side, and nature, combination, and proportion of the formulation components on the other side (Furlanetto *et al*., 2011). However, once this is optimized, a versatile and powerful procedure is at hand, as indicated by recent MEEKC-MS applications (Bytzek *et al*., 2010) or separations in multiplexed (48 capillaries in parallel) format (Lynen *et al*., 2011).

3.4   Capillary Electrochromatography

CEC is unique in a sense that it combines selectivity of HPLC with efficiency of CE. Capillaries with usually 75–150 μm ID are filled with stationary phase, and separation is mainly based on chromatographic interaction. Yet, the buffer is pushed through the capillary by applying voltage and the resulting EOF. Fundamental studies in this respect have been performed by Knox and coworkers, who showed that particle diameters as low as 1.5 μm have no reducing effect on EOF (Knox and Grant, 1991). Compared to CE, the latter is more complex to explain for CEC. Here, the externally applied electrical field interacts with space charges in the electrical double layer of the stationary phase, caused by ionizable surface functionalities, the adsorption of buffer components or molecular/particulate coatings (Ganzera and Nischang, 2010). This explains the fact why usually mixed mode separations are observed in CEC, as the same surface that is responsible for generating the EOF may also provide chromatographic selectivity (Freitag and Hilbrig, 2007).

There is a vast choice of stationary phases suitable for CEC use, even if only a few ready-made CEC capillaries are commercially available till date. The options include packed columns, which are filled with regular LC stationary phases of small particle size, monolithic materials, or open tubular columns. The latter are capillaries with modified inner surface, which, for example, is covered by a thin porous layer, molecularly imprinted polymers, or nanoparticles. Compared to packed beds, they are easier to prepare (e.g., no frits are required) and versatile, but their separation efficiency is low because of a reduced phase ratio (Yang, Zhao, and Li, 2010). A good alternative in this respect are porous monoliths. They can be either silica or polymer (styrene/acrylate) based, are comparatively easy to prepare by *in situ* polymerization directly in the capillary, are extremely flexible in their separation properties because of the large variety of functional precursors, and can be fabricated in diverse formats (Eeltink and Kok, 2006). Accordingly, these materials are in the main focus of research, as indicated by interesting contributions such as surface modifications for enantioselective CEC (Lämmerhofer and Gargano, 2010), CEC-MS by photoionization (Gu and Shamsi, 2010), or the preparation of organic-silica hybrid monoliths (Wu *et al*., 2011). Regardless of these promising attempts, CEC is still in some kind of experimental stage, with little impact on routine analysis. Possible reasons for this (current) status might be the facts that separation mechanisms are difficult to predict, the technique is vulnerable to methodological variables (e.g., sample composition), and, even easy at first sight, a lot of experience is required in the fabrication of columns with reproducible performance.

4   INSTRUMENTAL AND OPERATIONAL OPTIONS

CE instruments comprise only a few essential components, which are two buffer and one sample vial, two electrodes connected to a high-voltage unit, a

separation capillary, and a suitable detection device. For CEC, the same instrumentation can be used, except that both buffer vials have to be pressurized during analysis in order to prevent air bubble formation and therefore fluctuations in electric current. Otherwise analysis is performed in a similar way, meaning the capillary has to be prepared for analysis by flushing with suitable solvent(s) (in CE, an NaOH solution is usually included in this step for deprotonation of the silanol groups) and buffer (equilibration). The sample solution is injected and by applying high voltage via two electrodes inserted in the same buffer vials as the capillary tips (inlet at injection side and outlet at detection side), the compounds start migrating toward the detector. Detection is achieved right through the capillary (on-line) most of the time, so that depending on the type of capillaries used (e.g., UV transparent or nontransparent) the coating has to be removed at the detector window by burning or treatment with acid. Considering that microscale separations are utilized (injected sample volumes are usually in the range 10–50 nL and the total column volume is a few microliter only), special emphasis has to be put on quality and purity of solvents. Accordingly, for satisfactory results, all solutions including sample need to be membrane filtered and degassed before use and replaced after a few analyses.

4.1 Injection

In CE, there are two injection modes with practical relevance, hydrodynamic and electrokinetic injection. In the first case, the sample is injected by applying pressure at the inlet side of the capillary or vacuum at the outlet or through a height difference between sample and outlet buffer vial (siphoning). For electrokinetic injection, voltage is applied for a short time when the sample vial is placed at the inlet side (and buffer at the outlet); depending on their electrophoretic mobility, the resulting EOF will transport sample constituents into the capillary. Thus, in this mode, an inherent discrimination of the injected analytes takes place.

4.2 Analysis

As mentioned earlier, CE separations can be performed in different modes and formats. Typically

bare fused glass capillaries with an ID between 50 and 100 µm are employed, at variable (approximately 30–80 cm) or instrument-dependent length. In order to suppress possible analyte – capillary adsorption effects, coated capillaries are an alternative. They can be prepared by dynamic (coating agents are continuously present in the BGE) or static (covalent attachment of the coating agent to the capillary wall) coating and optimize the wall properties for particular analytes, for example, for protein analysis cationic coatings are favorable (Huhn *et al.*, 2010). Most commercial instruments are equipped with an autosampler and provide voltages of up to 30 kV, either in positive (cathode at the outlet electrode) or in negative direction (anode at the outlet). The capillaries are placed in temperature-controlled cassettes (air or liquid cooling), assuring an efficient heat dissipation. Buffers generating low electric currents ($\leq$150 µA) during analysis should be preferred, as otherwise Joule heating and related problems may occur. The majority of reports describe separations in analytical scale, but by changing the outlet vial(s) during analysis, CE can also be used for fractionation.

4.3 Detection

The standard detection technique in CE, regardless which mode is selected, is definitely UV–vis photometry in the form of a diode array or fixed wavelength detector. They are robust, show a good linear range, and are fairly sensitive. Yet, as detection is usually performed on-line through the separation capillary of small ID (and therefore short optical path length) and the injected sample volume is extremely small, sensitivity is lower than for other analytical techniques; detection limits are typically two to three orders of magnitude worse than LC (Breadmore, Dawod, and Quirino, 2011; Suntornsuk, 2010). Capillaries with extended light path (bubble cells) or z-shaped detector cells can be used to solve this problem. Other options include the use of specific sample preconcentration techniques (Section 4.4) or alternative detectors. Concerning the latter electrochemical (EC), chemiluminescence (CL), conductless conductivity (C^4D), or fluorescence detectors (FD) as well as a mass spectrometer (MS) are possible options. Each of them has unique characteristics and applications, for example, C^4D is well suited for nonchromophore species such as inorganic and

organic ions (Zemann *et al*., 1998), and CL detection is more sensitive but applicable to certain compounds only. For example, primary and secondary amines are not good coreactants of $Ru(bpy)_3^{2+}$, the most common CL emitter, and therefore, they cannot be determined precisely (Guo, Fu, and Chen, 2011). The use of CE-MS has gained popularity in the recent past even its routine application has not kept pace with technical and instrumental developments (Timerbaev, 2009). The choice of electrolytes for separation is limited as nonvolatile buffer salts usually impair detection and should be avoided. However, the advantages of MS still prevail. They include higher sensitivity and the assignment or identification of compounds based on their specific mass-to-charge ratios; for exact qualitative and quantitative determinations, the baseline separation of each compound is not required. Other previous drawbacks such as interface design have been resolved by now. Currently, sheath flow interfaces are the most popular ones. At the end of the capillary (in the MS ionization source), an additional flow of liquid (sheath or make up liquid) is added by an external pump, assuring an electrical contact between BGE and electrode and enhancing ionization. This technique creates no dead volumes but results in sample dilution; sheathless interfaces with conductive emitter tips or metal sleeve junctions have been described already but are not widely used (Maxwell and Chen, 2008). Among the sources suitable for CE-MS coupling are electrospray ionization (ESI), atmospheric pressure chemical ionization (APCI), and atmospheric pressure photoionization (APPI). APPI is a more recent approach and a good alternative for compounds that are not amendable by other ionization techniques. Here, ionization is achieved by a photon beam from a kypton or a xenon lamp and the help of a dopand, a preferably ionized substance (Marchi, Rudaz, and Veuthey, 2009). The choice of mass analyzer is flexible too, as quadrupole, ion-trap, time-of-flight (TOF), and Fourier transform-ion cyclotron resonance MS instruments have successfully been utilized in combination with CE (Suntornsuk, 2010).

4.4 Sample Preconcentration

A major disadvantage of CE is its generally lower sensitivity compared to other approaches such as HPLC, especially when the most common UV–vis detectors are used (see above). Simply increasing sample volume or concentration does not solve the problem as respective systems are easily overloaded, and deteriorated separations or distorted peaks are observed. Nevertheless, the injection step offers interesting options that can lead to drastically enhanced sensitivity. Most of them are based on the velocity change of the analytes between sample zone and separation solution zone (Simpson, Quirino, and Terabe, 2008). For example, in field-enhanced sample stacking, conductivity of the injected sample solution is lower than the separation buffer (Figure 1a). Owing to the fact that field strength is higher at lower conductivity, the sample zones are focused during analysis and peak heights increase. Another powerful on-line sample preconcentration technique is sweeping. It can be described as a procedure in which analytes are accumulated by a pseudo-stationary phase (see MEKC or MEEKC) penetrating the sample zone. Conductivity of buffer and sample solution should be the same, and as shown in Figure 1b, a comparatively large sample volume can be injected. Once voltage is applied, micelles from the buffer move through the sample zone; they concentrate the analytes in a narrow zone, so that separation itself can take place depending on the type of buffer used. The obtained results are impressive and sensitivities can be increased 5000-fold (Aranas, Guidote, and Quirino, 2009); even concentration factors up to 3,900,000 have been reported using this approach (Kitagawa *et al*., 2006). A rather new approach is SPE-CE (solid-phase extraction capillary electrophoresis), which has been described in in-line (the SPE sorbent is incorporated in the CE capillary) and on-line modes (the SPE column is separated but interfaced with a CE capillary). The applicable materials for extraction include monoliths, molecularly imprinted polymers, or impregnated membranes positioned between two capillaries (Ramautar, Somsen, and de Jong, 2010); this technique is still in developmental stage and applications on natural products are limited.

4.5 Novel Trends

Concerning instrumental innovation, there is a clear tendency toward further miniaturization. A good example is microchip CE, where separations are performed in channels with usually 10–100 μm ID etched in a planar substrate. Because of chemical stability and the EOF generation, glass is still the
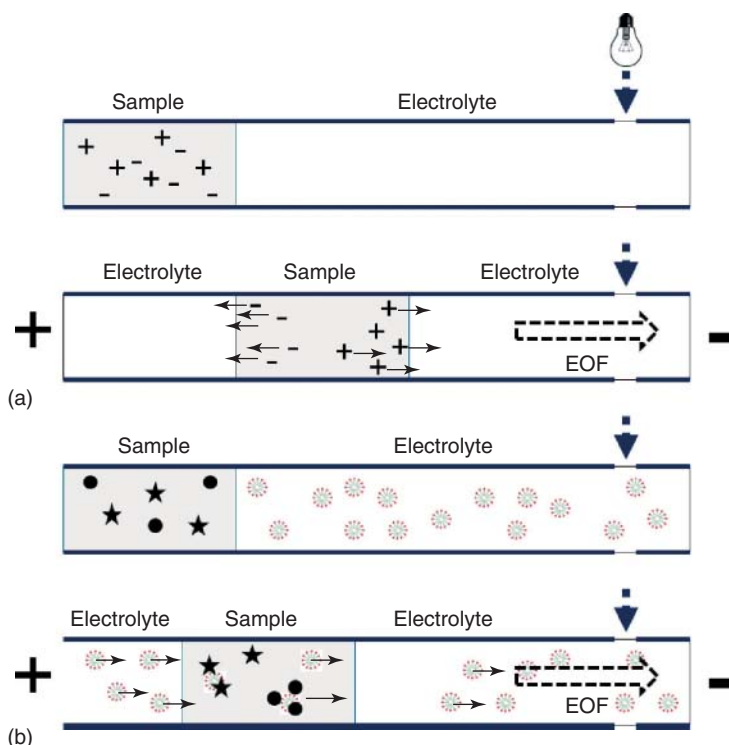
Figure 1 Sample preconcentration techniques in CE: (a) stacking due to differences in conductivity of sample solution and BGE; (b) sweeping due to a pseudo-stationary phase migrating through the sample zone.

preferred material for chips, but certain polymers such as poly(dimethylsiloxane) can be used as well. EC detection is the standard for this format as it is sensitive and can be miniaturized without loss of performance (detection limits close to 100 nM have been reported). The basic design of a CE microchip is shown in Figure 2. In its simplest form, injection is performed in the unpinched approach, in which high voltage is applied to the sample reservoir with the detection reservoir held at ground; owing to electrokinetic forces, the sample is introduced into the separation channel. After a few seconds, high voltage is switched back to the buffer reservoir and separation can start. The benefits of this design are high speed and throughput, minute solvent consumption, and the possibilities for portable/disposable devices (Martin, 2006; Suntornsuk, 2010). Another innovative trend is the use of carbon nanotubes (CNTs) or boron-doped diamond (BDD) films as electrodes for EC detection on chips or in conventional CE (Chen, 2007). These materials show unique properties, for example CNTs provide a large active surface on small

dimension electrodes, enhance electronic transfer, and possess strong sorption capacity; their practical applicability for the analysis of diverse analytes including natural products (e.g., isoflavones, vitamins, and dietary antioxidants) has been reported (Escarpa *et al*., 2008).

Other recent developments in the fields of detection, sample preconcentration, or novel materials have been described in respective chapters of this manuscript already. What additionally should be mentioned is immunoaffinity capillary electrophoresis (IACE). Its unique feature is an analyte concentrator-microreactor device (AMC), which is connected online to the CE separation capillary (Guzman and Phillips, 2011). The AMC contains a piece of capillary filled with beads coated with antibodies or other affinity molecules. When the sample is introduced through a separate transport tube in cross-sectional direction, matching compounds are retained, and later on they can be assayed by CE. Therefore, IACE is especially useful in drug screening and the search for novel biomarkers.
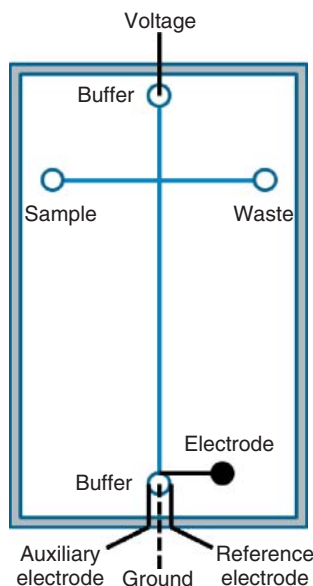
Figure 2 Schematic design of a CE microchip with electrochemical detector.

5 CHARACTERISTICS OF TECHNIQUE

Even they have been discussed individually earlier, a brief recapitulation of the major pros and cons of CE will be helpful to describe the technique in a nutshell. Main advantages are (i) high separation efficiency and selectivity, combined with short analysis time; (ii) versatility in respect to analytes, separation modes, and instrumentation; and (iii) small sample and solvent consumption. The downsides on the other hand are (i) robustness, especially when complex samples are assayed; (ii) sensitivity, in particular for UV detectors; and (iii) handling and availability of materials (e.g., CEC columns and specific interfaces). Some of these disadvantages can be minimized by optimizing sample preparation or changing the mode of detection; others are just the same as for any microscale separation technique that is relatively new. This indicates that the advantages of CE, once a system is truly optimized, should actually prevail. Whether this is a theoretical statement or one of practical relevance will be verified in the chapters to follow, which summarize CE applications for the analysis of natural products.

6 APPLICATIONS

6.1 Alkaloids

Owing to the fact that most alkaloids contain alkaline nitrogen atoms in their scaffold, this class of compounds is especially suitable for analysis by electro-driven separation techniques such as CE. Accordingly, a large number of respective reports can be found in literature. Taking one prominent alkaloid drug, opium, as first example already indicates the versatility in the different methods described so far. Several manuscripts report on the successful use of "simple" CZE to separate alkaloids such as thebaine, morphine, papaverine, and codeine with acetate buffers in the acidic pH range [e.g., a 3 : 7 mixture of 100 mM sodium acetate buffer with pH 3.1 in methanol, applied voltage and detection wavelength were 15 kV and 224 nm, respectively (Reddy *et al.*, 2003)]. Other groups used a comparable setup but added chiral selectors such as 5 mM α-CD (Fakhari *et al.*, 2010) or the mixture of 25 mM hydroxypropyl-β-CD and 75 mM dimethyl-β-CD to the BGE to enhance selectivity (Lurie *et al.*, 2003). These methods were well suited to differentiate different *Papaver* species (*P. dubium* L., *P. fugax* Poir., *P. glaucum* Hort. Leichtlin, and *P. tenuifolium* Bois. & Hohen.) and to determine major alkaloids with adequate sensitivity (LOD 0.2 μg mL^{-1}) and precision [relative standard deviation below 2.9% (Fakhari *et al.*, 2010)]. Options to enhance sensitivity are CL detection or CE-MS. Regarding the first, Gao and colleagues have described an interesting approach using a buffer containing an ionic liquid (1-ethyl-3-methylimidazolium tetrafluoroborate, EMImBF$_4$). It was added at a concentration of 8 mM to a 25-mM borax solution with pH 9.18; detection itself was achieved with tris(2,2-bipyridyl)ruthenium(II) as CL reagent (Gao *et al.*, 2006). With an analysis time of <6 min, the determination of four major opium alkaloids in real samples was possible, with detection limits ranging from 10^{-9} M (morphine) to 10^{-6} M (narcotine). For more details regarding this matter, the review by Francis *et al.*, 2008 is suggested, which gives an excellent overview to the determination of poppy alkaloids by CL detection. A CE-MS study of using gold nanoparticle-coated capillaries for the separation of diverse *Papaver* alkaloids [plate numbers close to 500,000 (Zhang *et al.*, 2009)], their analysis

by a chitosan-modified open tubular capillary (Zhou et al., 2010) or an assessment by monolithic CEC (Lin et al., 2007) are mainly of theoretical interest as practical applications are missing.

Tropane alkaloids such as atropine, scopolamine, and anisodamine have been investigated by CE as well. Yuan et al., 2010 described the analysis of these three constituents by NACE and dual detection. The electrophoretic buffer comprised 2.5 mM tetrabutylammonium perchlorate, 1 mM acetic acid, and 20 mM sodium acetate in a mixture of acetonitrile (ACN) and 2-propanol (8 : 2). This BGE facilitated the separation of the alkaloids in 5 min (18-cm bare-fused silica capillary with 25 μm ID, 20 kV), and their simultaneous detection by EC and CL detection. The latter showed to be less sensitive (e.g., LOD for scopolamine 5 vs 50 μM, respectively). This range is comparable to similar CE-UV applications, but less sensitive than another CE-CL assay for the same substances (Li, Chun, and Ju, 2007). In an aqueous 50 mM phosphate buffer (pH 5.0) containing 20% tetrahydrofuran (THF), the observed LOD for scopolamine was 0.2 μM; in both studies, $Ru(bpy)_3^{2+}$ was used as CL emitter. In another study, two different MS detectors, TOF and ion trap, were utilized for the analysis of *Atropa belladonna* L. by CE-MS (Arraez-Roman et al., 2008). The authors used an alkaline ammonium acetate buffer for separation and detected seven alkaloids in positive ESI mode. The compounds were assigned in extracted ion current mode using CE-ESI-TOF-MS, and the ion-trap instrument was used for MS/MS experiments. This study presented qualitative data only, but owing to excellent selectivity, the author considered this approach to be a valuable addition to existing LC-MS or GC-MS procedures.

Among the most toxic alkaloids are aconitine (in *Aconitum napellus* L. or *Aconitum carmichaelii* Debeaux) and strychnine (*Strychnos nux-vomica* L.). Despite their potential danger, respective drugs are used in traditional Chinese medicine (TCM), for example *Aconitum* preparations to alleviate pain, rheumatism, and neurological symptoms. Before use, the crude aconite roots are processed by prolonged boiling, heating, or steaming, so that their toxicity is within an acceptable limit. Nevertheless, to avoid any toxic incidents, stringent quality control is required; CE showed to be a suitable approach for this purpose. Six alkaloids (aconitine, mesaconitine, hypaconitine, and their benzoyl derivatives) could be determined by CZE in *A. carmichaelii* and *Aconitum kusnezoffii*

Rchb., two species listed in the Chinese Pharmacopeia (Song et al., 2010). The optimized running buffer comprised 200 mM Tris, 150 mM percholic acid, and 40% 1,4-dioxane. A nonaqueous system (60 mM ammonium acetate in methanol, with 0.5% acetic acid and 15% ACN) has been described as well (Li et al., 2004b). At a detection wavelength of 214 nm, three major alkaloids could be determined in two TCM preparations (*Chuanwu* and *Caowu*). The highest alkaloid levels were found in *Chuanwu* for hypaconitine (450 mg kg^{-1} root). Another study emphasized on the benefits of FASS concentration for the analysis of aconite alkaloids (Hu, Cui, and Liu, 2010). By injecting a short water plug at the column inlet, sample stacking was achieved, resulting in a 5–142-fold increase in detection sensitivity (e.g., LOD for hypaconitine 0.02 μg mL^{-1}). The running electrolyte was prepared by dissolving 35-mM 1B-3MI-TFB (1-butyl-3-methylimidazolium tetrafluoroborate; ionic liquid) in water and adjusting the pH to 8.5 (Figure 3). For the determination of strychnine and brucine, the major alkaloids in *S. nux-vomica*, a 100-mM NaH_2PO_4 solution with pH 4.5, showed to be an ideal buffer (Zhang et al., 2003); applied voltage, temperature, and detection wavelength were 20 kV, 25°C, and 214 nm, respectively. As for the previous report, sensitivity was enhanced by injecting a short water plug. In a more recent study, comparable separation conditions were reported (40 mM ammonium acetate buffer with pH 3.64, 25 kV, and 203 nm), but the number of analytes was extended to four (strychnine, brucine, and respective N-oxides). Sensitivity was typical for CE with LOD values ranging from 1.0 to 2.5 μg mL^{-1}, and the authors observed that by fermenting the drug the major alkaloids are converted to the respective N-oxides (Yang et al., 2011).

Isoquinoline alkaloids are found in many important medicinal plants, for example in *Hydrastis canadensis* L. (berberine and hydrastine), *Chelidonium majus* L. (sanguinarine, coptisine, and protopine), or different *Sophora* (matrine, oxymatrine, and sophocarpine) and *Coptis* (berberine, coptisine, and palmatine) species. All of them have successfully been assayed by CZE using similar aqueous acidic BGE systems (pH range from 2.5 to 4), sometimes containing an organic solvent as modifier (e.g., 10% ACN or 3.3% isopropanol) or different cyclodextrins as selectivity enhancer (e.g., 1% 1-hydroxypropyl-β-CD). The preferred detection technique was UV in the short wavelength range from
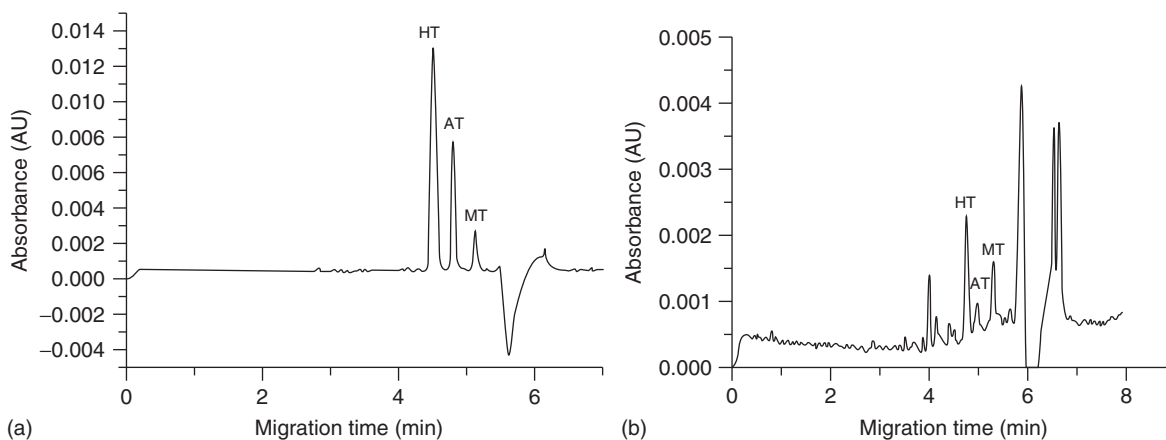
Figure 3 CE separation of hypaconitine (HT), aconitine (AT), and mesaconitine (MT) using an ionic liquid electrolyte system ((a): standard mixture; (b): *Aconitum carmichaelii* extract). (*Source*: X. Hu, *et al.* (2010). Reproduced by permission of Springer.)

200 to 225 nm. Another uncommon alternative was described by Kulp *et al*. (2011). They used ultraviolet light-emitting diode-induced native fluorescence (UV-LEDIF) detection for the determination of alkaloids in *C. majus*. Excitation was achieved by an LED at 280 nm, and fluorescence recorded with a wideband filter from 341 to 600 nm. This setup permitted the baseline separation of seven alkaloids in crude plant extracts (no sample cleanup was required) with excellent sensitivity (e.g., LOD for protopine 0.05 µg mL^{-1}). The method was validated

and the obtained quantitative results were in good agreement to HPLC-UV data (Figure 4). Besides CZE, isoquinoline alkaloids were also assessed by NACE and CEC. Concerning the latter, the described stationary phases were either a silica-based monolith, prepared by sol–gel process and chemical modification using 3-mercaptopropyltrimethoxysilane (Xie *et al*., 2005), or an acrylamide-based monolith with strong cation-exchange properties (Dong *et al*., 2007). Both were used for the separation of *Coptis chinensis* Franch. extracts to indicate
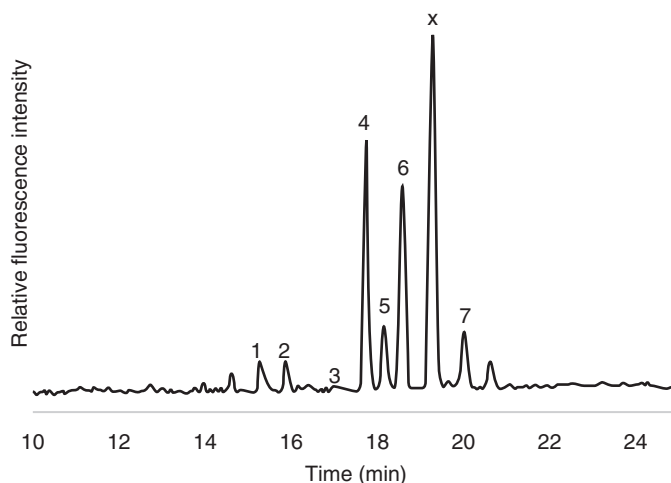
Figure 4 CE-LED separation of a crude *Cheledonium majus* extract; (1) sanguinarine, (2) coptisine, (3) chelerythrine, (4) stylopine, (5) chelidonine, (6) protopine, (7) allocryptopine, and (x) unknown component. (*Source*: Reprinted from J Chromatogr A, 1218, Kulp M, Bragina O, Kogerman P, Kaljurand M, Capillary electrophoresis with LED-induced native fluorescence detection for determination of isoquinoline alkaloids and their cytotoxicity in extracts of *Chelidonium majus* L, 5298–5304, Copyright (2011), with permission from Elsevier.)

practical applicability, without method validation or presenting quantitative results. The major alkaloids in this plant were also assayed under nonaqueous CE conditions using a buffer containing 35 mM ammonium acetate, 0.25% acetic acid, and 5% ACN in methanol (Liu *et al.*, 2006). The required analysis time to determine three major alkaloids (berberine, palmatine, and jatrorrhizine) was 4 min, using a 75 μm ID fused silica capillary with an effective length of only 20 cm and applying 10 kV. Detection was achieved by a fluorescence spectrometer (excitation at 488 nm, emission recorded from 500 to 650 nm). Similar conditions have been described for the CE separation of *Sophora* alkaloids (e.g., matrine and oxymatrine) using UV detection at 205 nm (Chen *et al.*, 2009a) or ESI-MS (Wang, Qu, and Cheng, 2007). In 2012, this group published another method for the analysis of five alkaloids in *Sophora flavescens* Ait. after sample preparation by subcritical water extraction (SWE). The buffer consisted of 110 mM monosodium phosphate at pH

3.0 mixed with isopropanol at the rate of 85 : 15 and FASS was optimized by adding 0.8 mM phosphoric acid to significantly improve reproducibility (Wang *et al.*, 2012).

The impact of sample preconcentration techniques was impressively shown by Sun and colleagues. They used either FASS (Sun and Tseng, 2004) or sweeping (Sun and Tseng, 2005) to determine berberine, coptisine, and palmatine in *C. chinensis* extracts; the observed concentration factors were 240- and 500-fold, respectively. In case of sweeping, the sample was dissolved in 50 mM phosphoric acid, the running electrolyte composed of 100 mM phosphoric acid with 15 mM SDS and 10% THF.

Besides the above-mentioned examples for the analysis of alkaloids by CE and related techniques, a large number of similar applications can be found in literature. As they cannot be discussed in detail at this point, a few selected methods are compiled in Table 1. The provided information includes name of analytes, separation technique and conditions, degree

Table 1 Selected applications of alkaloid analysis by CE and CEC.

Application/analytes	Technique/buffer	Detection	Validated	Application	References
Citrus aurantium (tyramine, synephrine, and octopamine)	CEC (monolith)/3:7 mixture of 10 mM ammoniumacetate (pH 9.5) in ACN/isopropanol (1 : 3)	UV (210 nm)	1–4	S, SA+	(Chizzali, Nischang, and Ganzera, 2011)
colchicine and related alkaloids (colchicoside and thiocolchicine)	MEKC on microchip/25 mM CHES buffer (pH 9.0) with 10 mM SDS	UV (350 nm)	1, 2	S	(Lu, Copper, and Collins, 2006)
Corydalis species (coptisine, palmatine, protopine, and others)	NACE-MS/50 mM ammonium acetate, 1 M acetic acid in ACN/methanol (9 : 1)	ESI +	1–4	S, SA+	(Sturm, Seger, and Stuppner, 2007)
Fritillaria species (verticine and verticinone)	CZE with ionic liquid/40 mM BMImBF$_4$ in 8 mM phosphate buffer (pH 8.0)	CL	1–4	S, SA+	(Gao *et al.*, 2009)
Lupinus species (sparteine, lupanine, and angustifoline)	NECE-MS/100 mM ammonium acetate in MeOH, ACN, and water (7 : 2 : 1), and 1% HOAc	UV (210 nm), ESI +	1–4	S, SA+	(Ganzera, Krüger, and Wink, 2010)
Pyrrolizidine alkaloids in TCM (senkirkine, retrorsine, and others)	MEKC/20 mM borate buffer (pH 9.1), 30 mM SDS, and 20% methanol	UV (220 nm)	1-4	S, SA+	(Yu *et al.*, 2005)
Sinomenium actum (sinomennine and berberine)	NACE/80 mM ammonium acetate in MeOH containing 20% ACN and 2% HOAc	UV (262 nm)	1–4	S, SA+	(Zhu *et al.*, 2010)
Stephania tetrandra (tetrandrine and fangcholioline)	NACE/50 mM ammonium acetate, 0.5% HOAc, in ACN/MeOH (1 : 1)	UV (214 nm)	1–4	S, SA+	(Gao *et al.*, 2005)
Tinospora sagittata (columbamin, palmatin, and deoxyecdysone)	CEC (C-18 packed) in pressure mode/0.04% FA in 5 mM ammonium acetate, ACN, and water	UV (247 nm)	1–4	S, SA+	(Yang *et al.*, 2007b)

Validated, 1 (sensitivity), 2 (specificity), 3 (accuracy) and 4 (precision); Application, S (standard), SA (sample) and + (quantification).

of validation, and type of application, for example, standard mixture or real-life sample.

6.2   Anthraquinones

Anthraquinones are the largest group of naturally occurring quinones. They can be considered as oxidized anthracene derivatives, and they usually are of intense yellow, orange, or red in color. The latter explains their traditional use as dyes for textiles and papers. Concerning medicinal relevance, many herbal drugs containing these compounds (mostly in the form of glycosides) are employed as laxatives, with *Senna* [the differentiation in *Cassia senna* L. and *Cassia angustifolia* Vahl. is not supported by recent data, both are now compiled to *Cassia alexandrina* Thell., Fabaceae (Rahfeld, 2011)], aloe or rhubarb being the most prominent examples. Especially for the analysis of rhubarb, different electrophoretic techniques have been reported. Several *Rheum* species (e.g., *R. palmatum* L., *R. tanguticum* Maxim. ex Balf., and *R. officinale* Baill.) play an important role in traditional Chinese medicine for purging heat, promoting blood circulation, and to treat gastric and renal disorders (Zhang and Liu, 2004); major bioactive compounds in these plants are 1,8-dihydroxyanthraquinones such as aloe-emodin, rhein, physcion, and chrysophanol. For their CE

separation, Li *et al*. utilized a 15-mM aqueous sodium borate solution (pH 9.3) containing 30 mM β-cyclodextrin, 20% ACN, and 1% ethanediol. The latter was found to be advantageous to improve peak shape and to reduce analysis time to 20 min. For extraction of the analytes, the samples (unspecified *Rheum* species and TCM preparations) were acidified with 20% sulfuric acid and exhaustively extracted with chloroform; then the organic solvent was evaporated and the residue dissolved in methanol (Li, Cao, and Ding, 2004a). The same compounds could also be well resolved with a BGE comprising 50 mM borate buffer (pH 8.2), 25% ACN, and 25% isopropanol; yet, migration times were prolonged by the factor of 2 approximately (Gong *et al*., 2005). Koyama and coworkers described a CZE approach enabling the determination of 11 anthraquinones (Figure 5), including the mentioned aglyca earlier, their respective glucosides and sennosides A and B (Koyama, Morita, and Kobayashi, 2007). To a large extent, their buffer resembled the one described by Li, but instead of β-cyclodextrin, it contained a mixture of 5 mM α-cyclodextrin and 2 mM 2,6-di-O-methyl-β-cyclodextrin. The authors also developed an HPLC assay for the same analytes, and in their final statement, they concluded that the CE approach has certain benefits (e.g., better separation of sennosides and rhein-8-β-D-glucoside) and also disadvantages too (LOD for CE: 0.1–0.8 μg mL^{-1}, HPLC: 0.02–0.2 μg mL^{-1}). General
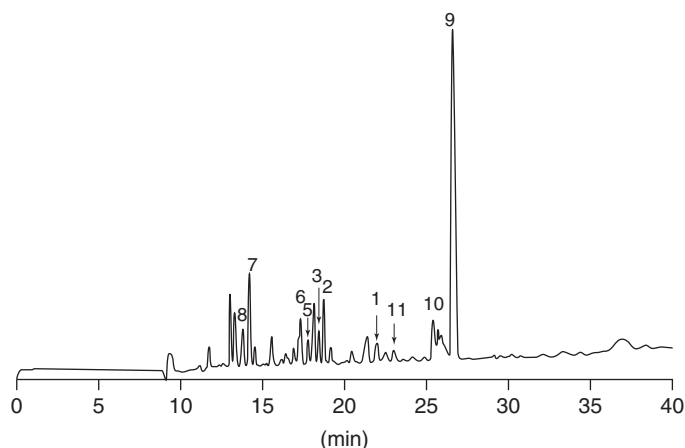
Figure 5   CZE separation of anthraquinones in a rhubarb extract; (1) emodin, (2) chrysophanol, (3) aloe-emodin, (5) emodin-1-β-D-glucoside, (6) emodin-8-β-D-glucoside, (7) chrysophanol-1-β-D-glucoside, (8) chrysophanol-8-β-D-glucoside, (9) rhein-8-β-D-glucoside, (10) sennoside A, and (11) sennoside B. (*Source*: Reprinted from J Chromatogr A, 1145, Koyama J, Morita I, Kobayashi N, Simultaneous determination of anthraquinones in rhubarb by high-performance liquid chromatography and capillary electrophoresis, 183–189, Copyright (2007), with permission from Elsevier.)

preferences for either of them were therefore not possible. Even they were described, the addition of SDS (Shang and Yuan, 2003) or the use of an ionic liquid as buffer (Tian *et al.*, 2007) did not significantly enhance separation efficiencies for these compounds. An innovative option for the analysis of rhubarb anthraquinones might be CEC (Lu *et al.*, 2007), as an in-house prepared methacrylate monolithic column (with iso-butyl methacrylate, ethylene dimethacrylate, and methacrylic acid as monomers) enabled the baseline separation of the main aglyca in <6 min. This was possible on a 20-cm long capillary in p-CEC mode, that is, during analysis voltage (−20 kV) and pressure (100 psi at inlet) were applied simultaneously. A mixture of 10 mM phosphate buffer (pH 6.2) in ACN (35/65) showed to be the optimum mobile phase for this application.

Several reports describe the CE analysis of the same 1,8-dihydroxyanthraquinones in other medicinal plants, for example in *Cassia obtusifolia* L. (Jiang, Lv, and Wang, 2005) or *Polygonum cuspidatum* Siebold & Zucc. (Tian *et al.*, 2006). The described procedures resemble the already mentioned ones to a large extent, for example, MEKC with sodium cholate as tenside or cyclodextrins were added to the BGE. A different technique was utilized by Li, who assessed less polar methoxyanthraquinones in *Xanthophytum attopvensis* Pierre by NACE (Li *et al.*, 2005). In this case, the running electrolyte consisted of 50 mM sodium cholate and 1% acetic acid in ACN and methanol (4/6, v/v). Cholate served two purposes, as electrolyte and surfactant; it is also well soluble in methanol. The addition of acid showed a pronounced and positive effect on the separation of compounds bearing ionizable groups such as an active aldehyde (e.g., 1-methoxy-2-formyl-3-hydroxy-9,10-anthraquinone). Applied voltage, temperature, and detection wavelength were kept constant at 25 kV, 21°C, and 254 nm, respectively, during analysis. For the three analytes of interest, method validation confirmed linear ranges from 1 to 250 µg mL^{-1}, excellent correlation coefficients of at least 0.9999, and detection limits below 0.4 mg L^{-1}. As the quantitative analysis of one plant extract was feasible with excellent repeatability too, this procedure can be considered an excellent alternative to established procedures such as HPLC.

6.3 Flavonoids

Flavonoids are polyphenolic compounds commonly found in plants. At least 6500 flavonoids have been identified so far and they can be divided into seven subgroups, namely flavones, flavanols, flavanones, isoflavones, catechins, anthocyanidins, and chalcones. They are well known for their health-promoting qualities and their pharmacological potential (Bachmann *et al.*, 2007). Owing to their reported antioxidant properties, flavonoids are protective substances against coronary heart diseases, cancer, and stroke (Molnarperl and Fuzfai, 2005). In addition, they can downregulate inflammatory processes and they show antimutagenic properties. Although HPLC is the most popular method for their qualitative and quantitative analysis, CE is known to be an excellent alternative, as it is very fast, simple in usage, highly efficient, and does not generate a lot of costs (Bachmann *et al.*, 2007).

The majority of CE assays for flavonoid analysis utilize CZE (Table 2). The buffers described are usually based on borate or sodium tetraborate (borax) with varying concentrations from 10 to 100 mM, whereby it was not always made clear in literature if the term "borate" equals "borax" or not. A typical example is the usage of a 50-mM borax buffer with pH 8.2 for the determination of hyperosid, quercetin, vitexin, and other flavonoids in *Crategus monogyna* Jaqu. Separation is performed at 25°C and an applied voltage of +25 kV, with UV detection set to 280 nm (Urbonaviciute *et al.*, 2006). Similar conditions were chosen for the investigation of flavonoids in *Garcinia kola* Heckel (Okunji *et al.*, 2002), *Fagopyrum esculentum* Moench (Hinneburg, Mrestani, and Neubert, 2004), or *Solanum lycopersicum* L. (Helmja *et al.*, 2008). An enhancement of sensitivity can be achieved via sample stacking techniques such as LVSS (low voltage sample stacking). For this purpose, a low voltage of −5 kV was applied for 83 s after injection of the sample, which consisted of flavonoids (quercetin and kaempferol) from *Brassica oleracea* L. This procedure was followed by a CZE separation with a simple 10 mM borax buffer at pH 8.4 (Lee, Boyce, and Breadmore, 2012).

In many applications, a certain amount, usually between 10% and 30%, of organic solvents (e.g., methanol and ACN) was added to the BGE, for example, for the separation of apigenin, luteolin, their 7-O-glycosides, and rutin in *Achillea millefolium* L. In this case, the buffer consisted of 30 mM borate

Table 2 Selected applications of flavonoid analysis by CE and CEC.

Applications/analytes	Buffer	Detection	Validated	Application	References
Fagopyrum esculentum (rutin, hyperoside, and chlorogenic acid)	CZE/60 mM borate; pH 10	UV (206 nm)	1,2,4	SA+	(Hinneburg, Mrestani, and Neubert, 2004)
Passiflora incarnata (vitrexin, swertisin, and orientin)	CZE/25 mM borax, 20% MeOH; pH 9.5	UV (275 nm)	1,2,4	SA+	(Marchart, Krenn, and Kopp, 2003)
Verbena officinalis (luteolin, apigenin; phenylethanoides, and iridoid-glycosides)	MEKC/50 mM borax, 50 mM SDS; pH 9.3	UV (205, 235 nm)	1–4	SA+	(Müller, Ganzera, and Stuppner, 2004)
Agrimonia pilosa (catechin, hyproside, quercitrin, quercetin, and rutin)	CZE/60 mM borax, 120 mM NaH_2PO_4; pH 8.8	ECD	1–4	SA+	(Xu *et al.*, 2005)
Flavonoids and carboxylic acids (e.g., apigenin, luteolin, and ferulic acid)	CEC (C18 packed)/20 mM Tris–HCl, in ACN and water (40 : 60); pH 6.5	UV (205, 210, and 254 nm)	1	S	(Stoggl *et al.*, 2006)
Olea europaea (flavonoids, lignans, secoiridoids, phenyl alcohols, and phenyl acids)	CE-MS/25 mM ammonium hydrogen carbonate, pH 9; isopropanol/water = 1 : 1 (sheat liquid)	ESI –	1–4	SA	(Carrasco-Pancorbo *et al.*, 2007)
Saussurea katochaete (e.g., apigenin and 7-hydroxy-coumarin)	CZE/20 mM sodium phosphate, 10% MeOH; pH 11	UV (214 nm)	1–4	SA+	(Yue, Li, and Shi, 2007)
Prunella vulgaris (rutin, quercetin, ursolic acid, and rosmarinic acid)	CZE/40 mM borax, 2 mM β-CD, 4% MeOH; pH 9.4	UV (210 nm)	1–4	SA+	(Cheung and Zhang, 2008)
Chinese herbal teas (apigenin, kaempferol, luteolin, rutin, quercetin, and ferulic acid)	CZE/400 mM potassium hydrogen carbonate, 200 mM borax, 0.2 mM β-CD; pH 7.6	ECD	1–4	SA	(Chi *et al.*, 2009)

Validated, 1 (sensitivity), 2 (specificity), 3 (accuracy) and 4 (precision); Application, S (standard), SA (sample) and + (quantification).

solution and 30% methanol; the sample was introduced by hydrodynamic injection. To enhance the detection sensitivity, a capillary with extended light path (bubble cell) was used. Other examples for the use of borax/borate buffers with organic solvents as modifiers can be found for *Lamiphlomis rotata* Kudo (Luo *et al.*, 2007), *Scorzonera austriaca* Willd. (Jiang *et al.*, 2007), *Halenia elliptica* D. Don (Yang, Yue, and Shi, 2006), *Myricaria bracteata* Royle, *Myricaria wardii* C. Marquand (Zhao *et al.*, 2005a), *Saussurea mongolica* Franch. (Jiang, Li, and Shi, 2004b), *Passiflora incarnata* L. (Marchart, Krenn, and Kopp, 2003), *Crataegus pinnatifida* Bunge (Chen and Liu, 2005), *Nelumbo nucifera* Gaertn. (Do *et al.*, 2012), and different *Epidemium* species (Chen *et al.*, 2009b).

Cyclodextrins are popular additives in CE, because they are able to enhance peak shape and separation efficiency (especially for chiral compounds). For the determination of two flavonoids and six other compounds in *Prunella vulgaris* L., a 40-mM borax

buffer containing 4% methanol and 2 mM β-CD enabled good results (Cheung and Zhang, 2008). Comparable procedures can be found for the analysis of *Hippophae rhamnoides* L. (Yue, Jiang, and Shi, 2004) and *Ginkgo biloba* L. (Shaban, Gajdošová, and Havel, 2006).

Apart from the very popular borax and borate buffers, other electrolytes were applied for flavonoid analysis too. The determination of 26 components in the aerial parts of *Genista tenera* (Jaqu.) Kuntze, including several flavonoids, was possible with a 10-mM ammonium carbonate solution in a water/propan-2-ol mixture (95 : 5, v/v; pH 9.25 adjusted with ammonium hydroxide). This volatile buffer was required because of the use of CE-MS (Edwards *et al.*, 2006). Another buffer system that, even not volatile, has also been used for CE-MS was a 200-mM boric acid solution at pH 9.5. It allowed the determination of different flavonoids in bitter (naringin and hesperidin) and

sweet (narirutin and neohesperidin) variations of *Citrus x aurantium* L. by CZE-ESI-IT-MS (Sawalha *et al.*, 2009). For the CZE-MS analysis of phenolic compounds in extra-virgin olive oil (*Olea europaea* L.), a 25-mM ammonium bicarbonate buffer at pH 9 has been described (Carrasco-Pancorbo *et al.*, 2007), whereas the same scientists later also published a NACE-MS method using electrospray ionization and time-of-flight mass spectroscopy (ESI-TOF-MS) for assessment of the same matrix. This water-free method was developed to simplify the injection of the hydrophilic sample matrix, whereby the buffer consisted of 25 mM ammonium acetate/acetic acid in methanol/ACN (1 : 1, v/v) at pH 5 (Gomez-Caravaca *et al.*, 2009). Last but not least the same authors also described a "regular" CZE procedure for the analysis of flavonoids and other compounds such as phenols, lignans, or phenolic acids in olive oil. Here, the optimum buffer consisted of an aqueous 45 mM borax solution at pH 9.3, enabling the separation of 26 compounds in Oleum olivae virginum within 10 min (Carrasco-Pancorbo *et al.*, 2006). In a later study, this method was used to compare the quality of extra-virgin olive oils from several Spanish and Italian Protected Designations of Origin's (PDO's) (Carrasco-Pancorbo *et al.*, 2009).

Another strategy for the CE analysis of flavonoids is the use of Tris buffers. In this respect, a 20-mM Tris electrolyte containing 3 mM HCl has been described for the fast analysis of baicalin, farrerol, and puerarin in *Scutellaria baicalensis* Georgi. The required separation time was only 1.9 min using a fused silica capillary with 9 cm length (8.8 cm effective length; 50 µm id) and applying a voltage of 4 kV (Tian and Qin, 2009). For the baseline separation of a standard mixture of (−)-epicatechin, hesperetin, myricetin, naringenin, and quercetin by CEC, Stöggl *et al.* used a solution of 10% aqueous 20 mM Tris–HCl, 40% ACN, and 50% water at pH 6.5. The acidic pH was required to assure protonation of the flavonoids, yet the pH should be as high as possible in order to produce sufficient EOF. The capillaries used were of 33.5 cm length (25 cm effective length) with an inner diameter of 100 µm and filled with 3 µm reversed phase material (Hypersil ODS, 120 Å). During the measurements, inlet and outlet of the capillary were set under pressure to prevent the formation of air bubbles (Stoggl *et al.*, 2006).

Sodium phosphate buffers are commonly employed for flavonoid analysis by CE as well; for example, Yue, Li, and Shi (2007) developed a CZE method to

assay six active flavonoids in *Saussurea katochaete* Maxim. using a 20-mM sodium phosphate buffer containing 10% MeOH (Yue, Li, and Shi, 2007), whereas Zhou *et al.* (2007) was able to separate three different plant constituents, one of them being the flavonoid rutin, in samples of *Acanthopanax senticosus* Harms. In this case, a phosphate/borate buffer was advantageous (7.5 mM sodium phosphate and 7.5 mM borax), and EC detection was used. In general, this technique is widely used in combination with CE for detection of flavonoids because of their oxidizable nature. Just to name a few out of many examples, Xu determined catechin, hyperosid, quercetin, quercitrin, and rutin in *Agrimonia pilosa* Ledeb. (Xu *et al.*, 2005), in *Houttuynia cordata* Thunb. and *Saururus chinensis* (Lour.) Bail. (Xu *et al.*, 2006), Cao and coworkers the structurally closely related derivatives liquiritigenin and isoliquiritigenin in *Glycyrrhiza uralensis* Fisch. ex DC. (Cao *et al.*, 2004), or Peng several flavonoids and phenolic acids (apigenin, luteolin, catechin, ferulic acid, rosmarinic acid, and caffeic acid) in *Perilla frutescens* L. (Peng, Ye, and Kong, 2005).

Multiple MEKC methods for the investigation of flavonoids can be found in literature. For the separation of apigenin, diplacone, and mimulone in *Paulownia tomentosa* (Thunb.) Steud., a 20-mM borax buffer in 5% methanol, containing 10 mM sodium dodecyl sulfate (SDS) as anionic surfactant, was used (Jiang, Du, and Shi, 2004a). Five compounds, including apigenin-7-O-β-D-diglucuronide and luteolin-7-O-β-D-diglucuronide, in *Verbena officinalis* L. have been investigated using a BGE consisting of 50 mM borax and 50 mM SDS (Müller, Ganzera, and Stuppner, 2004). The separation of 12 compounds, mainly flavonoids (e.g., luteolin and quercetin-3-O-β-D-glucoside) but also caffeic acids and leontopodic acids, in *Leontopodium alpinium* Cass. was possible with a 60 mM borax buffer containing 25 mM SDS and 35% ACN at pH 6.75 (Ganzera *et al.*, 2012). To separate three flavonoids in *S. baicalensis* Georgi, 10 mM sodium cholate (SC) was added to a solution of 20 mM borax and 20 mM SDS, in order to improve the separation without increasing the migration time (Dong *et al.*, 2009). Zhu *et al.* (2007) compared two different preconcentration techniques, which were stacking with reverse migrating micelles (SRMMs) and anion-selective electrokinetic injection with water plug-sweeping and reverse migrating micelles (ASIW-sweep-RMMs) before separating

six flavonoids (tangeretin, nobiletin, hesperetin, naringenin, hesperidin, and naringin) with MEKC. In respect to preconcentration, both approaches worked well, yet with ASIW-sweep-RMM, slightly better efficiencies (a 25–66-fold increase of peak area) were observed. The CE buffer consisted of 20 mM aqueous phosphoric acid, 100 mM SDS, and 20% ACN at pH 2 (Zhu *et al*., 2007). A similar procedure was successfully applied for flavonoid analysis in *Swertia mussotii* Franch. and *Swertia franchetiana* Harry Sm. (Li *et al*., 2008b). Fan *et al*. (2010) additionally added 50 mM γ-CD to a borax buffer with 50 mM SDS to be able to separate six structurally similar flavonoid gallate esters occurring in *Nepenthes gracilis* Korth. (Fan *et al*., 2010).

An ITP-CZE method was implemented to determine quercetin, quercitrin, chlorogenic acid, isoquercitrin, hyperoside, and rutin in *Hypericum perforatum* L. This required optimization of leading electrolyte (LE) and a terminating electrolyte (TE), as well as the running buffer for CZE. The optimum LE for this separation problem showed to be a 10 mM HCl solution with Tris as counterion and 0.2% 2-hydroxyethylcellulose at pH 7.2. The TE was 50 mM boric acid at pH 8.2, with the pH being adjusted with barium hydroxide. For the CZE separation, a 50 mM Tris-buffer containing 25 mM β-hydroxy-4-morpholino-propanesulfonic acid (MOPSO) and 65 mM boric acid was prepared. All six flavonoids could be baseline separated in 30 min with this method (Hamoudova, Pospisilova, and Spilkova, 2006). Another innovative option is the coupling two different CE methods for flavonoid analysis. Zhang *et al*. created a heart-cutting two-dimensional (2D) CE method combining MEKC in the first and CZE in the second dimension to effectively separate kaempferol, hesperetin, apigenin, rutin, hyperoside, and quercetin in *Leonurus cardiaca* L. In between the two separation steps, sweeping by means of analyte focusing by micelle collapse (AFMC) was implemented. The latter was essential to destroy the micelles required for MEKC, narrow the analyte zone, and to decrease the peak width for an optimized separation (Zhang and Zhang, 2011).

6.4 Terpenes

Terpenes, also known as *terpenoids* or *isoprenoids*, consist of chains of C_5-isoprene units that are linked to each other via their head or tail ends. Depending on the amount of isoprene units connected, the chains form mono- (10 carbons), sesqui- (15 carbons), di- (20 carbons), or triterpenes (30 carbons). The formation of longer chains is also possible but does not occur frequently. General statements about the chemical and physical properties of terpenes are not possible because of their structural variability, so that this aspect will be briefly discussed in the chapters to follow (Hänsel and Sticher, 2007a). An overview to selected CE procedures for terpene analysis is shown in Table 3.

6.4.1 Mono- and Sesquiterpenes

Monoterpenes may be present in an acyclic, monocyclic, or bicyclic constitution. They are commonly found in the aerial parts of most plants in the form of essential oils. As those oils are responsible for protection against herbivores, they usually show antibacterial, antihelmintic, and antifungal activities (Gotti, 2011).

A CZE method for the separation of the monoterpenes α-pinene, β-pinene, camphene, and limonene as well as tetralin and terbutalin was published by Gahm, using a 10-mM sodium phosphate buffer at pH 3.3 with an addition of 6.5 mM sulfated β-CD. The separation was performed in reversed mode, that is, the applied voltage at the cathode was negative (−20kV). Special attention was paid to the separation of chiral analytes, and the addition of different types and amounts of CDs studied in detail (Gahm, Chang, and Armstrong, 1997). In another study that is also dealing with the difficulties of separating chiral monoterpenes, Diagone *et al*. tried to compare two different CE-methods, CZE and MEKC, to capillary liquid chromatography (C-LC). The investigated sample was a standard mixture of monoterpenes typically found in *Citrus* essential oils. Again, it was confirmed that in any case cyclodextrins were necessary for the separation of these chiral analytes. In CZE, an acidic 20 mM phosphate buffer with the addition of 50 mM SDS was used, and the impact of (2,3,6-tri-O-methyl)-β-cyclodextrin (met-β-CD) and α-CD studied in detail. The cyclodextrins were added individually or in combination at different concentrations (0 to 25 mM). It was observed that using met-β-CD resulted in decreased migration times, but neither of the two CDs permitted the separation of racemic carvone. The optimum MEKC electrolyte

Table 3 Selected applications of terpene analysis by CE and CEC.

Applications/analytes	Buffer	Detection	Validated	Application	References
Monoterpenes					
Citrus essential oil (e.g., citral, linalool, pinene, and carvone)	MEKC/20 mM phosphate, 50 mM SDS, 20 mM met-β-CD; pH 8.1	UV (214, 254 nm)	2	S	(Diagone, Ogawa, and Lancas, 2003)
Piper nigrum (e.g., caryophyllene, piperine, and limonene)	CEC (C-18 packed)/50 mM ammonium acetate/ acetonitrile (1 : 9); pH 6.0	UV (210, 265, and 338 nm)	1–4	SA+	(Musenga et al., 2007)
Diterpenes					
Ginkgo biloba (diterpene trilactones, sesquiterpenes, and flavonoids)	MEKC/20 mM phosphoric acid, 40 mM SDS, 12 mM β-CD; pH 2.2	UV (190, 250 nm)	1–4	SA	(Dubber and Kanfer, 2006)
Salvia miltiorrhiza (abietane-type diterpenes, e.g., tanshinone I)	MECK/140 mM SC in methanol	UV (275 nm)	1–4	SA+	(Chen et al., 2005)
Iridoids					
Picrorhiza kuroa (picrosid, veronicosid, and apocynin)	MEKC/100 mM borate, 30 mM SDS, 1% acetonitrile; pH 8.6	UV (205 nm)	1–4	SA+	(Sturm and Stuppner, 2001)
Triterpenes					
Glycyrrhiza glabra (saponins, e.g., glycyrrhizin)	CZE/50 mM borax, 5 mM β-CD; pH 9.0	UV (254 nm)	1	SA+	(Sung and Li, 2004)
Boswellia serrata (boswellic acids)	CEC (C-18 packed)/20 mM ammonium formate/ acetonitrile (1:9); pH 6.5	UV (210, 254 nm)	1–4	SA+	(Ganzera et al., 2003)

Validated, 1 (sensitivity), 2 (specificity), 3 (accuracy) and 4 (precision); Application, S (standard), SA (sample) and + (quantification).

consisted of a 20-mM phosphate buffer (pH 8.1) with the addition of 50 mM SDS and 20 mM met-β-CD. This setup showed better results; nevertheless, a baseline separation of racemic carvone was still not possible. Compared to a parallel-developed C-LC method, the authors concluded that MEKC is equally suitable for this separation problem (Diagone, Ogawa, and Lancas, 2003).

A CEC method for the separation of 11 terpenoid compounds in *Piper nigrum* L., such as limonene, piperine, and 3-carene, using a 32 cm long capillary packed with C18 material (LiChrospher 100 RP18, 5 µm) was reported more recently. The buffer consisted of 50 mM ammonium acetate solution and ACN (ratio 1 : 9, v/v) at pH 6 (Musenga *et al.*, 2007). The required analysis time was below 25 min, and method validation indicated adequate sensitivity (LOD values ranging from 1.0 to 10 µg mL^{-1}) and repeatability (RSD below 3%).

The number of reports describing the analysis of sesquiterpenes by CE is much smaller. One of the few examples is a CZE procedure for the separation of eremophilenolides in *Ligulariopsis shichuana* Y.L. Chen. These compounds are typical for the genus *Ligularia* and *Cacalia*, so that their occurrence in the

just mentioned species was considered as proof for a close chemosystematic relationship. Four respective compounds could be well resolved in a plant extract in <10 min using a simple 20 mM borate buffer at pH 10, without the need of additional modifiers; detection was performed at 214 nm (Jiang *et al.*, 2006). Another manuscript reports on the separation of the (+)- and (−)-enantiomers of gossypol in *Gossypium*. As (−)-gossypol is more toxic for animals, it is important to determine the ratio of both compounds in order to cultivate cotton with a low toxicity. This was possible by HPCE using a 50 mM borate buffer at pH 9.3, with a separation voltage of 15 kV and a cassette temperature of 15°C. Both enantiomers could be resolved within 6 min on a relatively short capillary of only 24.5 cm effective length and an inner diameter of 50 µm (Vshivkov *et al.*, 2012).

6.4.2 Iridoids

Biogenetically, iridoids are monoterpene derivatives. They comprise two isoprene units forming a unique cyclopentane ring fused to a six-membered

oxygen heterocycle. Iridoids can further be subdivided into three groups, the iridoid glycosides, the secoiridoidglycosides, and the nonglycosidic iridoids. The CE assays for iridoid analysis do not differ much from already described approaches. To give two typical examples only, four iridoid glycosides (6-O-methylcatalpol, aucubin, harpagid, and harpagosid) occurring in *Scrophularia ningpoensis* Hemsl. were separated by CZE using a 100 mM borax buffer containing 20% methanol at pH 9.3 (Li *et al.*, 2008a). The iridoid glycosides occurring in *Picrorhiza kuroa* Royle ex Benth, including picrosid, apocynin, and veronicosid, have successfully been analyzed by MEKC. A total of 13 different derivatives could be baseline separated using a 100 mM borate buffer containing 30 mM SDS and 1% ACN at a pH of 8.6. This method was compared to HPLC-MS and it was found that both methods deliver comparable qualitative and quantitative results, but the MEKC method was advantageous in respect to required analysis time (15 vs 40 min) (Sturm and Stuppner, 2001).

6.4.3 Diterpenes

More than 3000 diterpenes are known so far, nearly all of them are cyclic. One of the most common basic structures is labdane, which can be further metabolized by ring modifications and additional functional groups. Accordingly, lipophilic and hydrophilic diterpenes are differentiated; yet, the polar derivatives possessing hydroxyl, epoxy, or carbonyl groups are usually the chemical and biological more active ones (Hänsel and Sticher, 2007a).

For the determination of the three isomeric diterpenes acanthoic acid, continentalic acid, and kaurenoic acid in *Acanthopanax* L., a CZE method using a 50 mM borax buffer containing 30 mg mL^{-1} of β-CD sulfobutyl ether at pH 8.5 has been reported (Phuong *et al.*, 2006). Roots and stem bark of these plants are used as sedative and tonic in traditional Asian medicine, as well as the treatment of diabetes and rheumatism. A CE-MS procedure for the determination of seven phenolic diterpenes and several phenolic acids in *Rosmarinus officinalis* L. can also be found in literature, but the separation conditions resemble previously described ones (e.g., mixture of 40 mM ammonium acetate and ammonium hydroxide at pH 9.0 as buffer) to a large extent (Herrero *et al.*, 2005a).

Another popular option for diterpene analysis is MEKC, and it was used for the separation of horminone and 7-O-methylhorminone in *Salvia chionantha* Boiss. and *Salvia kronenburgii* Rech.f. for example. The optimized background electrolyte for this application was found to be 50 mM SDS in 25% methanol at pH 11.5 (Oztekin *et al.*, 2010). Another sage species was investigated via nonaqueous-micellar electrokinetic chromatography (NAMEKC). The abietane-type diterpenes cryptotanshinone, tanshinone IIA, and tanshinone I in *Salvia miltorrhiza* Bunge were separated using a nonaqueous buffer consisting of 140 mM sodium cholate in methanol. This bile salt served as biological surfactant (CMC in methanol 111 mM) as well as electrolyte. Two compounds, tanshinones IIA and I, could not be baseline separated; therefore, a second-order derivative electropherogram had to be constructed to perform quantifications (Chen *et al.*, 2005). A reversed-flow micellar electrokinetic chromatography (RF-MEKC) method was developed for the determination of terpene trilactones (ginkgolides A, B, and C) together with the sesquiterpene bilabolid and several flavonoids in *G. biloba* L. The buffer consisted of 20 mM phosphoric acid with 40 mM SDS and 12 mM β-CD at pH 2.2. During analysis (run time 22 min), the applied voltage was changed after 11 min from −17.5 to −20 kV, an adjustment that helped to reduce the required separation time (Dubber and Kanfer, 2006). For the analysis of andrographolide and dehydroandrographolide in *Andrographis paniculata* Nees, two MEEKC methods have been reported. The first one using an ethyl acetate oil-based BGE (0.5% ethyl acetate, 0.6% SDS, and 6.0% 1-butanol in 30 mM borax buffer at pH 9.5; applied voltage −15 kV) (Zhao *et al.*, 2005b), the second a BGE consisting of 0.81% heptane, 3.31% SDS, and 6.61% butan-1-ol in 10 mM borax buffer at + 20 kV (Yanfang *et al.*, 2006). Both were found equally suitable for sample analysis; yet, the first procedure required an analysis time of <3 min only.

6.4.4 Triterpenes and Saponins

Triterpenes represent the largest group within terpenes, and they can be divided into three classes, lipophilic triterpenes, such as phytosterols or triterpene alcohols, highly oxidized triterpenes

with medium polarity and hydrophilic, glycosidic triterpens such as saponins. Saikosaponins, which are found in different *Bupleurum* species, are typical examples for the latter. A simple CZE procedure to separate three of them (saikosaponins a, c, and d) was developed by Lin, using a 40 mM borate buffer at pH 10 (Lin *et al*., 2005). For the investigation of seven bioactive triterpene acids (including maslinic acid, arjunic acid, and euscaphic acid) in *Rubus chingii* Hu, a CZE method using a 200 mM borax buffer containing 15 mM β-CD and 12.5% methanol was best suited (Guo, Gao, and Yang, 2005b). Similar electrolyte systems were reported for the separation of oleanolic acid and ursolic acid in *Pterocephalus hookeri* (C.B. Clarke) Airy Shaw & M.L. Green (Yang *et al*., 2007a), or a selection of different acids, some of them triterpenic acids (e.g., betulinic acid) in *P. vulgaris* L. (Cheung and Zhang, 2008). A nonaqueous system comprising 65% methanol, 35% ACN, and 90 mM Tris has been reported for the determination of oleanolic and ursolic acids, together with 2α,3β,24-trihydroxy-urs-12-en-28-oic acid, in different Chinese herbs (Qi *et al*., 2006). The investigated species included relevant TCM plants such as *Piper kadsura* (Choisy) Ohwi, *Ligustrum lucidum* Hort. ex K.Koch, and *V. officinalis* L.; in most species, ursolic acid (content 3–13 mg g^{-1} plant material) was dominant.

One of the most popular herbal remedies is licorice, *Glycyrrhiza glabra* L., a plant containing saponins such as glycyrrhizin and glycyrrhetic acids as major bioactive ingredients. Not surprisingly, these compounds have also been investigated by CE several times. The applied buffers range from alkaline borax solutions containing cyclodextrins (5 mM β-CD) (Sung and Li, 2004), a carbonate buffer containing methanol, ethylene glycol and β-CD (Sabbioni *et al*., 2005), or simple borax solutions (70 mM borax) without any additives. The latter permitted the quantitative determination of glycyrrhizin, glycyrrhetic acid, glabridin, liquiritin, and licochalcone A in two *Glycyrrhiza* L. species (*G. glabra* and *G. uralensis* Fisch) and allowed the differentiation of Asian and European licorice (Rauchensteiner *et al*., 2005). Another well-known medicinal plant, also containing characteristic saponins (ginsenosides), is ginseng (*Panax ginseng* C.A. Mey., Araliaceae). These compounds are mainly based on tetracyclic aglyca of the dammarane-type, belonging either to the protopanaxadiol or protopanaxatriol subtype.

Special attention has to be paid on the exact authentication of the plant material because the term *ginseng* is also used for other *Panax* species [e.g., *P. quinquefolius* L., *P. notoginseng* (Burkill) F.H.Chen ex C.Y.Wu & K.M.Feng] or even unrelated ones (e.g., *Eleutherococcus senticosus* Maxim., the so-called Siberian Ginseng). For the separation of 10 ginsenosides (including Rd, Rc, Rb1, Rg1, and Re) in *P. notoginseng*, MEKC was successfully utilized (Figure 6). The optimum buffer contained 10 mM phosphoric acid, 140 mM SDS, 20% ACN, and 15% 2-propanol. A methanolic extract of the root material was prepared, and to enhance sensitivity of the method, an on-line concentration of the saponins was achieved using field-enhanced sample injection with reverse migrating micelles (FESI-RMMs). The then observed LOD values ranged from 2 to 6 μg mL^{-1} (Wang, Ye, and Cheng, 2006).

Among the few applications of CEC for triterpene analysis is the analysis of several boswellic acids in *Boswellia serrata* Roxb. The authors prepared packed capillaries with different reversed phases (Hypersil MOS 3 μm, Hypersil ODS 3 μm, and Prontosil C30 3 μm) and evaluated their performance on the separation of a standard mixture of six boswellic acids and a plant extract. Best results were obtained with the C-18 material and a mobile phase comprising 20 mM ammonium formate and ACN in the ratio 1 : 9; samples were injected by simultaneously applying voltage (20 kV) and pressure (10 bar at inlet) for 12 s. The quantitative results obtained were in good agreement to respective HPLC data (Ganzera *et al*., 2003).

6.5 Proteins and Lectins

Proteins are biological macromolecules formed by amino acids, lectins on the other hand are proteins linked to a glycan residue (Hänsel and Sticher, 2007b). Both are found in plants and especially lectins usually show pronounced pharmacological effects; therefore, they represent an interesting type of natural products. Because of their natural occurrence in small quantities only, CE is not a very common analytical approach for the analysis of proteins. Regardless, it is sometimes used (often coupled to MS), especially for smaller lectins or the clarification of amino acid compositions.

One example for the analysis of lectins by CE is the investigation of *Urtica dioica* L. The lectins in
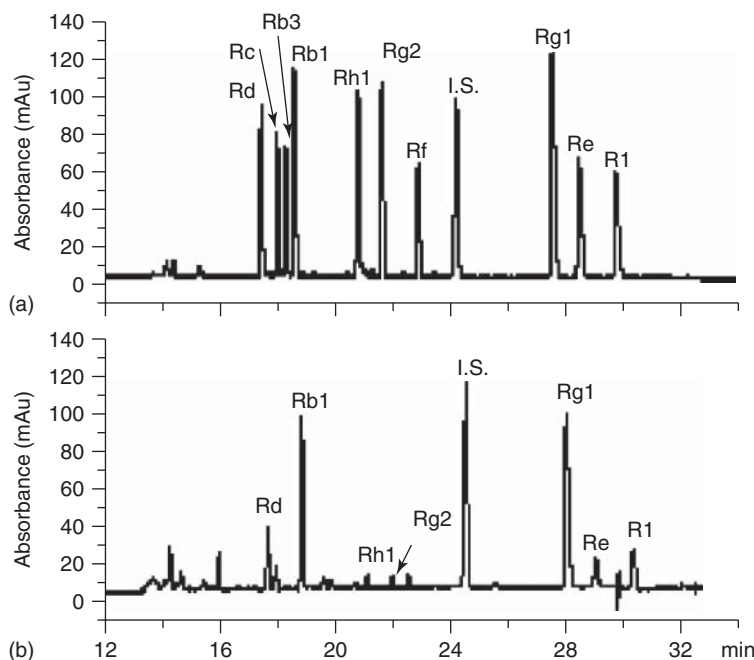
Figure 6   Analysis of saponins (ginsenosides Rd, Rc, Rb3, Rb1, Rh1, Rg2, Rf, Rg1, Re, and R1) in *Panax notoginseng* by MEKC; (a) standard mixture and (b) *P. notoginseng* extract. (*Source*: Reprinted from J Chromatogr A, 1109, Wang S, Ye S, Cheng Y, Separation and on-line concentration of saponins from *Panax notoginseng* by micellar electrokinetic chromatography, 279–284, Copyright (2006), with permission from Elsevier.)

stinging nettle root are called *UDA* (*Urtica dioica agglutinin*), a mixture of several isolectins composed of 80–90 amino acids. As UDA is strongly stimulating lymphocyte proliferation, it might be an explanation for the health benefits of this plant in the treatment of benign prostate hyperplasia (BPH). In a study by Ganzera *et al.*, three isolectins (ILs I, II, and VI) were isolated from plant material by chromatographic means and then used for developing a CZE assay. Most suitable for separation was a 200 mM sodium acetate buffer at pH 3.75, adjusted with acetic acid (Figure 7). Applied voltage, temperature, and detection wavelength were 25 kV, 30°C and 228 nm, respectively. By changing to a 100 mM ammonium formate buffer with the same pH, CE-MS experiments were possible. They confirmed identity of isolated lectins and allowed the assignment of additional ones. Quantitative results of different nettle root samples indicated varying lectin profiles, ranging from 0% to 0.42% (Ganzera *et al.*, 2005). Another publication presented the CZE separation of six leguminous lectins isolated from different species of the *Erythrina* genus, for example, *E. caffra* Thunb. or

E. belliformis Kearn. Once again, separation was achieved in the acidic pH range and CE-ESI-MS utilized for detection (Bonneil *et al.*, 2004).

The analysis of proteins is done predominantly with CZE, coupled to different detection methods such as MS, UV, or FLD. In order to monitor the extraction of phycobiliproteins from the microalga *Spirulina platensis*, a CZE-ESI-MS assay was reported. The biological material was extracted by pressurized liquid extraction (PLE) and the resulting extracts separated using a BGE containing 40 mM ammonium hydrogen carbonate in 45% water, 50% ACN, and 5% 2-propanol; pH of this solution was 7.8 (Herrero *et al.*, 2005b). Another study investigated the composition of 58 peptides in the seeds of *Brassica napus* L. by CZE-ESI-MS. Commercial rapeseed meal was concentrated for proteins and then hydrolyzed with Alcalase® at pH 9 and subjected to gelchromatography on Sephadex material to obtain a solution rich in small proteins (≤1000 Da). Afterward this solution was analyzed by HPLC-MS and CE-MS, the latter with a buffer only containing 50 mM formic acid in water (pH of 2.75).
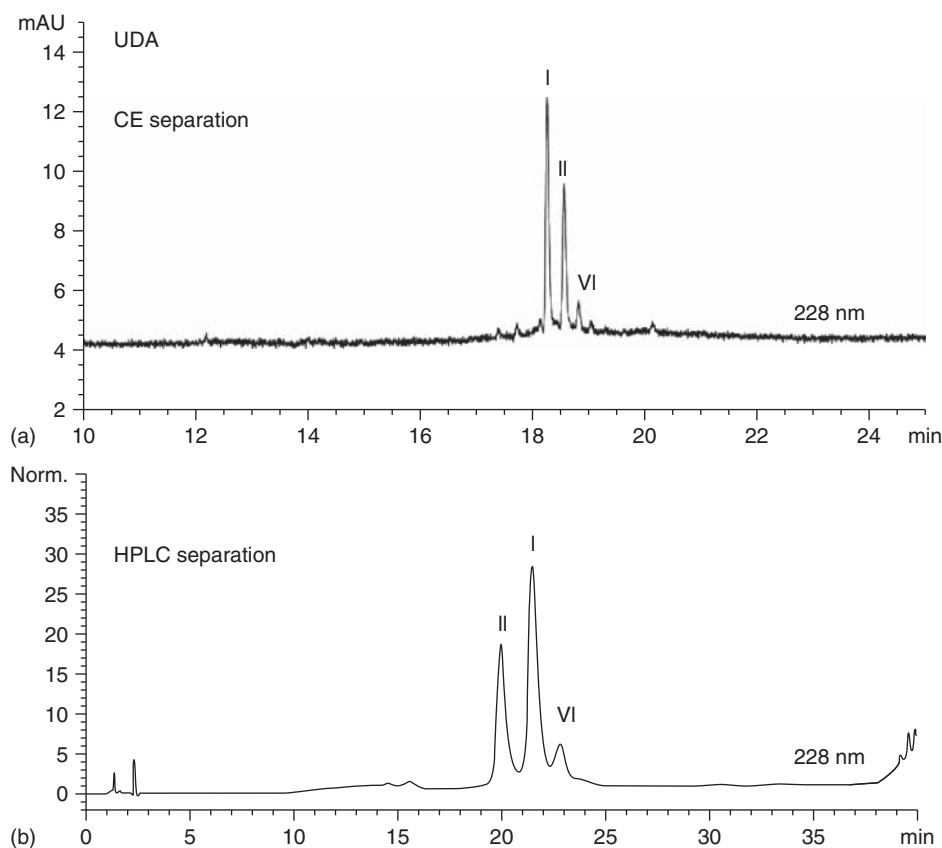
Figure 7 Separation of *Urtica dioica* agglutinin (UDA) by CZE (a) and HPLC (b); (I) isolectin I, (II) isolectin II, and (VI) isolectin VI. (*Source*: Reproduced with permission from Ganzera M, Piereder D, Sturm S, Erdelmeier C, Stuppner H, Electrophoresis. Copyright © 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.)

The sheath liquid, introduced at a flow rate of 6 μL min^{-1} by a syringe pump, was methanol containing 0.2% formic acid (Tessier *et al.*, 2005). The analysis of SDS complexes with subunits of the protein ribulose-1,5-bisphosphatecarboxylase/oxygenase (rubisco) in *Vigna unguiculata* L. Walp and *Eucalyptus diversicolor* F. Muell was possible using a not further specified "SDS-polymer running buffer." Main focus of this study was the question to which extent the presence of a nitrogen source has an influence on the concentration of rubisco in these plant species (Warren, Chen, and Adams, 2000).

6.6 Lipids

The term lipids encompasses fatty acids, sterols, glycerophospholipids, sphingolipids, and glycerolipids (Gotti, 2011). Especially for the first three

compound classes, CE has been utilized for analysis. The major free fatty acids (e.g., stearic, oleic, and linoleic acids) occurring in peanut (*Arachis hypogaea* L.) seeds were analyzed using a partially aqueous electrolyte system. It was composed of 40 mM Tris, 2.5 mM adenosine-5′-monophosphate (AMP), and 7 mM α-CD in a mixture of *N*-methylformamide, dioxane, and water (5 : 3 : 2). In this application, AMP served as background absorber for indirect UV detection of the fatty acids at 254 nm, whereas α-CD was required as selectivity modulator. Before analysis, a rather complicated sample cleanup was required, including several partitioning steps; yet, excellent recovery rates of 97% and higher were observed. In all samples analyzed, oleic acid was found to be the dominant fatty acid (Bannore *et al.*, 2008). A similar approach was reported for the analysis of 10 fatty acids in *Bertholletia excelsa* O. Berg oil (brazilnut oil), which were formed

during hydrogenation of the oil. Again, the analytes were monitored by indirect UV detection, this time using a phosphate buffer containing 4 mM SDS, 10 mM polyoxyethylene 23 lauryl ether (Brij 35), 2% 1-octanole, and 45% ACN. Under mild hydrogenation conditions, larger amounts of oleic and linoleic acids remained, under severe conditions predominantly stearic and elaidic acids were formed (de Oliveira *et al.*, 2003). For the separation of six anacardic acids (6-alkylsalicylic acids) from the shell liquid of cashew nuts, *Anacardium occidentale* L., two CE methods were developed. Conventional MEKC in uncoated silica capillaries was compared to RF-MEKC in polydimethylacrylamide-coated capillaries; in these capillaries, the residual EOF was reduced to 1% of the original value for bare-fused silica material. Nevertheless, by utilizing a 10 mM phosphate buffer containing 1 M urea, 20% ACN, 10 mM β-CD, and 1 mM heptakis-6-sulfo-β-CD better results were obtained at comparable analysis time (20 min, applied voltage −17.5 kV) (Cesla *et al.*, 2006).

Important membrane phospholipids such as phosphatidylcholine (PC), phosphatidylethanolamine (PE), and phosphatidylinositol (PI) were investigated by MEKC and NACE. The analytes were determined in chloroform/methanol extracts of *Glycine max* L., *Juglans regia* L., *Helianthus annuus* L., *Arachis hypogea* L., *Prunus armeniaca* L., and *Corylus avellana* L.; the solvent was removed and the obtained residue dissolved in the respective mobile phase. NACE with a BGE containing 0.3% acetic acid and 60 mM ammonium acetate in an ACN/2-propanol mixture (6 : 4) enabled faster separations compared to MEKC utilizing a cholate buffer. Otherwise, both approaches were comparable in many aspects (e.g., linear range, precision, and sensitivity) and the obtained quantitative results were in good agreement (Guo *et al.*, 2005a).

The CEC separation of unsaponifiable lipids, tocopherols, and plant sterols in rice bran oil (*Oryza sativa* L.) and soybean oil (*G. max* L.) was attempted by Abidi and colleagues. For this purpose, they compared packed capillaries filled with different stationary phases (pentafluorophenylsilica, triacontylsilica, and octadecylsilica) with regular HPLC, to conclude that CEC provided a much better resolution. The best results were obtained on pentafluorophenylsilica material (25 cm effective capillary length, 3 μm particle size; material from ES Industries, West Berlin, NJ, USA) with an ACN/Tris buffer at pH 8 (6:4, v/v;
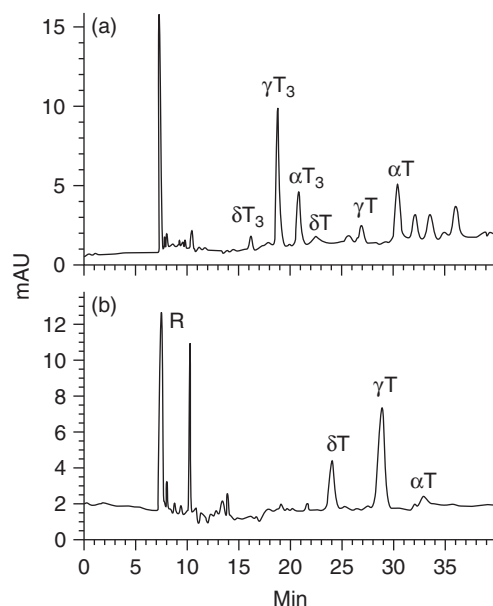
Figure 8 Analysis of rice bran oil (a) and soybean oil (b) by CEC using a pentafluorophenylsilica stationary phase; (T_3) tocotrienol derivatives and (T) tocopherol derivatives. (*Source*: Reprinted from J Chromatogr A, 949, Abidi SL, Thiam S, Warner IM, Elution behavior of unsaponifiable lipids with various capillary electrochromatographic stationary phases, 195–207, Copyright (2002), with permission from Elsevier.)

Figure 8). Even if the application on real life samples is presented, this study emphasized mainly on separation mechanisms and theoretical aspects of CEC (Abidi, Thiam, and Warner, 2002).

6.7 Organic Acids

The use of capillary electrophoretic methods for the analysis of organic acids, especially for those with a low molecular mass, is rather common and usually shows good results (Table 4). Many applications describe the use of alkaline buffer systems, so that negative separation voltages are applied for achieving faster results. One exemption is the analysis of 13 short chain organic acids, including oxalate, formate, and citrate, in natural rubber latex derived from *Hevea brasiliensis* (Willd. ex A. Juss.) Müll. Arg. In this case, the buffer was acidic (pH 6.25) and contained 0.5 mM CTAB, a cationic surfactant to suppress EOF; the applied voltage during analysis was −10 kV. Because of low sensitivity of some

Table 4 Selected applications of organic acid analysis by CE.

Applications/analytes	Buffer	Detection	Validated	Application	References
Hevea brasiliensis (e.g., nitrate, oxalate, formate, fumarate, aconitate, and succinate)	CZE/0.5 M phosphoric acid, 0.5 mM CTAB (pH 6.25)	UV (200 nm)	1–4	SA	(Galli, 2003)
Acer pseudoplatanus, Arabidopsis thaliana, Ranunculus glacialis (e.g., citric acid, malic acid, succinic acid, and oxalic acid)	CZE/20 mM trimellitate, 0.3 mM TTAB (pH 9.0)	UV (indirect at 240 nm)	1–4	SA+	(Rivasseau *et al.*, 2006)
Chrysanthemum morifolium, Prunella vulgaris, Morus alba (e.g., oxalate, valeriate, malate, and glutamate)	CZE/15 mM Tris, 3 mM 1,2,4-benzenetricarboxylic acid, 1.5 mM TEPA, and 20% methanol (pH 8.4)	UV (indirect at 254 nm)	4	SA	(Fung and Tung, 2001)
Chrysanthemum morifolium Morus alba, "*Var splendens*" (e.g., benzoate, lactate, glycolate, and formate)	CZE/200 mM histidine, 0.1 mM CTAB, 0.025% hydroxypropyl-β-CD, and 10% methanol; pH 6.0	C^4D	1, 4	SA+	(Law *et al.*, 2007)
Aristolochia fangchi (aristolochic acid I and aristolochic acid II)	CZE/100 mM phosphate, 10% acetonitrile, and 3 mM cucurbit(7)uril; pH 7.5	UV (254 nm)	1, 4	SA+	(Wei and Feng, 2008)

Validated, 1 (sensitivity), 2 (specificity), 3 (accuracy) and 4 (precision); Application, S (standard), SA (sample) and + (quantification).

of the analytes, sample stacking was utilized, which required samples and standards being dissolved in a 1 : 1 mixture of ACN and 0.5% sodium chloride solution. Enhancement factors over 16 were observed using this technique (Galli, 2003). Similar analytes were also investigated in a common Chinese vegetable, which cannot be further defined as the provided botanical name (*Var splendens*) is incorrect. Again, an EOF suppressing additive (0.6 mM tetradecyltrimethylammoniumbromide, TTAB) was added to the BGE, which also contained 15 mM phthalate for indirect detection of the analytes at 254 nm; the applied voltage was −15 kV (Wang *et al.*, 2003). Comparable procedures have been reported for the analysis of organic acids in *Acer pseudoplatanus* L., *Arabidopsis thaliana* Schur., and *Ranunculus glacialis* L. (Rivasseau *et al.*, 2006), in corn (*Zea mays*) exposed to cadmium stress (Guo *et al.*, 2007), in *Brachiaria brizantha* Stapf (Simas Vaz *et al.*, 2012), or in diverse TCM plants such as *Chrysanthemum morifolium* Ramat. and *Morus alba* L. (Law *et al.*, 2007). In all of them, an EOF suppressant is used in buffers with pH values ranging from slightly acidic to pH 9.0. The latter enabled excellent and fast separations (analysis time 1.8 min) combined with detection limits in the low pmol range (e.g., LOD for oxalate: 0.082 pmol) (Rivasseau *et al.*, 2006). Detection was

performed either in indirect UV-mode or by C^4D, using a 200-mM histidine buffer solution containing CTAB, hydroxypropyl-β-CD, and methanol (Law *et al.*, 2007).

To investigate the antioxidant capacity of four phenolic acids (danshensu, salvianolic acid B, caffeic acid, and lithospermic acid) occurring in the root and rhizome of *Salvia miltiorrhiza* Bunge, a FASS-CZE-ESI-MS method has been described. This approach combined a simple CZE separation, using a 10 mM ammonium acetate buffer at pH 7.0, with a FASS preconcentration step in order to increase detection sensitivity. The antioxidant potential of the analytes was studied by adding 0.3 mM DPPH, a free radical, to the sample solutions. In case an antioxidant was present the peak area of DPPH decreased, so that the activity of the respective phenolic acid could be deduced indirectly (Duan, Cao, and Zhang, 2012).

For the separation of three low molecular organic acids (oxalate, malate, and citrate), sugars and glutamate in different vegetables such as tomato (*S. lycopersicum* L.), pepper (*Capsicum annuum* L.), or orange (*Citrus sinensis* Osbeck), a running buffer consisting of 20 mM 2,6-pyridine dicarboxylic acid at pH 12.1 and 0.1% hexadimethrinebromide (HDM) was most suitable. To increase repeatability, the capillary was rinsed for 2 min with 58 mM SDS solution
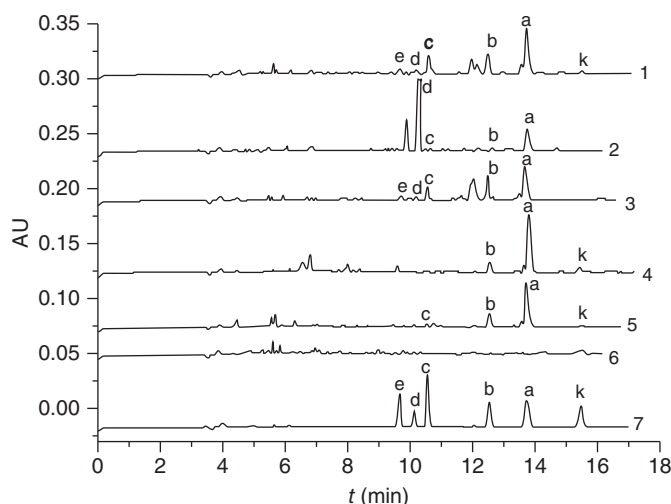
Figure 9 Determination of aristolochic acids by MEEKC; (e) 7-hydroxy-aristolochic acid A, (d) aristolochic acid D, (c) aristolochic acid C, (b) aristolochic acid B, (a) aristolochic acid A, (k) aristolochic acid. Traces: Herba Aristolochiae (1), Fructus Aristolochiae (2), Radix Aristolochiae (3), Radix Aristolochiae Fangchi (4), Caulis Aristolochiae Manchuiriensis (5), Herba Asari (6), and standards (7). (*Source*: Reprinted from Anal Chim Acta, 561, Zhai Z, Luo X, Shi Y, Separation and determination of aristolochic acids in herbal medicines by microemulsion electrokinetic chromatography, 119–125, Copyright (2006), with permission from Elsevier.)

between two injections before flushing it with BGE for 5 min. This procedure showed to be very effective to increase peak resolution and to reduce analysis time. In addition, up to 200 consecutive injections were possible, if the buffer is replaced every 5 h (Cebolla-Cornejo *et al.*, 2012).

Among the group of medicinal plants associated with safety concerns is *Aristolochia fangchi* Y.C. Wu ex L.D. Chow & S.M. Hwang, a TCM remedy used as an traditional aid in childbirth. Several years ago herbal slimming products were adulterated with this plant, and severe kidney diseases were observed (Chinese herb nephropathy). Responsible for these effects are nitrophenanthrene carboxylic acids (aristolochic acids), and one option to determine these toxic compounds is by CE. Wei *et al*. used a simple CZE method with a 100 mM phosphate buffer (pH 7.5), containing 10% ACN, and 3 mM cucurbit(7)uril (CB[7]) for the assessment of aristolochic acids I and II. CB[7] is a macrocyclic molecule, being able to form inclusion complexes just like cyclodextrins. This additive, in combination with sample preconcentration by FESI, enabled the selective and sensitive determination of the two target compounds (Wei and Feng, 2008). Another CE method for the same two analytes used a 20 mM phosphate buffer at pH 10 in combination with EC detection (Zhou *et al*., 2006). A third option for separation of these compounds

is MEEKC. An O/W buffer system was prepared by mixing 0.81% octane, 3.31% SDS, and 6.61% butan-1-ol with 89.27% of a 10 mM borax solution at pH 9.2. After method validation ($R \geq 0.005$; LOD $\leq 4.6\ \mu g\ mL^{-1}$; recovery rates $\geq 97.2\%$), this method showed to be ideal for the quantitative analysis of diverse aristolochic acids (AA-a, AA-b, AA-c, AA-d, AA-e, and AA-k) in plant material (Figure 9). In all *Aristolochia* specimens, AA-1 was dominant (4.7–$13.0\ mg\ g^{-1}$), whereas AA-c and AA-k were detected in traces only (Zhai, Luo, and Shi, 2006). To analyze aristolochic acids in 61 different Chinese herb samples, laser-induced fluorescence detection was used after an initial CE separation under MEKC conditions (50 mM borax containing 10 mM SDS at pH 9.0). The analytes had to be reduced with iron powder before detection because aristolochic acids show no intrinsic fluorescence (Hsieh *et al*., 2006).

Many respective publications deal with the analysis of organic acids in food and beverages, especially in wine. For example, a CZE method for the determination of tartaric, malic, succinic, acetic, lactic, and citric acid in red, white, and rosé wine, as well as tartaric, malic, and citric acid in grape juice was developed by Mato, Suárez-Luque, and Huidobro (2007). The reported separation conditions do not differ much from already described ones, except that direct UV detection at 185 nm was performed and

the buffer contained a specific EOF modifier, tetrade-cyltrimethylammonium hydroxide (TTAOH) (Mato, Suárez-Luque, and Huidobro, 2007). Other groups were also working on this topic, including the analysis of organic acids in 39 red wines from Spain (Castiñeira, Peña, and Herrero, 2002), or the CZE separation of the six organic acids in the same matrix using a microchip (Masár *et al.*, 2005). The list of food items analyzed by CE for organic acids can further be extended to coffee (Galli, 2004), beer (Cortacero-Ramírez *et al.*, 2005), fruit juices (Tang and Wu, 2005), honey (Mato *et al.*, 2006), or milk products such as whey (Buiarelli *et al.*, 2003) and cheese (Izco, Tormo, and Jiménez-Flores, 2002).

6.8 Tannins and Other Polyphenols

Tannins are water soluble, weakly acidic polyphenolic compounds (usual mass range is from 500 to 3000 Da) that occur in plants as secondary metabolites; they are subdivided into two groups, which are hydrolysable tannins and proanthocyanidins (Bronze, Vilas Boas, and Belchior, 1997). Most of the CE methods reported focus on the analysis of monomeric or smaller building blocks such as gallic acid or ellagic acid (Table 5). For example, both have been determined in extracts of *Quercus robur* L. (oak) and barrel-aged brandy using a buffer prepared from 0.1 M borax and 5% ethanol at pH 9.2 (Bronze,

Vilas Boas, and Belchior, 1997); similar electrolyte systems have been described for the separation of trans-resveratrol, catechin, and gallic acid in *Myrica gale* L., *H. rhamnoides* L., and *Reynoutria japonica* Houtt. (Vaher and Koel, 2003) or in different types of berries (e.g., blueberries and raspberries) (Ehala, Vaher, and Kaljurand, 2005). The analysis of hydrolysable tannins in *Terminalia chebula* Retz. fruits was attempted by CZE and MEKC. The determined analytes include punicalagin, pentagalloyl glucose, casuarinin, and the two above-mentioned acids. Better results in respect to resolution were obtained with MEKC (25 mM borax, 5 mM sodium phosphate, 20 mM SDS, and 10% ACN at pH 7.0), in which the sample was injected in electrokinetic mode (5 kV for 3 s) using protocatechuic acid as internal standard (Juang, Sheu, and Lin, 2004).

6.9 Fingerprinting

The characteristic pattern of a large number of plant constituents shown in a chromatogram or electropherogram can be a suitable way for identification and differentiation. Instead of a few marker compounds, the whole metabolic profile is reflected and compared, an approach that is especially popular in Asia where traditional medicine is mostly based on holistic principles. Accordingly, fingerprinting is recommended by the Chinese Pharmacopeia as a

Table 5 Selected applications of tannin analysis by CE.

Applications/analytes	Buffer	Detection	Validated	Application	References
Quercus robur (e.g., gallic acid, ellagic acid, rutin, and ferulic acid)	CZE/0.1 M borax buffer, 5% ethanol; pH 9.2	UV (280 nm)	1, 4	SA	(Bronze, Vilas Boas, and Belchior, 1997)
Myrica gale, *Hippophae rhamnoides*, and *Reynoutria japonica* (e.g., trans-resveratrol, catechin, quercetin, and gallic acid)	CZE/25 mM borax buffer; pH 9.4	UV (240 nm)	1	SA	(Vaher and Koel, 2003)
Vaccinium myrillus, *Vaccinium vitis-idaea*, *Vaccinium oxycoccus*, *Fragaria ananassa*, *Ribes nigrum*, and *Ribes rubrum* (e.g., (+)-catechin, cinnamic acid, ferulic acid, quercetin, morin, and gallic acid)	CZE/35 mM borax, 5% methanol; pH 9.3	UV (210 nm)	1–4	SA+	(Ehala, Vaher, and Kaljurand, 2005)
Fucus vesiculosus (phloroglucinol, gallic acid, and catechin)	CZE/20 mM borax, 50 mM sodium acetate, and acetonitrile/methanol/acetic acid = 74 : 25 : 1; at pH 9.0	UV (210 nm)	1	SA	(Truus *et al.*, 2004)

Validated, 1 (sensitivity), 2 (specificity), 3 (accuracy) and 4 (precision); Application, S (standard), SA (sample) and + (quantification).
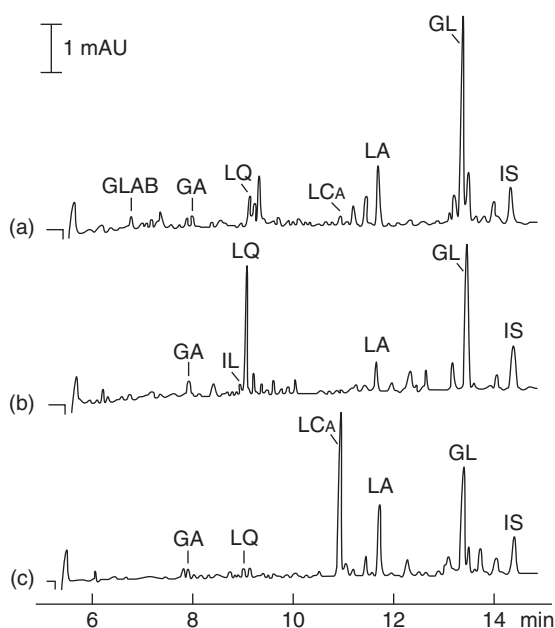
Figure 10 Fingerprinting of *Glycyrrhiza glabra* (a), *Glycyrrhiza uralensis* (b) and *Glycyrrhiza inflata* (c) by CE with the assignment of several marker compounds; GLAB, glabridin; GA, glycyrrhetic acid; IL, isoliquiritin; LQ, liquiritin; LC$_A$, licochalcone A; LA, liquiritin apioside; GL, glycyrrhizin; and IS, internal standard = cinnamic acid. (*Source*: Reprinted from J Pharm Biomed Anal, 38, Rauchensteiner F, Matsumura Y, Yamamoto Y, Yamaji S, Tani T, Analysis and comparison of Radix Glycyrrhizae (licorice) from Europe and China by capillary-zone electrophoresis (CZE), 594–600, Copyright (2005), with permission from Elsevier.)

potential and reliable strategy for the quality control of complex mixtures such as herbal medicines.

Many CE fingerprinting procedures can be found in literature, most of them also enabling the determination of individual marker compounds. The above-mentioned method by Rauchensteiner *et al*. (Section 6.4.4; Figure 10) is one example. It utilizes fingerprints as well as the quantitative determination of individual compounds for the differentiation of several licorice species (Rauchensteiner *et al*., 2005). Another typical example is the comparison of different batches of *S. baicalensis* Georgi and *S. amoena* Wright (Wang *et al*., 2005). The authors first optimized the extraction procedure to observe the maximum number and height of peaks. Then samples were extracted accordingly and compared to find signals occurring in all of them (common peaks). These signals were used to construct an average electropherogram, and based on the relative peak

areas of common peaks in the individual samples, the "sameness" of specimens could be deduced. The latter is possible by calculating correlation coefficients or angles of cosine. The CE separation itself was performed using a BGE typical for flavonoid analysis, that is, an alkaline borax buffer (Wang *et al*., 2005). Another interesting application in this respect is the differentiation of Flos Carthami (*Carthamus tinctorius* L.) from its substitute in Chinese medicine, Stigma Croci (*Crocus sativus* L.), and an adulterant, Flos Hemerocallis (*Hemerocallis citrina* L.). Two methods were developed for this purpose, one CZE assay (Sun *et al*., 2003) and another one by CEC (Xie *et al*., 2005). Sun *et al*. used three reference compounds (adenosine, rutin, and quercetin) to develop the assay, and based on the results of 10 sample batches, a clear differentiation of the three plant species was possible. The CEC procedure was performed on a commercially available packed capillary (EP-150-30/50-5-C18; Unimicro Technologies) in gradient mode (solvent A: 0.02% TFA in water; solvent B: 0.02% TFA in 95% aqueous methanol). Hydroxy safloryellow A served as reference to identify 43 "common peaks" (by HPLC only 21 could be identified). Accordingly, the authors considered CEC advantageous as it "showed better performance and contained more chemical information"; validation of the assay was not attempted. One last example out of many more is the characterization of *S. miltiorrhiza* by NACE fingerprint. The optimum electrolyte contained 125 mM borax and 50 mM sodium deoxycholate in methanol at a pH of 10; it allowed the assignment of seven compounds (some were identified as tashinones) in the plant extracts. These results were compared to the separation achieved by high speed countercurrent chromatography (HSCCC), an approach that was found to be better suited for tashinones. Yet, by NACE, more peaks could be resolved (Gu *et al*., 2004).

7   CONCLUSIONS

CE, with all its versatility in respect to separation and detection mode as well as many desirable operational features, definitely deserves a better standing in the field of analytical chemistry in general and natural products analysis in particular. Sometimes raised concerns regarding poor reproducibility and

sensitivity can be resolved most of the time or at least reduced to an acceptable minimum by utilizing a combination of optimum techniques and investing time in method development and validation. Then, the obtained results will often not only be comparable to established approaches but also superior. That this final statement is based on solid scientific evidences rather than on wishful thinking has hopefully been shown in this chapter.

8 RELATED ARTICLES

Thin-layer Chromatography, with Chemical and Biological Detection Methods; HPLC and Ultra HPLC: Basic Concepts; Near Infrared Spectroscopy (NIR) in Natural Product Research; Headspace Sampling and Gas Chromatography: A Successful Combination to Study the Composition of a Plant Volatile Fraction; LC and LC-MS: Techniques and Applications; NMR as Analytical Tool for Crude Plant Extracts.

REFERENCES

Abidi, S. L., Thiam, S. and Warner, I. M. (2002) *J. Chromatogr. A*, **949**, 195–207.

Altria, K. D. (2010) Fundamentals of capillary electrophoresis theory, in *Capillary Electrophoresis Guidebook*, ed. K. D. Altria, Humana Press, Totowa, pp. 3–14.

Aranas, A. T., Guidote, A. M. Jr., and Quirino, J. P. (2009) *Anal. Bioanal. Chem.*, **394**, 175–85.

Arraez-Roman, D., Zurek, G., Bassmann, C., *et al.* (2008) *Electrophoresis*, **29**, 2112–6.

Bachmann, S., Huck, C. W., Bakry, R., *et al.* (2007) *Electrophoresis*, **28**, 799–805.

Bannore, Y. C., Chenault, K. D., Melouk, H. A., *et al.* (2008) *J. Sep. Sci*, **31**, 2667–2676.

Bonneil, E., Young, N. M., Lis, H., *et al.* (2004) *Arch. Biochem. Biophys.*, **426**, 241–249.

Breadmore, M. C., Dawod, M. and Quirino, J. P. (2011) *Electrophoresis*, **32**, 127–148.

Bronze, M. R., Vilas Boas, L. F. and Belchior, A. P. (1997) *J. Chromatogr. A*, **798**, 143–152.

Buiarelli, F., Cartoni, G., Coccioli, F., *et al.* (2003) *J. Sep. Sci.*, **26**, 425–428.

Bytzek, A. K., Reithofer, M. R., Galanski, M., *et al.* (2010) *Electrophoresis*, **31**, 1144–50.

Cao, Y., Wang, Y., Ji, C., *et al.* (2004) *J. Chromatogr. A*, **1042**, 203–209.

Carrasco-Pancorbo, A., Gómez-Caravaca, A. M., Cerretani, L., *et al.* (2006) *J. Sep. Sci.*, **29**, 2221–2233.

Carrasco-Pancorbo, A., Neususs, C., Pelzing, M., *et al.* (2007) *Electrophoresis*, **28**, 806–821.

Carrasco-Pancorbo, A., Gómez-Caravaca, A. M., Segura-Carretero, A., *et al.* (2009) *J. Sci. Food Agric.*, **89**, 2144–2155.

Castiñeira, A., Peña, R. M., Herrero, C., *et al.* (2002) *J. Food Compos. Anal.*, **15**, 319–331.

Cebolla-Cornejo, J., Valcarcel, M., Herrero-Martinez, J. M., *et al.* (2012) *Electrophoresis*, **33**(15), 2416–23.

Cesla, P., Blomberg, L., Hamberg, M., *et al.* (2006) *J. Chromatogr. A*, **1115**, 253–259.

Chen, G. (2007) *Talanta*, **74**, 326–32.

Chen, G. H. and Liu, W. M. (2005) *J. Liq. Chromatogr. Related Technol.*, **28**, 223–232.

Chen, A., Li, C., Gao, W., *et al.* (2005) *J. Pharm. Biomed. Anal.*, **37**, 811–816.

Chen, Q., Li, P., Cheng, F., *et al.* (2009a) *Chromatographia*, **69**, 1443–1446.

Chen, X. J., Tu, P. F., Jiang, Y., *et al.* (2009b) *J. Sep. Sci.*, **32**, 275–81.

Cherkaoui, S., Geiser, L. and Veuthey, J. L. (2000) *Chromatographia*, **52**, 403–407.

Cheung, H. Y. and Zhang, Q. F. (2008) *J. Chromatogr. A*, **1213**, 231–238.

Chi, L., Li, Z., Dong, S., *et al.* (2009) *Microchim. Acta*, **167**, 179–185.

Chizzali, E., Nischang, I. and Ganzera, M. (2011) *J. Sep. Sci.*, 34.

Cortacero-Ramírez, S., Segura-Carretero, A., Hernáinz-Bermúdez de Castro, M., *et al.* (2005) *J. Chromatogr. A*, **1064**, 115–119.

Diagone, C. A., Ogawa, C. A. and Lancas, F. M. (2003) *J. Liq. Chromatogr. Related Technol.*, **26**, 505–516.

Do, T. C. M. V., Nguyen, T. D., Tran, H., *et al.* (2012) *Planta Med.*, **78**, 1796–1799.

Dong, J., Ou, J., Dong, X., *et al.* (2007) *J. Sep. Sci.*, **30**, 2986–92.

Dong, Y., Leu, Y.-L., Chien, K.-Y., *et al.* (2009) *Anal. Lett.*, **42**, 1444–1457.

Duan, Q., Cao, J. and Zhang, J. (2012) *Anal. Methods*, **4**(9), 3027.

Dubber, M. J. and Kanfer, I. (2006) *J. Chromatogr. A*, **1122**, 266–274.

Edwards, E. L., Rodrigues, J. A., Ferreira, J., *et al.* (2006) *Electrophoresis*, **27**, 2164–2170.

Eeltink, S. and Kok, W. T. (2006) *Electrophoresis*, **27**, 84–96.

Ehala, S., Vaher, M. and Kaljurand, M. (2005) *J. Agric. Food Chem.*, **53**, 6484–6490.

El Deeb, S., Iriban, M. A. and Gust, R. (2011) *Electrophoresis*, **32**, 166–83.

Escarpa, A., Gonzalez, M. C., Lopez Gil, M. A., *et al.* (2008) *Electrophoresis*, **29**, 4852–61.

Fakhari, A. R., Nojavan, S., Ebrahimi, S. N., *et al.* (2010) *J. Sep. Sci.*, **33**, 2153–9.

Fan, D.-H., Wang, H., Zhi, D., *et al.* (2010) *Chromatographia*, **72**, 1013–1016.

Francis, P. S., Adcock, J. L., Costin, J. W., *et al.* (2008) *J. Pharm. Biomed. Anal.*, **48**, 508–18.

Freitag, R. and Hilbrig, F. (2007) *Electrophoresis*, **28**, 2125–44.

Fung, Y. S. and Tung, H. S. (2001) *Electrophoresis*, **22**, 2242–2250.

Furlanetto, S., Cirri, M., Piepel, G., *et al.* (2011) *J. Pharm. Biomed. Anal.*, **55**, 610–7.

Gahm, K. H., Chang, L. W. and Armstrong, D. W. (1997) *J. Chromatogr. A*, **759**, 149–155.

Galli, V. (2003) *Anal. Chim. Acta*, **482**, 37–45.

Galli, V. (2004) *J. Chromatogr. B*, **1032**, 299–304.

Ganzera, M. and Nischang, I. (2010) *Curr. Org. Chem.*, **14**, 1769–1780.

Ganzera, M., Stöggl, W. M., Bonn, G. K., *et al.* (2003) *J. Sep. Sci.*, **26**, 1383–1388.

Ganzera, M., Piereder, D., Sturm, S., *et al.* (2005) *Electrophoresis*, **26**, 1724–1731.

Ganzera, M., Krüger, A. and Wink, M. (2010) *J. Pharm. Biomed. Anal.*, **53**, 1231–5.

Ganzera, M., Greifeneder, V., Schwaiger, S., *et al.* (2012) *Fitoterapia*, **83**(8), 1680–6.

Gao, W., Lin, S., Chen, Y., *et al.* (2005) *J. Sep. Sci.*, **28**, 639–46.

Gao, Y., Xiang, Q., Xu, Y., *et al.* (2006) *Electrophoresis*, **27**, 4842–8.

Gao, Y., Xu, Y., Han, B., *et al.* (2009) *Talanta*, **80**, 448–53.

Geiser, L. and Veuthey, J. L. (2009) *Electrophoresis*, **30**, 36–49.

Gomez-Caravaca, A. M., Carrasco-Pancorbo, A., Segura-Carretero, A., *et al.* (2009) *Electrophoresis*, **30**, 3099–3109.

Gong, Y. X., Li, S. P., Wang, Y. T., *et al.* (2005) *Electrophoresis*, **26**, 1778–82.

Gotti, R. (2011) *J. Pharm. Biomed. Anal.*, **55**, 775–801.

Gu, C. and Shamsi, S. A. (2010) *Electrophoresis*, **31**, 1162–74.

Gu, M., Zhang, S., Su, Z., *et al.* (2004) *J. Chromatogr. A*, **1057**, 133–140.

Guo, B., Wen, B., Shan, X., *et al.* (2005a) *J. Chromatogr. A*, **1074**, 205–213.

Guo, Q. L., Gao, J. Y. and Yang, J. S. (2005b) *Chromatographia*, **62**, 145–150.

Guo, B. Y., Peng, Z. L., Han, F., *et al.* (2007) *J. Sep. Sci.*, **30**, 2742–2747.

Guo, L., Fu, F. and Chen, G. (2011) *Anal. Bioanal. Chem.*, **399**, 3323–43.

Guzman, N. A. and Phillips, T. M. (2011) *Electrophoresis*, **32**, 1565–78.

Ha, P. T., Hoogmartens, J. and Van Schepdael, A. (2006) *J. Pharm. Biomed. Anal.*, **41**, 1–11.

Hamoudova, R., Pospisilova, M. and Spilkova, J. (2006) *Electrophoresis*, **27**, 4820–4826.

Hänsel, R. and Sticher, O. (2007a) *Pharmakognosie Phytopharmazie*, Springer Medizin Verlag, Heidelberg, pp. 810; 882; 815–817;916.

Hänsel, R. and Sticher, O. (2007b) *Pharmakognosie – Phytopharmazie*, Springer Medizin Verlag, Heidelberg, p. 706.

Helmja, K., Vaher, M., Pussa, T., *et al.* (2008) *Electrophoresis*, **29**, 3980–3988.

Herrero, M., Arraezroman, D., Segura, A., *et al.* (2005a) *J. Chromatogr. A*, **1084**, 54–62.

Herrero, M., Simo, C., Ibanez, E., *et al.* (2005b) *Electrophoresis*, **26**, 4215–4224.

Hinneburg, I., Mrestani, Y. and Neubert, R. H. H. (2004) *Chromatographia*, **59**, 591–594.

Hjerten, S. (1967) *Chromatogr. Rev.*, **9**, 122–219.

Hsieh, S. C., Huang, M. F., Lin, B. S., *et al.* (2006) *J. Chromatogr. A*, **1105**, 127–134.

Hu, X., Cui, S. and Liu, J. Q. (2010) *Chromatographia*, **72**, 993–997.

Huhn, C., Ramautar, R., Wuhrer, M., *et al.* (2010) *Anal. Bioanal. Chem.*, **396**, 297–314.

Izco, J. M., Tormo, M. and Jiménez-Flores, R. (2002) *J. Diary Sci.*, **85**, 2122–2129.

Jiang, T. F., Du, X. and Shi, Y. P. (2004a) *Chromatographia*, **59**, 255–258.

Jiang, T. F., Li, Y. and Shi, Y. P. (2004b) *Planta Med.*, **70**, 284–287.

Jiang, T.-F., Lv, Z.-H. and Wang, Y.-H. (2005) *J. Sep. Sci.*, **28**, 2225–2229.

Jiang, T.-F., Wang, Y.-H., Lv, Z.-H., *et al.* (2006) *Chromatographia*, **65**, 101–104.

Jiang, T. F., Wang, Y. H., Lv, Z. H., *et al.* (2007) *J. Pharm. Biomed. Anal.*, **43**, 854–858.

Jorgenson, J. and Lukacs, K. D. (1981) *Anal. Chem.*, **53**, 1298–1302.

Juang, L. J., Sheu, S. J. and Lin, T. C. (2004) *J. Sep. Sci.*, **27**, 718–724.

Karger, B. L. and Foret, F. (1993) Capillary electrophoresis: introduction and assessment, in *Capillary Electrophoresis Technology*, ed. N. A. Guzman, Dekker, New York, pp. 3–65.

Karniadakis, G., Beskok, A. and Aluru, N. (2005) *Microflows and nanoflows*, Springer, New York, pp. 266–291.

Kitagawa, F., Tsuneka, T., Akimoto, Y., *et al.* (2006) *J. Chromatogr. A*, **1106**, 36–42.

Knox, J. H. and Grant, I. H. (1991) *Chromatographia*, **32**, 317–328.

Koyama, J., Morita, I. and Kobayashi, N. (2007) *J. Chromatogr. A*, **1145**, 183–9.

Kulp, M., Bragina, O., Kogerman, P., *et al.* (2011) *J. Chromatogr. A*, **1218**, 5298–304.

Lämmerhofer, M. and Gargano, A. (2010) *J. Pharm. Biomed. Anal.*, **53**, 1091–123.

Law, W. S., Zhao, J. H., Hauser, P. C., *et al.* (2007) *J. Sep. Sci.*, **30**, 3247–3254.

Lee, I. S. L., Boyce, M. C. and Breadmore, M. C. (2012) *Food Chem.*, **133**(1), 205–211.

Li, S. F. Y. (1992) *Capillary Electrophoresis*, Elesvier, Amsterdam, The Netherlands, pp. 128–201.

Li, F., Cao, Q.-E. and Ding, Z. (2004a) *Chromatographia*, **59**.

Li, Y., Qi, S., Chen, X., *et al.* (2004b) *Electrophoresis*, **25**, 3003–9.

Li, Y., Qi, S., Chen, X., *et al.* (2005) *Talanta*, **65**, 15–20.

Li, J., Chun, Y. and Ju, H. (2007) *Electroanalysis*, **19**, 1569–1574.

Li, J., Huang, X., Lai, D., *et al.* (2008a) *Chromatographia*, **67**, 989–993.

Li, Y. L., Ding, C. X., Wang, H. L., *et al.* (2008b) *J. Anal. Chem.*, **63**, 574–579.

Lin, X., Xue, L., Zhang, H., *et al.* (2005) *Anal. Bioanal. Chem.*, **382**, 1610–1615.

Lin, X., Wang, J., Li, L., *et al.* (2007) *J. Sep. Sci.*, **30**, 3011–7.

Liu, F. K. (2009) *J. Chromatogr. A*, **1216**, 9034–47.

Liu, Q., Liu, Y., Li, Y., *et al.* (2006) *J. Sep. Sci.*, **29**, 1268–1274.

Lu, Q., Copper, C. L. and Collins, G. E. (2006) *Anal. Chim. Acta*, **572**, 205–11.

Lu, H., Wang, J., Wang, X., *et al.* (2007) *J. Pharm. Biomed. Anal.*, **43**, 352–7.

Luo, M., Lu, H., Ma, H., *et al.* (2007) *J. Pharm. Biomed. Anal.*, **44**, 881–886.

Lurie, I. S., Panicker, S., Hays, P. A., *et al.* (2003) *J. Chromatogr. A*, **984**, 109–120.

Lynen, F., Saveedra, L., Nickerson, B., *et al.* (2011) *Talanta*, **84**, 724–9.

Marchart, E., Krenn, L. and Kopp, B. (2003) *Planta Med.*, **69**, 452–456.

Marchi, I., Rudaz, S. and Veuthey, J. L. (2009) *Talanta*, **78**, 1–18.

Martin, R. S. (2006) Interfacing amperometric detection with microchip capillary electrophoresis, in *Microchip Capillary Electrophoresis*, ed. C. S. Henry, Humana Press, Totowa 85–112.

Masár, M., Poliaková, K., Danková, M., *et al.* (2005) *J. Sep. Sci.*, **28**, 905–914.

Mato, I., Huidobro, J. F., Simallozano, J., *et al.* (2006) *J. Agric. Food. Chem.*, **54**, 1541–1550.

Mato, I., Suárez-Luque, S. and Huidobro, J. F. (2007) *Food Chem.*, **102**, 104–112.

Maxwell, E. J. and Chen, D. D. (2008) *Anal. Chim. Acta*, **627**, 25–33.

Molnarperl, I. and Fuzfai, Z. (2005) *J. Chromatogr. A*, **1073**, 201–227.

Müller, A., Ganzera, M. and Stuppner, H. (2004) *Chromatographia*, **60**, 193–197.

Musenga, A., Mandrioli, R., Ferranti, A., *et al.* (2007) *J. Sep. Sci.*, **30**, 612–619.

Natishan, T. (2005) *J. Liq. Chromatogr. Related Technol.*, **28**, 1115–1160.

Okunji, C. O., Ware, T. A., Hicks, R. P., *et al.* (2002) *Planta Med.*, **68**, 440–444.

Olivares, J. A., Nguyen, N. T., Yonker, C. R., *et al.* (1987) *Anal. Chem.*, **59**, 1230–1232.

de Oliveira, M. A., Solis, V. E., Gioielli, L. A., *et al.* (2003) *Electrophoresis*, **24**, 1641–1647.

Oztekin, N., Baskan, S., Evrim Kepekci, S., *et al.* (2010) *J. Pharm. Biomed. Anal.*, **51**, 439–442.

Peng, Y. Y., Ye, J. N. and Kong, J. L. (2005) *J. Agric. Food Chem.*, **53**, 8141–8147.

Phuong, N. T., Lee, K. A., Jeong, S. J., *et al.* (2006) *J. Pharm. Biomed. Anal.*, **40**, 56–61.

Qi, S., Ding, L., Tian, K., *et al.* (2006) *J. Pharm. Biomed. Anal.*, **40**, 35–41.

Rahfeld, B. (2011) *Mikroskopischer Farbatlas pflanzlicher Drogen*, Spektrum Akademischer Verlag, Heidelberg.

Ramautar, R., Somsen, G. W. and de Jong, G. J. (2010) *Electrophoresis*, **31**, 44–54.

Rauchensteiner, F., Matsumura, Y., Yamamoto, Y., *et al.* (2005) *J. Pharm. Biomed. Anal.*, **38**, 594–600.

Reddy, M. M., Suresh, V., Jayashanker, G., *et al.* (2003) *Electrophoresis*, **24**, 1437–1441.

Rivasseau, C., Boisson, A. M., Mongelard, G., *et al.* (2006) *J. Chromatogr. A*, **1129**, 283–290.

Ryan, R., Donegan, S., Power, J., *et al.* (2009) *Electrophoresis*, **30**, 65–82.

Ryan, R., McEvoy, E., Donegan, S., *et al.* (2011) *Electrophoresis*, **32**, 184–201.

Sabbioni, C., Mandrioli, R., Ferranti, A., *et al.* (2005) *J. Chromatogr. A*, **1081**, 65–71.

Sawalha, S. M. S., Arráez-Román, D., Segura-Carretero, A., *et al.* (2009) *Food Chem.*, **116**, 567–574.

Scriba, G. K. (2008) *J. Sep. Sci.*, **31**, 1991–2011.

Shaban, E., Gajdošová, D. and Havel, J. (2006) *J. Sep. Sci.*, **29**, 1174–1179.

Shang, X. and Yuan, Z. (2003) *Bioorg. Med. Chem. Lett.*, **13**, 617–622.

Shpak, A. V., Pirogov, A. V. and Shpigun, O. A. (2004) *J. Chromatogr. B*, **800**, 91–100.

Silva, M. (2007) *Electrophoresis*, **28**, 174–92.

Simas Vaz, F. A., da Silva, P. A., Passos, L. P., *et al.* (2012) *Phytochem. Anal.*, **23**(6), 569–75.

Simpson, S. L. Jr., Quirino, J. P. and Terabe, S. (2008) *J. Chromatogr. A*, **1184**, 504–41.

Song, J. Z., Han, Q. B., Qiao, C. F., *et al.* (2010) *Phytochem. Anal.: PCA*, **21**, 137–43.

Stoggl, W. M., Huck, C. W., Stecher, G., *et al.* (2006) *Electrophoresis*, **27**, 787–792.

Sturm, S. and Stuppner, H. (2001) *Chromatographia*, **53**, 612–618.

Sturm, S., Seger, C. and Stuppner, H. (2007) *J. Chromatogr. A*, **1159**, 42–50.

Sun, S.-W. and Tseng, H.-M. (2004) *J. Pharm. Biomed. Anal.*, **36**, 43–48.

Sun, S. W. and Tseng, H.-M. (2005) *J. Pharm. Biomed. Anal.*, **37**, 39–45.

Sun, Y., Guo, T., Sui, Y., *et al.* (2003) *J. Chromatogr. B*, **792**, 147–152.

Sung, M. W. and Li, P. C. (2004) *Electrophoresis*, **25**, 3434–3440.

Suntornsuk, L. (2010) *Anal. Bioanal. Chem.*, **398**, 29–52.

Tang, Y. and Wu, M. (2005) *Talanta*, **65**, 794–798.

Terabe, S. (2008) *Chem. Rec.*, **8**, 291–301.

Terabe, S., Otsuka, K., Ichikawa, K., *et al.* (1984) *Anal. Chem.*, **56**, 111–113.

Tessier, B., Schweizer, M., Fournier, F., *et al.* (2005) *Food Res. Int.*, **38**, 577–584.

Tian, J. and Qin, W. (2009) *Anal. Sci.*, **25**, 1119–1123.

Tian, K., Zhang, H., Chen, X., *et al.* (2006) *J. Chromatogr. A*, **1123**, 134–7.

Tian, K., Wang, Y., Chen, Y., *et al.* (2007) *Talanta*, **72**, 587–93.

Timerbaev, A. R. (2009) *Trends Anal. Chem.*, **28**, 416–425.

Truus, K., Vaher, M., Koel, M., *et al.* (2004) *Anal. Bioanal. Chem.*, **379**, 849–852.

Unger, M. (2009) *Planta Med.*, **75**, 735–45.

Urbonaviciute, A., Jakstas, V., Kornysova, O., *et al.* (2006) *J. Chromatogr. A*, **1112**, 339–344.

Vaher, M. and Koel, M. (2003) *J. Chromatogr. A*, **990**, 225–230.

Vshivkov, S., Pshenichnov, E., Golubenko, Z., *et al.* (2012) *J. Chromatogr. B*, **908**, 94–7.

Wang, M., Qu, F., Shan, X., *et al.* (2003) *J. Chromatogr. A*, **989**, 285–292.

Wang, L. C., Cao, Y. H., Xing, X. P., *et al.* (2005) *Chromatographia*, **62**, 283–288.

Wang, S., Ye, S. and Cheng, Y. (2006) *J. Chromatogr. A*, **1109**, 279–284.

Wang, S., Qu, H. and Cheng, Y. (2007) *Electrophoresis*, **28**, 1399–406.

Wang, H., Lu, Y., Chen, J., *et al.* (2012) *J. Pharm. Biomed. Anal.*, **58**, 146–51.

Warren, C. R., Chen, Z.-L. and Adams, M. A. (2000) *Physiol. Plantarum*, **110**, 52–58.

Wätzig, H., Degenhardt, M. and Kunkel, A. (1998) *Electrophoresis*, **19**, 2695–2752.

Wei, F. and Feng, Y. Q. (2008) *Talanta*, **74**, 619–624.

Weinberger, R. (2000) *Practical Capillary Electrophoresis*, Academic Press, San Diego, CA, pp. 25–72.

Wu, M., Ra, W., Zhang, Z., *et al*. (2011) *Electrophoresis*, **32**, 105–115.

Xie, C., Hu, J., Xiao, H., *et al*. (2005) *J. Sep. Sci.*, **28**, 751–6.

Xu, X., Qi, X., Wang, W., *et al*. (2005) *J. Sep. Sci.*, **28**, 647–652.

Xu, X., Ye, H., Wang, W., *et al*. (2006) *Talanta*, **68**, 759–764.

Yanfang, Z., Xingping, L., Zongde, Z., *et al*. (2006) *J. Pharm. Biomed. Anal.*, **40**, 157–161.

Yang, H.-P., Yue, M.-E. and Shi, Y.-P. (2006) *Chromatographia*, **63**, 449–452.

Yang, P., Li, Y., Liu, X., *et al*. (2007a) *J. Pharm. Biomed. Anal.*, **43**, 1331–1334.

Yang, Y., Sk, Y., Zy, L., *et al*. (2007b) *J. Liq. Chromatogr. Related Technol.*, **30**, 2669–2679.

Yang, F. Q., Zhao, J. and Li, S. P. (2010) *Electrophoresis*, **31**, 260–77.

Yang, G., Pan, Y., Liu, L., *et al*. (2011) *Chromatographia*, **73**, 1223–1228.

Yu, L., Xu, Y., Feng, H., *et al*. (2005) *Electrophoresis*, **26**, 3397–404.

Yuan, B., Zheng, C., Teng, H., *et al*. (2010) *J. Chromatogr. A*, **1217**, 171–4.

Yue, M. E., Jiang, T. F. and Shi, Y. P. (2004) *Talanta*, **62**, 695–699.

Yue, M. E., Li, Y. and Shi, Y. P. (2007) *Biomed. Chromatogr.*, **21**, 376–381.

Zaitseva, E. A. (2001) Deutsche an der Moskauer Universität im 19. Jahrhundert, in *Deutsch-russische Beziehungen in Medizin und Naturwissenschaften*, eds. D. Engelhardt and I. Kästner, Aachen, Shaker Verlag, pp. 209–226.

Zemann, A. J., Schnell, E., Volgger, D., *et al*. (1998) *Anal. Chem.*, **70**, 563–567.

Zhai, Z., Luo, X. and Shi, Y. (2006) *Anal. Chim. Acta*, **561**, 119–125.

Zhang, H. X. and Liu, M. C. (2004) *J. Chromatogr. B*, **812**, 175–81.

Zhang, X. and Zhang, Z. (2011) *J. Chromatogr. B*, **879**, 1641–1646.

Zhang, J., Wang, S., Chen, X., *et al*. (2003) *Anal. Bioanal. Chem.*, **376**, 210–3.

Zhang, Z., Yan, B., Liu, K., *et al*. (2009) *Electrophoresis*, **30**, 379–87.

Zhao, D. B., Liu, X. H., Cui, S. Y., *et al*. (2005a) *Chromatographia*, **61**, 643–646.

Zhao, Y., Ming, Y., Zhang, H., *et al*. (2005b) *Chromatographia*, **62**, 611–615.

Zhou, X., Zheng, C., Sun, J., *et al*. (2006) *J. Chromatogr. A*, **1109**, 152–159.

Zhou, X., Zheng, C., Huang, J., You, T. (2007) Identification of herb *Acanthopanax senticosus* (Rupr. et Maxim.) Harms by capillary electrophoresis with electrochemical detection. *Anal. Sci.* **23**, 705–711.

Zhou, S., Tan, J., Chen, Q., *et al*. (2010) *J. Chromatogr. A*, **1217**, 8346–51.

Zhu, J., Yu, K., Chen, X., *et al*. (2007) *J. Chromatogr. A*, **1166**, 191–200.

Zhu, Y., Chen, Q., Zeng, A., *et al*. (2010) *Chromatographia*, **71**, 447–454.

LC and LC–MS: Techniques and Applications

Nadja Arens, Stefanie Doell and Hans-Peter Mock

Department of Physiology and Cell Biology, Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany

1   INTRODUCTION

Mass spectrometry coupled with LC (liquid chromatography) separation has developed into a technique routinely applied for targeted as well as for nontargeted phytochemical analysis of plant samples. Earlier on, LC–MS (liquid chromatography–mass spectrometry) was mostly part of the efforts for identification of one or few unknown metabolites of interest as part of a phytochemical study. As a major strategy, unknown compounds had to be purified in sufficient quantities. The purified fractions were then subjected to LC–MS/MS as part of the structural elucidation, mostly complemented by NMR (nuclear magnetic resonance) analysis. With the advancement of mass spectrometrical instrumentation, LC–MS is now widely applied for analysis of crude plant extracts and large numbers of samples. By that, it has become an essential part of metabolomic studies (See **Metabolomics**), aiming at the comprehensive coverage of the metabolite profiles of cells, tissues, or organs. Owing to the huge chemical diversity of small molecules, conditions for the extraction will restrict the subfraction of the metabolome which can be actually analyzed. The conditions for LC have to be adjusted to allow good separation of the particular metabolites from any extract. Major consideration will be the selection of an appropriate column and suitable eluents, the establishment of gradient profiles, temperature conditions, and so on. The reader is referred to (See **Extraction Methodologies: General Introduction**, **Supercritical Fluid Extraction**, **New Trends in Extraction of Natural Products: Microwave-Assisted Extraction and Pressurized Liquid Extraction**, and **Solid-Phase Microextraction (SPME) and Its Application to Natural Products**) for sample preparation and extraction and to **HPLC and Ultra HPLC: Basic Concepts** for more information on HPLC (high performance liquid chromatography) and ultra HPLC.

The reader is also referred to other recent reviews covering the application of LC–MS for plant metabolite analysis (Allwood and Goodacre, 2010; Dunn, 2008; Marston, 2007). LC–MS analysis is also the central approach for plant proteome analysis and has replaced to a large extent the former gel-based methods (Matros *et al.*, 2011).

2   LC AND LC–MS INSTRUMENTATION

LC coupled with mass spectrometry is most commonly performed by (ultra) HPLC instruments (see **HPLC and Ultra HPLC: Basic Concepts**). Alternatively, capillary electrophoresis is applied (**Analysis of Natural Products by Capillary Electrophoresis and Related Techniques**). A variety of detectors can be used to monitor the elution of compounds.

Most commonly, absorbance detection is practiced, either in a single wavelength mode recording or by using a diode array system and extracting the chromatogram(s) at appropriate wavelength(s) as part of the data evaluation after finishing sample runs. Other principles to monitor elution of separated compounds are based on fluorescence, electrochemical detection, evaporative light scattering, radioactivity, or refractivity (Meyer, 2010). Compared with these detection methods, mass spectrometry provides the most universal applicability. All detectors based on detections other than mass spectrometry are covered in (See **HPLC and Ultra HPLC: Basic Concepts**) in detail. This section will focus on mass spectrometrical detection in LC–MS, and the relevant instrumental configurations. Technical aspects of the combination of MS with other detection systems will however be briefly considered.

2.1 Type of MS Instruments for LC–MS

Any MS instrument consists of three components, namely, a source to create ions, a mass analyzer, and a detection system (Figure 1).

Ion sources used in LC–MS include electrospray ionization (ESI), atmospheric pressure chemical ionization (APCI), and atmospheric pressure photoionization (APPI).

In an ESI source, ion formation takes place in a capillary held at a high voltage. Nebulization of the liquid and desolvatation are required to enable transfer of the ions into the gas phase. The development of the electrospray technology was awarded with the noble prize in 2002 to J. B. Fenn and the reader is referred to his noble lecture (Fenn, 2003). Fragmentation can already occur in the source, depending on the chemical structure of the analytes. For example, loss of side chains of flavonoids or the glucose moiety of the coumarin scopolin could easily be observed. Important parameters to control are the cone voltage
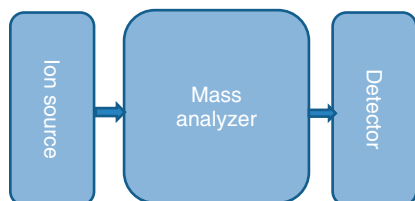
Figure 1   Components of MS instruments

and the dry gas temperature. In-source fragmentation can be used to generate information on the chemical composition of individual metabolites in case the instrument has no MS/MS capability. Formation of positive and negative ions occurs in ESI. In addition to the $[M + H]^+$ and $[M - H]^-$ ions, alkali adducts are frequently observed. Sample preparation and the composition of the solvent will impact the degree of various adducts formed.

ESI is also characterized more as a method of ion transfer, as it mostly relies on analyte ions already present in the liquid (Gross, 2011).

On the contrary, ions are actively generated in an APCI source. The liquid from the LC is transferred through a pneumatic nebulizer into a heater cartridge operated at a temperature of about 500°C for vaporization. Ion formation is supported by a corona discharge; for details of the processes leading to ion formation, see Gross (2011).

APCI has been used for analysis of apolar metabolites such as triacylglyerols (Rezanka and Sigler, 2007) and carotenoids (Rezanka *et al*., 2009).

In APPI, generation of ions is initiated using UV (ultraviolet) light, requiring the use of a quartz tube in the source design. APPI has been compared with APCI as a source for the analysis of pentacyclic triterpenoids extracted from the bark of plane birch. It was found that APPI provided higher sensitivity than APCI in positive ion mode, whereas opposite with APCI, greatest sensitivity for acidic triterpenes was observed in the negative ion mode (Rhourri-Frih *et al*., 2009).

Imbert *et al*. (2012) compared ESI, APCI, and APPI for a wide range of lipids. They used standard mixtures and biological extracts prepared from *Leishmania donovani* promastigotes resistant to amphotericin B (AmB). APPI provided the highest signal, signal-to-noise (S/N), and sensitivity for nonpolar and low-polarity lipids. ESI and APCI provide higher sensitivity for the most of the polar lipids.

Recently, a fast targeted method for wax ester for semiquantitative profiling using nano-ESI with a triple quadrupole (QqQ) detector after direct infusion was presented (Iven *et al*., 2013). The method was applied to characterize jojoba seed oil and Arabidopsis lines modified by expression of wax ester biosynthetic genes. Earlier results based on APCI-MS and GC–MS (gas chromatography–mass spectrometry) could be confirmed, and it is expected that the novel method will be instrumental in screening larger sets of plants (Iven *et al*., 2013).

When comparing the three modes, ESI-MS is by far the mostly used technique. However, for particular applications, APPI or APCI sources might be advantageous as outlined in the selected references. Current instruments allow rapid switching between the different sources, so that different applications can be conveniently scheduled.

Further development of existing and the introduction of novel types of *mass analyzers* have fundamentally contributed to the applicability of LC–MS for metabolite analysis.

Time-of-flight (TOF) analyzers are based on the fact that ions with different *m/z* are separated in a field-free flight tube. Ions having lower ratios will have higher velocity.

Linear TOF as well as reflector TOF analyzers were designed. The reflector mode provides an improvement in the resolving power [for technical details, see (Gross, 2011) and references therein]. A modified design is present in the TOF analyzers with orthogonal acceleration of ions (oaTOF, orthogonal acceleration time-of-flight). From a continuous flow of ions, pulses are orthogonally extracted into the TOF module. The oaTOF design bears several analytical advantages and has therefore been successfully used in novel instrument generations afterward (Gross, 2011).

A linear quadrupole (Q) mass analyzer consists of four rods, which are of cylindrical or hyperbolic shape. The opposite rods are pairwise connected electrically, and voltage is applied having DC and a RF component. The selected potentials enable a trajectory to some of the ions so that they can move through the quadrupole, whereas ions with higher or lower *m/z* values are lost by collision with the rods. The Q mass analyzer operates as a mass filter, and with the variation in the potential settings, a mass scan can be performed (Gross, 2011). When the constant voltage of the Q mass analyzer is set to zero, the quadrupole acts as a wide band pass filter for ions and *q* is now used as abbreviation.

The design of a quadrupole can be changed in a way so that it can work as a linear ion trap (LIT). A trapping potential can be achieved with electrodes at both ends of the quadrupole. A LIT is frequently used as an accumulating device for other mass analyzers such as the already mentioned oaTOF.

In a Fourier transform ion cyclotron resonance (FT-ICR) type of mass analyzer, ions are guided by the Lorentz force to circulate perpendicular to a strong magnetic field in a cell with ultrahigh vacuum. The cyclotron frequency is proportional to the ion charge and the magnetic field, and is inversely proportional to the ion mass. In practice, the real motion is rather complex, as other effects impose on the pure motion derived from the action of the Lorentz force; see Gross (2011) for a detailed outline. As frequencies can be measured very precisely and along with the introduction of very strong magnets, FT-ICR instruments provide high resolution and mass accuracy. FT-ICR-MS instruments are frequently operated with ESI sources. The ion current provided is additionally controlled by further components between the ion source and the ICR cell, such as a linear ion trap (Wang *et al.*, 2000).

The *orbitrap* mass analyzer represents a novel principle of an ion trap, where the ions are kept moving in an electrostatic field. The field is created from a central electrode with a spindle-like form and an outer electrode split in two halves. An image current is recorded between these two halves and the *m/z* values are calculated from the oscillation frequencies using Fourier transformation (Gross, 2011; Scigelova and Makarov, 2006). A linear quadrupole ion trap was added to the orbitrap mass analyzer to reach a higher application range of the instrument. When compared with FT-ICR instruments, the resolving power of the orbitrap analyzer is lower in the lower *m/z* range and better for higher *m/z* values, with the actual range settings being dependent on the given instruments being compared.

2.2 Hybrid Instruments

It has already been mentioned that many of the available instruments have several mass analyzer components such as in the FT-ICR-MS and in the orbitrap MS.

Another prominent class of hybrid instruments frequently used in metabolite analysis are the so-called Q-TOF (quadrupole time-of-flight) mass spectrometers. Between the quadrupole and the TOF mass analyzer, a higher pressure collision cell is installed. Hence, tandem mass spectrometry can be performed by collision-induced fragmentation of ions selected in the quadrupole.

Tandem MS is also performed with QqQ systems, in which instead of the TOF, a quadruple mass analyzer is used to measure fragment ions from the collision cell. The type of MS/MS experiments

include product ion scanning, neutral loss scanning, and single reaction monitoring (SRM) as well as multiple reaction monitoring (MRM) modes. SRM and MRM operation provide a very selective and sensitive way for targeted analysis of molecules. For the analysis of low abundant metabolites such as many plant hormones, this type of instrument is ideally suited. However, the MRM mode in the triple quade systems has been limited to a selected range of analytes. In contrast, Q-TOF type of tandem MS is applied for a broader profiling of the metabolic spectrum in a nontargeted way.

More recent additions to the hybrid instruments were the *QqLIT* and *QqTOF* type of mass analyzers providing enhanced sensitivity, rapid acquistion, and resolution when compared with QqQ instruments. The technical developments of the recent instruments allowed the introduction of the so-called sequential window acquistion of all theoretical fragment ion spectra (SWATH) strategy, where the acquisition of the precursion ions is performed in small mass windows of about 20 mass units in a sequential manner.

2.3 Coupling of MS Detection with Other Detection Methods

As discussed in the already mentioned chapters, a range of other detection systems are in use to monitor elution of compounds after LC separation. Most frequently, absorbance has been measured, in a single or dual wavelength mode detector, or in the now common diode array type of instruments. Other detectors, for example, fluorescence detectors, have been used alternatively or in combination with absorbance detectors. Selection of a specific detection system will be guided by the class of compounds under investigation. If the user wishes to keep this information present when switching to an LC–MS approach, an additional detector can be integrated into the setup. When doing so, the specific requirements imposed by the MS detector have to be considered, such as the backpressure exerted on the flow cell of the first detector. In practice, additional use of a UV/VIS detector is quite common.

2.4 The LC Part of the LC–MS System

Although most of the developments in LC–MS-based metabolite analysis were achieved in mass

spectrometrical instrumentation, the LC part has also experienced technical advancements. HPLC systems were designed with higher pressure in the range of 15,000 psi, exceeding the 6000 psi being a common limit reached in the earlier HPLC generations. These types of HPLC systems are commonly named as UPLC (ultra performance liquid chromatography) and instruments are now available from different companies. The higher pressure range allows the use of columns with smaller particles (<2 μm), leading to improved resolution and shorter separation cycles. The shorter separation time per sample enables a higher throughput of samples for large-scale experiments. This could be screening of mutant collections or genotypes for metabolic traits, or the comprehensive analysis of larger sample sets in "omics" studies. Optimized separation of compounds before MS analysis reduces the suppression of signals by co-eluting compounds and eases the distinction of related compounds with similar MS or MS/MS properties. An example will be discussed later on (Vrhovsek *et al*., 2012).

3 APPLICATIONS

The metabolome of plant organs, tissues, or even cell populations is characterized by a huge chemical diversity, covering a wide range in terms of chemical properties from very hydrophilic to rather lipophilic molecules and different molecular weight. Sugars, polymeric carbohydrates, amino acids, and other organic acids as well as lipids are among the so-called primary metabolites, complemented by a vast array of metabolites from specialized or secondary metabolism. The abundance of primary metabolites as well as secondary metabolites is dependent on many control mechanisms, to allow plant adaptation to environmental conditions and developmental programs. The metabolome is therefore highly dynamic. Changes in concentrations of individual metabolites might, for example, occur in seconds or might follow diurnal patterns. Hence, experiment setup and sampling strategies have to be carefully considered to ensure reproducible comprehensive metabolic analysis. The complexity of the metabolome however cannot be covered by any single analytical method, despite the capabilities of recent MS technology. Whole metabolome analysis is therefore dependent on several analytical
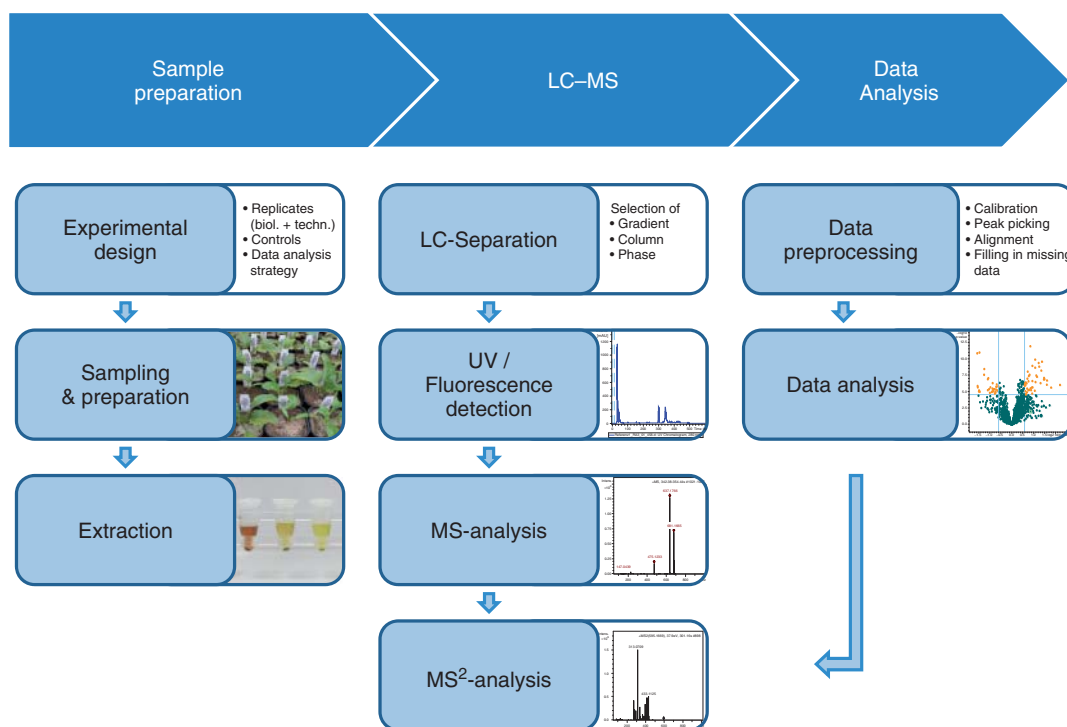
Figure 2 Schematic overview of the LC-MS workflow. It is divided in 3 major segments: Sample preparation, LC-MS analysis and data analysis. Sample preparation comprises the experimental design, which is the first and most important step. Further sample preparation and extraction is conducted. The second segment involves chromatographic separation and mass spectrometric analysis. Optionally UV & fluorescence detection and MS^2 analysis can be implemented. The third segment data analysis consists of data preprocessing (mainly MS-data) and the actual analysis with the help of statistical tools.

approaches and is frequently performed in a combinatorial manner. Nontargeted approaches allow pinpointing novel candidates in the individual biological context; a typical workflow is depicted in Figure 2. Although nontargeted approaches provide a comprehensive overview of the metabolic pattern of the extracted tissue, targeted strategies have their benefits. Individual protocols defined by specific extraction and analytical procedures in most cases will be superior in quantification.

3.1 Targeted Metabolite Analysis

A targeted method was established for the quantification of phenolics occurring in fruits and beverages (Vrhovsek *et al.*, 2012). The method was based on UPLC separation with a triple quad MS detection system and allowed the quantification of 135 phenolic

substances, representing different branches of this class of natural compounds, such as flavonoids, coumarins, stilbens, and simple cinnamic acid conjugates among others. With the optimized LC separation method, one analytical cycle lasted 15 min. Separation was provided for most of the isomeric compounds included in the study, such as for the chlorogenic acid isomers allowing their separate quantification. In some cases, the methods only allowed the quantification of the isomers together. This observation emphasizes the value of good LC separation before MS analysis, as already outlined in an earlier part of the manuscript.

Almost 1000 commercially available compounds were used to develop a targeted metabolomics, also using MRM on a triple quad MS system after HPLC separation. The conditions could be successfully established for 497 compounds of the library, and the established procedure was applied to quantify

patterns of 100 metabolites in 14 plant accessions from three plant species. Data were evaluated with a batch-learning, self-organizing map analysis and allowed prediction of family-specific metabolites (Sawada *et al*., 2009). The data sets and technical features of the MRM setup were made available through the PRIMe website (Akiyama *et al*., 2008) of the RIKEN institute.

3.2 Nontargeted Analysis

LC–MS has been applied for nontargeted profiling of compounds with medium polarity, typically represented by plant phenylpropanoids. Analysis of Arabidopsis roots and leaves were performed by LC–MS using a capillary HPLC system followed by a Q-TOF type of instrument (Roepenack-Lahaye *et al*., 2004). In addition to compounds from the phenylpropanoid pathway, glucosinolates and their degradation products could be tentatively annotated to the mass signals. Overall, approximately 2000 mass signals were recorded in this study. The analytical setup was tested by comparing the metabolic profiles of wild-type plants with mutants lacking chalcone synthase. As expected, the profile of the mutant plants lacked flavonoid glycosides present in the controls. Interestingly, the root extracts showed higher accumulation of compounds assigned to the glucosinolate pathway (Roepenack-Lahaye *et al*., 2004).

The metabolome of tomato fruits was characterized by LC–MS and data were deposited in a database (Moco *et al*., 2006). The detection system used was composed of a PDA (photodiode array) detector followed by a Q-TOF type of MS instrument. With the extraction and separation parameters used, the metabolite profile was dominated by a vast array of phenolic compounds, but alkaloids were also found. The LC system was characterized by a small variation in retention time for all peaks of about 2 s when the chromatograms of a single series of samples were compared. The chromatographic separation allowed resolving many isomeric forms. In addition to a range of already known components, the LC-PDA–MS/MS approach revealed a number of novel compounds that were tentatively assigned. It was emphasized that NMR will be required to establish the full structures of the novel compounds (Moco *et al*., 2006). Data evaluation was assisted by the use of a novel software tool termed *MetAlign*

(see Lommen, 2012 for a recent reference), which will be further described in the section on Bioinformatic tools. Data and information from literature were stored in a database called *MoTo* accessible via http://appliedbioinformatics.wur.nl. A detailed protocol for the whole work flow from extraction, instrument setup, and data analysis has been provided (De Vos *et al*., 2007).

With the use of different LC column materials, the LC–MS-based profiling can also be extended to monitor more hydrophilic compounds. A pentafluorophenylpropyl column was found superior to a reversed-phase C18 column when analyzing the polar cellular metabolites of the blue-green algae Synecchocystis (Narainsamy *et al*., 2013). The chromatographic system was coupled to a LTQ-orbitrap MS system. The method was established using more than 100 reference compounds. Global reprogramming of metabolism stimulated by glucose could be quantitatively assessed when testing the influence of various growth conditions (Narainsamy *et al*., 2013). A combination of reversed-phase and hydrophilic interaction liquid chromatography (HILIC) for separation of metabolites and detection in positive as well as negative MS mode was also introduced by the group of G. Siuzdak. The approach increased the spectrum of metabolites detected in extracts from bacterial cell, human plasma, or cancer cells (Ivanisevic *et al*., 2013).

3.3 Combination of Targeted with Nontargeted LC–MS Analysis

A combination of nontargeted analysis of metabolites with a FT-ICR-MS type of instrument with several approaches for targeted analysis was performed on poplar genotypes with modified emission of isoprene emission (Way *et al*., 2013). Data sets resulting from the nontargeted analysis were evaluated using the MassTrix tool (Suhre and Schmitt-Kopplin, 2008). The metabolite profiles of the plants with suppressed isoprene emission capacity showed alterations in several classes of compounds, when compared with wild-type controls. Interestingly, the metabolic differences between poplar plants emitting isoprenes and those with suppressed emission were diminished at higher CO_2 concentrations. This outcome of a metabolomics study will be relevant when discussing future climate warming (Way *et al*., 2013).

3.4 Mass-Spectrometry-Based Imaging

Within recent years, methods for spatial resolved analysis of metabolites and proteins using mass spectrometry have been introduced in plant biology (Kaspar *et al*., 2011; Peukert *et al*., 2012). For example, MS imaging has been applied to resolve the patterns of metabolites in seed tissues by measuring the surface of seed sections. In many of the plant metabolite studies cited so far, the extraction of whole organs such as roots or leaves is performed before LC–MS-based separation. Accordingly, information on the cell-type specific accumulation of metabolites is lost. Future modeling of plant metabolic networks will benefit from the spatial information of metabolites and corresponding enzymatic activities. Hence, the information obtained by MS imaging will be highly complementary to the data obtained by LC–MS analysis of extracted samples. The reader is referred to the cited papers for further information.

4 DATA EVALUATION

LC–MS in MS and MS/MS mode provides large sets of data, which have to be evaluated to provide information in the context of targeted as well as nontargeted applications (Figure 3). The evaluation starts with the task to preprocess the raw data into a data format ready for further bioinformatic analysis. Statistical analysis with approaches such as principal component analysis (PCA) or independent component analysis (ICA) are common steps, followed by data visualization such as mapping into metabolic pathways. The software solutions offered by the instrument vendors differ in their capabilities with respect to the downstream part of the data analysis. Meanwhile, a large set of resources are accessible over the web (Tohge and Fernie, 2009) and a large number of tools for bioinformatic analysis of LC–MS data is available within the scientific community (Dunn, 2008; Sugimoto *et al*., 2012).

A range of vendor-independent software suites have been generated for data extraction from the raw data files. This step of data analysis is addressed as (pre)processing and consists of several specific tasks. Data conversion often has to be performed to transform data from any vendor-specific format into a format for independent software tools. Feature detection is a central step in the preprocessing pipeline and

provides its own challenges owing to the complex peak shapes of the MS data. The comparison of data from several samples requires an alignment, another challenging process for bioinformatic software tools. Shifts in the elution time occur despite the progress in instrument performance and efforts to control the system parameters affecting stability such as column and eluent temperature.

Different mathematical approaches have been introduced [see Sugimoto *et al*. (2012) for further discussion].

The software tool MetAlign (Lommen, 2009) has been developed initially to evaluate LC–MS data (De Vos *et al*., 2007; Matsuda *et al*., 2009; Moco *et al*., 2006), but can also be used to analyze GC–MS data (Lommen, 2012). Meanwhile, this software is freely available.

XCMS is another open software tool for data extraction (Smith *et al*., 2006), which is widely used (Narainsamy *et al*., 2013) and is now available as a web-based tool (Tautenhahn *et al*., 2012a).

Databases storing information from plant metabolite analysis have become an important resource for sharing information in the scientific community (Fukushima and Kusano, 2013). Several databases were established to store spectra from LC–MS analysis.

METLIN (http://metlin.scripps.edu/index.php) is a database storing MS/MS data from model compounds generated at different collision energies (Tautenhahn *et al*., 2012b). Currently, the data from more than 10,000 compounds have been collected. In the recent version of the database, an advanced workflow has been established allowing to upload own data and perform comparisons with the stored reference spectra in an automated manner. The capability to support compound identification has been verified by analyzing the MS data from 23 standard compounds on five different MS instruments; correct results were obtained for nearly 90% of the data sets (Tautenhahn *et al*., 2012b). However, it has to be noted that many plant metabolites are still not available as commercial standards, in particular, those of the different classes of specialized metabolism. This lack of available references for a majority of plant secondary metabolites is limiting this database strategy.

Mass Bank (http://www.massbank.jp/) was initiated to create a public repository of mass spectral data provided by a number of research groups (Horai *et al*., 2010).
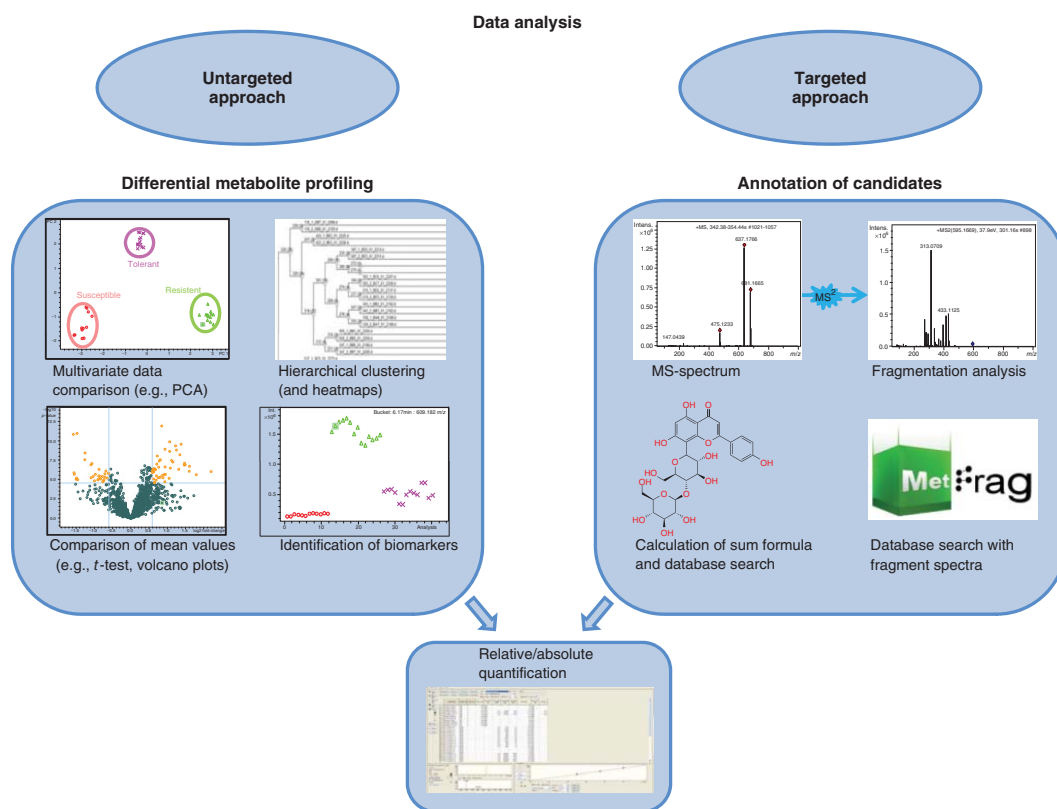
Figure 3 Overview of the most common data analysis strategies. With an untargeted approach differential metabolite profiling can be implemented. Multivariate statistic is used to compare large sets of data and identify potential biomarkers. These candidates are tentatively annotated in a targeted approach by fragmentation analysis and database searches.

AtMetExpress is a database storing data from the analysis of Arabidopsis, which is harbored at http://prime.psc.riken.jp/lcms/AtMetExpress/. Other databases store information on chemical compounds such as PubChem, ChemSpider, and chemical abstract service (CAS). There are more databases than those few mentioned here, providing a wealth of information to assist the researcher in identification of compounds [see Fukushima and Kusano (2013) for a recent review of metabolome databases]. A novel tool called *MetFusion* was created to combine identification results from different identification processes (Gerlich and Neumann, 2013).

Further data analysis includes a range of statistical tools and visualization of data (Fukushima and Kusano, 2013; Sugimoto *et al.*, 2012; Theodoridis *et al.*, 2012). At this point, collaboration between the application chemist and a scientist representing specific bioinformatic expertise is a strategy for effective data mining. A close interaction between experimentalists (wet lab) and bioinformatic scientists (dry lab) is also effective when integrating data from different "omics" approaches. A combination of metabolomics and transcriptomics has been used to further annotate the phenylpropanoid pathway in Arabidopsis (Saito and Matsuda, 2010).

REFERENCES

Akiyama, K., Chikayama, E., Yuasa, H., *et al.* (2008) *Silico Biol.*, **8**(3–4), 339–345.

Allwood, J. W. and Goodacre, R. (2010) *Phytochem. Anal.*, **21**(1), 33–47.

De Vos, R. C. H., Moco, S., Lommen, A., *et al.* (2007) *Nat. Protoc.*, **2**(4), 778–791.

Dunn, W. B. (2008) *Phys. Biol.*, **5**(1), 11001.

Fenn, J. B. (2003) *Angew. Chem. Int. Ed.*, **42**(33), 3871–3894.

Fukushima, A. and Kusano, M. (2013) *Front. Plant Sci.*, **4**.

Gerlich, M., and Neumann, S. (2013) *J. Mass Spectrom.*, **48**(3), 291–298.

Gross, J. H. (2011) *Mass Spectrometry: A Textbook*, Springer, Heidelberg.

Horai, H., Arita, M., Kanaya, S., *et al.* (2010) *J. Mass Spectrom.*, **45**(7), 703–714.

Imbert, L., Gaudin, M., Libong, D., *et al.* (2012) *J. Chromatogr. A*, **1242**, 75–83.

Ivanisevic, J., Zhu, Z. J., Plate, L., *et al.* (2013) *Anal. Chem.*, **85**(14), 6876–6884.

Iven, T., Herrfurth, C., Hornung, E., *et al.* (2013) *Plant Methods*, **9**(1), 24.

Kaspar, S., Peukert, M., Svatos, A., *et al.* (2011) *Proteomics*, **11**(9), 1840–1850.

Lommen, A. (2009) *Anal. Chem.*, **81**(8), 3079–3086.

Lommen, A. (2012) *Methods Mol Biol.*, **860**, 229–253.

Marston, A. (2007) *Phytochemistry*, **68**(22–24), 2786–2798.

Matros, A., Kaspar, S., Witzel, K., *et al.* (2011) *Phytochemistry*, **72**(10), 963–974.

Matsuda, F., Yonekura-Sakakibara, K., Niida, R., *et al.* (2009) *Plant J.*, **57**(3), 555–577.

Meyer, V. R. (2010) *Practical High-Performance Liquid Chromatography*, John Wiley & Sons, Chichester, UK.

Moco, S., Bino, R. J., Vorst, O., *et al.* (2006) *Plant Physiol.*, **141**(4), 1205–1218.

Narainsamy, K., Cassier-Chauvat, C., Junot, C., *et al.* (2013) *Metabolomics*, **9**(1), 21–32.

Peukert, M., Matros, A., Lattanzio, G., *et al.* (2012) *New Phytolog.*, **193**(3), 806–815.

Rezanka, T., and Sigler, K. (2007) *Curr. Anal. Chem.*, **3**(4), 252–271.

Rezanka, T., Olsovska, J., Sobotka, M., *et al.* (2009) *Curr. Anal. Chem.*, **5**(1), 1–25.

Rhourri-Frih, B., Chaimbault, P., Claude, B., *et al.* (2009) *J. Mass Spectrom.*, **44**(1), 71–80.

Roepenack-Lahaye, E., Degenkolb, T., Zerjeski, M., *et al.* (2004) *Plant Physiol. (Rockville)*, **134**(2), 548–559.

Saito, K., and Matsuda, F. (2010) *Annu. Rev. Plant Biol.*, **61**(1), 463–489.

Sawada, Y., Akiyama, K., Sakata, A., *et al.* (2009) *Plant Cell Physiol.*, **50**(1), 37–47.

Scigelova, M., and Makarov, A. (2006) *Proteomics*, **6**(S2), 16–21.

Smith, C. A., Want, E. J., O'Maille, G., *et al.* (2006) *Anal. Chem.*, **78**(3), 779–787.

Sugimoto, M., Kawakami, M., Robert, M., *et al.* (2012) *Curr. Bioinform.*, **7**(1), 96–108.

Suhre, K., and Schmitt-Kopplin, P. (2008) *Nucleic Acids Res.*, **36**(suppl 2), W481–W484.

Tautenhahn, R., Patti, G. J., Rinehart, D., *et al.* (2012a) *Anal. Chem.*, **84**(11), 5035–5039.

Tautenhahn, R., Cho, K., Uritboonthai, W., *et al.* (2012b) *Nat. Biotechnol.*, **30**(9), 826–828.

Theodoridis, G. A., Gika, H. G., Want, E. J., *et al.* (2012) *Anal. Chim. Acta*, **711**, 7–16.

Tohge, T., and Fernie, A. R. (2009) *Phytochemistry*, **70**(4), 450–456.

Vrhovsek, U., Masuero, D., Gasperotti, M., *et al.* (2012) *J. Agric. Food Chem.*, **60**(36), 8831–8840.

Wang, Y., Shi, S. D. H., Hendrickson, C. L., *et al.* (2000) *Int. J. Mass Spectrom.*, **198**(1–2), 113–120.

Way, D. A., Ghirardo, A., Kanawati, B., *et al.* (2013) *New Phytolog.*, **200**(2), 534–546.

NMR as Analytical Tool for Crude Plant Extracts

Anna R. Bilia

Department of Chemistry, University of Florence, Sesto Fiorentino (Firenze), Italy

1 INTRODUCTION

Plant extracts are extremely complex matrices containing different chemical classes of constituents having diverse polarity and solubility and present at various concentrations. Frequently, the dominant ones are not involved in the biological and pharmacological activity, whereas markers or active principles generally represent only minor constituents. Therefore, the choice of an appropriate analytical method is firmly correlated to the identification and quantification of multiple analytes in the extracts, a sort of multitarget botanical standardization for profiling the totality of the chemical constituents being recognized that botanicals exert their effects as a whole.

In the Pharmacopoeias, standardized analytical methods for the quality control and stability testing of extracts, based on the analysis of markers (single constituents or groups) or active constituents (single constituents or groups), are generally achieved using UV spectroscopy or chromatographic techniques, mainly high performance thin layer chromatography (HPTLC), high performance liquid chromatography (HPLC) coupled with UV–vis detector, and capillary gas liquid chromatography (GLC).

In recent two decades, chromatographic and spectroscopic methods of analysis have becoming more and more sophisticated, mainly based on hyphenated chromatographic and spectroscopic techniques to obtain constituent information online, directly from crude plant extracts. HPLC or ultra-performance liquid chromatography (UPLC) coupled with different mass spectrometers (MS) and tandem (MS–MS, MS^n) mass spectrometers such as chemical ionization (CI), electron impact (EI), electrospray ionization (ESI), fast atom bombardment (FAB), field desorption/field ionization (FD/FI), matrix-assisted laser desorption ionization (MALDI), and thermospray ionization (TSP). Alternatively, LC is combined with evaporative light scattering detector (ELSD) and nuclear magnetic resonance (NMR). Even electrophoretic techniques such as capillary electrophoresis (CE) coupled with UV and MS have been extensively used in the analysis of plant extracts (Marston, 2007).

All of these analytical techniques are recognized as very powerful methods; however, frequently, it is not easy to develop a simple and rapid chromatographic separation of the constituents because of their varying solubility, polarity, and size. In addition, these techniques cannot reveal unknown plant metabolites that can also be present and contribute to the biological activity, rising problems concerning the specificity and sensitivity of the method of revelation. Furthermore, the baseline instability and the separation are generally time consuming and solvent consuming. In most of the cases, crude extracts need a prepurification or derivatization steps and thus face the possibility of losing some constituents during the analysis. Generally, these analytical methods

provide a fingerprint of active constituents or arbitrarily chosen "marker" compounds (and their percentage) but no information with regard to the other metabolites of the extract that can represent even more than 95% w/w can be acquired. In addition, some unknown substances (single or groups) can be involved in the pharmacokinetic or pharmacodynamic properties of the extracts (Bilia *et al.*, 2001). Finally, the above-mentioned methods require pure reference compounds, which are difficult to obtain commercially because of their limited availability or stability and high price.

NMRs, in particular ^{1}H and ^{13}C NMR spectra, are not frequently used for the direct analysis of plant extracts but this tool has an outstanding position in this field because it is nondestructive and noninvasive, very rapid, selective, modest time consuming, and capable of simultaneous detection of a great number of constituents in complex mixtures, both major and minor components (Fan, 1996; Lindon and Nicholson, 1997).

In addition, quantitative NMR procedures for the determination of constituents are rarely reported but could be truly useful because selectivity, reproducibility, and sensitivity are comparable with HPLC, besides the great advantage that no reference substances (identical authentic pure phytochemical reference) are needed to establish a standard curve for quantification (Pauli, Jaki, and Lankin, 2005). Furthermore, NMR spectroscopy could select appropriate signals in fast and simple one-dimensional NMR that can serve as parameters in the chemometric analysis.

Within the last years, buzz words such as metabolic profiling or metabolomics have appeared in the literature. Key concepts in metabolome analysis involve the qualitative and quantitative evaluations of the spectra recorded by NMR using chemometric methods as fingerprinting technologies (Lindon and Nicholson, 2008).

1.1 Constituent Detection by NMR

Six decades after its discovery, NMR can be considered a leading analytical tool extensively employed for qualitative structure elucidation of natural products and also for quantitative determinations of purified target analytes.

Any molecule containing one or more atoms with a nonzeromagnetic moment is potentially detectable by NMR, as the isotopes with nonzero magnetic moments include ^{1}H, ^{13}C, ^{14}N, ^{15}N, and ^{31}P, and all biologically important molecules have at least one NMR signal. These signals are characterized by their frequency (chemical shift, in δ or ppm), intensity, fine structure, and magnetic relaxation properties, all of which reflect the precise environment of the detected nucleus.

While early NMR structure analysis was predominantly based on the interpretation of chemical shift data because of the modulation of the resonance energy by the electronic environment of a given atomic nucleus resulting in chemical shift dispersion, coupling analysis has become progressively more important as a result of the explosive development of methods for the analysis of spin networks based on the mutual influence that neighboring nuclei exert on each other via chemical bonds (scalar coupling), namely spin–spin coupling and dipolar coupling (J, leading to connectivity), by interaction through space (nuclear Overhauser effect, NOE) and relaxation ($T1/T2$) (Fan, 1996).

In addition, in recent two decades, the technique of NMR spectroscopy has improved markedly in terms of accuracy, skills, and sensitivity. As the increasing power of the magnets boosted the super-conductivity boundary, the increase in the magnetic field allowed higher resolution spectra to be obtained with smaller amounts of compound. The development of new pulse sequences led to new two-dimensional (2D) techniques where second dimension increases spectral resolution by distributing the signals along two frequency axes.

These experiments exploit the interactions between the NMR detectable isotopes in a molecule, and they result either in homonuclear correlation, where the two-frequency axes of the spectrum correspond to the same nucleus, usually ^{1}H, or in heteronuclear correlation, where one-frequency axis corresponds to ^{1}H and the other corresponds to ^{13}C, ^{15}N, or, occasionally, ^{31}P. Homonuclear correlation experiments are particularly useful in metabolite profiling, allowing linked subsets of peaks in the conventional one-dimensional (1D) ^{1}H NMR spectrum to be identified and assigned to particular compounds.

In typical cases, 2D ^{1}H–^{1}H correlation experiments, recorded with different mixing times, are used to assign the proton spin systems via scalar couplings.

The homonuclear *co*rrelation *s*pectroscopy (COSY) experiment is probably the most widely used 2D experiment. The experiment reveals cross

peaks between directly coupled protons. The most widely used variant of this experiment is the basic COSY-90 and the COSY-45 sequences that are both routinely available in gradient-selected mode.

The relayed COSY sequences are called *total correlation spectroscopy* (*TOCSY*) and it can give the sequence that would give cross peaks between all of the protons compromising a sequence of coupled protons. There are different sequences, but the name TOCSY is commonly used for the experiment, regardless of the specific sequence used. It can be run in either absolute value or phase-sensitive mode, although there are resolution advantages to the latter approach. A gradient-selected version of the TOCSY sequence is also available in spectrometer software packages. The resolution and unequivocal structure elucidation of complex mixtures can also be achieved by 1D-selected TOCSY experiments. The acquisition of a series of 1D-TOCSY spectra with different mixing times takes far less time than a corresponding set of 2D-TOCSY spectra.

Heteronuclear correlation experiments are also successfully used for the analysis of plant extracts. Proton-detected ^1H–^{13}C correlation experiments with gradient selection of magnetization transfer, for example, heteronuclear multiple quantum coherence (HMQC) and heteronuclear single quantum coherence (HSQC), are used to correlate the ^{13}C chemical shifts with protons via one bond to find the correlations between protons and their neighboring carbons. Heteronuclear multiple bound coherence (HMBC) correlates the ^{13}C chemical shifts with protons via two or three bonds and to confirm the connections of quaternary carbons to adjacent protons through two- or three-bond couplings.

Both HSQC and HMQC involve detection of protons directly bonded to ^{13}C, and consequently, suppression of magnetization of the 99% of protons that are bonded to ^{12}C is a significant concern.

Two-dimensional NOESY and ROESY experiments provide information about dipolar couplings through space (NOE) or through the lattice (T1) linkage of structural fragments, respectively. The NOESY sequence is most commonly used even if there are two significant problems with this sequence, particularly when applied to natural products. First, the maximum observable NOE changes as molecular tumbling slows in solution from +38% for small molecules to −100% for macromolecules and the exact crossover point where NOE is near zero depends on molecular size and shape, internal mobility, solution viscosity, and the spectrometer frequency. The second problem is that the NOESY sequence is also prone to produce COSY cross peaks as artifacts.

The ROESY sequence replaces the mixing period with a spinlock giving positive NOE cross peaks over a wide range of tumbling rates from small molecules to the largest proteins. However, ROESY tends to produce TOCSY peaks as artifacts, but this can be avoided using a modified version called the *T-ROESY sequence*. One-dimensional NOE difference experiments have been available for many years before the development of NOESY and ROESY sequences and recently have been improved using shaped pulses and pulsed field gradients, which gives much cleaner spectra than NOE difference experiments.

In addition, pairs of sequences can be combined together to provide hybrid sequences giving information from both types of experiments. These can be carried out as 3D and even 4D experiments. The most generally used 3D experiments are HMQC-TOCSY and HSQC-TOCSY. This is particularly true for crowded ^1H spectra, where the superior ^{13}C resolution helps in interpreting the TOCSY data. These experiments require considerably longer times and generally suffer from limited resolution along the two time-incremented axes (Reynolds and Enríquez, 2002).

Modern NMR instruments are also able to measure self-diffusion constants with an accuracy approaching 1% for objects in solution ranging from the size of molecules to that of aggregates such as micelles and liposomes (Bilia *et al*., 2002c). These experiments are denominated *d*iffusion-*or*dered *s*pectroscop*y* (DOSY) and diffusion rates are measured using gradient pulses such as pulsed gradient spin echo (PGSE), pulsed field gradient-stimulated spin echo (PFG-SSE), and bipolar pulse longitudinal eddy current delay (BPP-LED). NMR diffusion experiments can be exploited as a routine analytical tool for the study of mixtures providing a way to separate the different compounds in a mixture based on the differing translation diffusion coefficients (and therefore differences in the size and shape of the molecule, as well as physical properties of the surrounding environment such as viscosity and temperature) of each chemical species in solution. In a certain way, it can be regarded as a special chromatographic method for physical component separation, but unlike those techniques, it does not require any particular sample preparation or

chromatographic method optimization and maintains the innate chemical environment of the sample during analysis. Diffusion spectra can be combined with any 2D technique in order to separate 2D spectra in the diffusion dimension, giving 3D-DOSY experiments such as DOSY-NOESY, DOSY-HMQC, or by incorporating diffusion weighting internally, where the parent pulse sequence can accommodate an extra diffusion delay as in the COSY-iDOSY and HMQC-iDOSY pulse sequences. Again, all these experiments require considerably long times with respect to the native DOSY (Barjat, Morris, and Swanson, 1998; McLachlan *et al.*, 2009).

2 PLANT EXTRACTS NMR ANALYSES

2.1 NMR Fingerprint Determination

NMR experiments can be considered as universally suitable method of assay of plant extracts and the existing high field strengths led to record highly resolved NMR signals. For this reason, in the past two decades, NMR has played an increasingly important role in the compositional study of food products, particularly wine, fruit pulps and juices, olive oil, beer, and beverages for the analysis primary metabolites such as sugars, organic acids, amino acids, and lipids (Cordella *et al.*, 2002).

The NMR analysis of secondary metabolites of crude plant extracts is more recent but now the literature offers remarkable examples of application of NMR analysis in their qualitative and quantitative analyses. NMR can reveal information about similarities and differences of the extracts by comparing their metabolic fingerprints and to conclude on their relationship to each other.

In general, signal (metabolite) identification is performed by comparison with reference compounds, either from a database or from the spectrum of an authentic sample, to confirm known compounds. Frequently, signal overlap is a problem in the complex spectra of plant extracts and hampers the identification and quantification of secondary constituents. A better signal resolution can be obtained using 2D-NMR spectroscopy reducing signal overlap by spreading the resonances in a second dimension but this approach retains some of the benefits of 1D-NMR spectroscopy, principally the longer acquisition time (Reynolds and Enríquez, 2002).

NMR instruments required for plant extract analysis have generally field strengths ranging from 300 to 600 MHz, obviously higher field strengths give better resolution of the overlapping resonances and improve the sensitivity. The sample is prepared by dissolution of an amount of the extract (10–50 mg) in an appropriate solvent, in specific proportions, usually in a 5-mm diameter NMR tube, using sample volumes ranging from 300 to 600 μL. A broad range of solvents suitable for most of the polarities encountered in natural product chemistry can be used, namely deuterated dimethyl sulfoxide (DMSO-d_6), deuterated methanol (CD$_3$OD), H$_2$O/D$_2$O, or D$_2$O buffered with sodium phosphate or mixtures thereof (e.g., DMSO-d_6/D$_2$O and CD$_3$OD/D$_2$O). The addition of chemical shift reference compounds (e.g., TSP, trimethylsilyl-1,1,2,2-d$_4$-propionic acid, or TMS, tetramethylsilane) to the NMR solvent is useful for both peak registration and quantitation purposes.

In selected cases, the process of NMR analysis can be tremendously improved by the optimization of temperature, pH value, or by adding some auxiliary reagents (shift reagents) (Choi *et al.*, 2003).

A variety of acquisition parameters have been described and optimized for the specific sample such as acquisition times and recovery delays. Spectral acquisition is normally carried out at room temperature or at a controlled temperature between 20°C and 30°C (Reynolds and Enríquez, 2002).

After the processing of the acquired data, many functional groups can be easily and conclusively identified by their characteristic ^{1}H and/or ^{13}C chemical shifts compared with those reported in the literature.

It is worth noting that the chemical shift range of proton is little and hardly extends beyond 10 ppm, with the significant disadvantage that the dispersion of the signals is rather small, resulting in extensive overlap in the signals in most regions of the spectrum, by comparison, the chemical shift range of ^{13}C is larger by a factor of about 20 (extending over about 200 ppm).

The first use of NMR for metabolite profiling is represented by the ^{13}C NMR analysis for the fingerprint of essential oils, in order to detect and identify the pattern of constituents and capable of measuring multiple components simultaneously, as an alternative or complimentary analytical method to hyphenated chromatographic techniques.

In the ^{13}C-NMR spectrum, the presence of ^{13}C–^{1}H couplings, which are needed to clarify exact structure, can be eliminated simplifying the spectrum and obtaining singlet signals of the C nuclei, usually well resolved. The natural occurrence of the ^{13}C nucleus is only 1.1%; consequently, homonuclear spin–spin coupling does not occur. At the same time, the signal to noise ratio and the sensitivity of the measurement are increased. The overlapping of individual resonance signals occurs rather seldom being the spectral width quite large in the range from 0 to 240 ppm (Formäcek and Kubeczka, 1982).

Generally, the use of ^{13}C NMR is quite limited, and proton NMR spectroscopy is commonly considered an effective technique for plant extract analysis with good chemical specificity for constituents. Increased specificity is further realized with the use of high magnetic fields that provide greater resolution and separation of chemical shifts. In the light of these considerations, it is unsurprising that ^{1}H NMR is the most commonly used NMR technique for metabolite profiling. The problems arising from the limited spectral dispersion can be circumvented or at least minimized using 2D experiments.

Studies on the combination of high resolution NMR spectroscopy with pattern recognition have been firstly reported by Bilia and coworkers. Searching for a universally suitable method of assay of plant extracts, 1D- and 2D-NMR experiments were considered as an analytical instrument to fully characterize them and evaluate the stability of their constituents. Fingerprint and semiquantitative analyses have been reported for extracts of *Piper methysticum* G. Forst (Bilia *et al*., 2002a), *Hypericum perforatum L.* (Bilia

et al., 2001a), and *Arnica montana L.* (Bilia *et al*., 2002b).

In a first study (Bilia *et al*., 2001a), a commercially dried extract of St John's wort was evaluated by NMR spectroscopy without fractionation steps. St John's wort is generally standardized in the hypericin content, although a number of constituents with documented biological activity are also present, namely other naphthodianthrones (principally pseudohypericin), flavonols (rutin, hyperoside, isoquercitrin, quercitrin, and quercetin), and phloroglucinols (hyperforin, adhyperforin, and others) (Figure 1).

As a first step of the investigation, a DMSO-d_6 solution of the extract was prepared and used to record resolution-enhanced proton NMR of the total extract (Figure 2). The spectrum had a sufficient line narrowing leading to the assignments of some characteristic resonances according to the data (chemical shifts and coupling constants) found in the literature or from the comparison of the chemical shifts and *J* couplings with those observed for single constituent solutions. Two-dimensional NMR spectra provided further information to confirm these assignments and complete the pattern recognition, in particular the HMQC spectra (Figure 3).

The proton NMR spectrum was separated in four main regions: a low field region between 6.0 and 9.0 ppm with signals principally due to aromatic and olefinic protons of flavonoids, naphthodianthrones, and caffeoyl quinic acid derivatives; a mid-low field region between 5.5 and 4.5 ppm with signals due to anomeric protons of sugar units and olefinic protons of hyperforins; a mid-field region
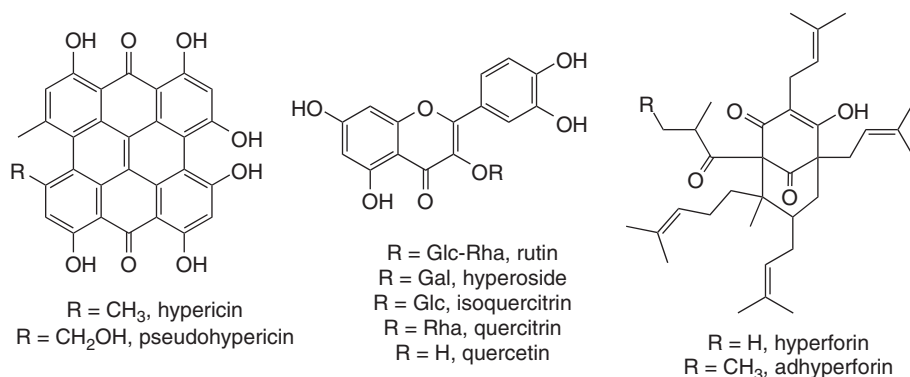
Figure 1 Characteristic constituents of St John's wort.
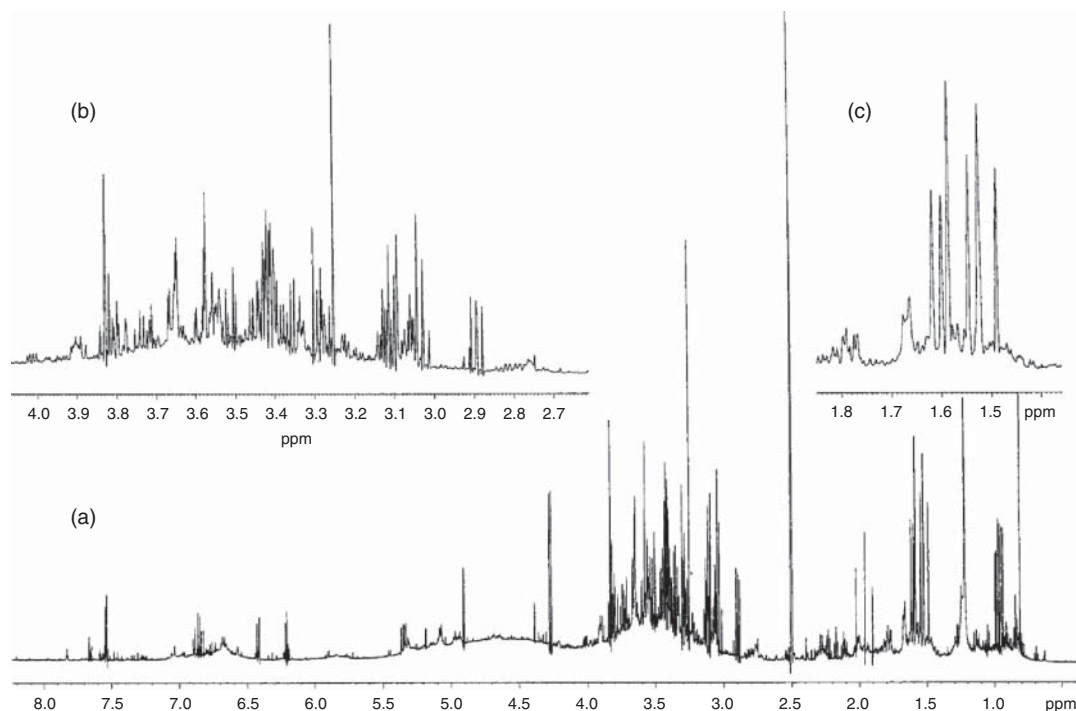
Figure 2 (a) Full-resolution-enhanced 600-MHz ^{1}H NMR spectrum of St John's wort commercial extract in dimethyl sulfoxide. (b and c) Horizontally expanded regions (2.6–4.1 and 1.35–1.85 ppm) of the spectrum to clarify the splitting pattern of individual resonances. (*Source*: Reprinted with permission from Bilia AR, Bergonzi MC, Mazzi G, Vincieri FF. Analysis of plant complex matrices by use of nuclear magnetic resonance spectroscopy: St. John's wort extract. 2001. J Agric Food Chem 49: 2115–2124. Copyright (2001) American Chemical Society.)

between 4.5 and 3.0 ppm with signals due principally to the sugar protons of the glycosides; and finally a high field region between 2.7 and 0.7 ppm with signals due to aliphatic protons of sugar residues (Me-6 of rhamnose), aromatic methyls of naphthodianthrones, and aliphatic protons of phloroglucinols.

By the detailed analysis of 2D-NMR experiments, in particular ^{1}H–^{1}H COSY and HMQC, the resonances of the characteristic constituents of St John's wort extract were unambiguously assigned, namely naphthodianthrones, flavonols, and phloroglucinols. In addition, other constituents such as caffeoylquinic acid derivatives, lipids, and traces of residual solvent were found.

A second part of the study was dedicated to the evaluation of stability of constituents of the extract, with special regard to the very unstable phloroglucinols (Bilia *et al.*, 2001b). On the basis of the assignments done, the decreased intensity of vinyl, methyl, and methylene proton signals of the isoprenylic chains of hyperforins, due to oxidative cleavage of the

isoprenyl side chains, was easily assessed, whereas signals of the other molecules were not significantly modified in the spectra obtained with samples submitted to stability testings. Owing to the difficulty in using an intensity reference peak, only qualitative and semiquantitative considerations were made, on the basis of relative intensity changes (Figure 4).

A further example of combination of high resolution NMR spectroscopy with pattern recognition is represented by the NMR analysis of commercially available extracts of kava kava containing from 30% to 70% w/w kavalactones (Bilia *et al.*, 2002a). These are represented by a mixture of more than 18 different γ-pyrones, being the major constituents kavain, methysticin, demethoxyyangonin, yangonin, dihydrokavain, dihydromethysticin, and 5,6-dehydromethysticin. These kavalactones are characterized by different double-bond linkage patterns in positions 5,6 and 7,8. Their pyrone moieties can be summarized by structures A–C as shown in Figure 5. Their activity is related to the presence of
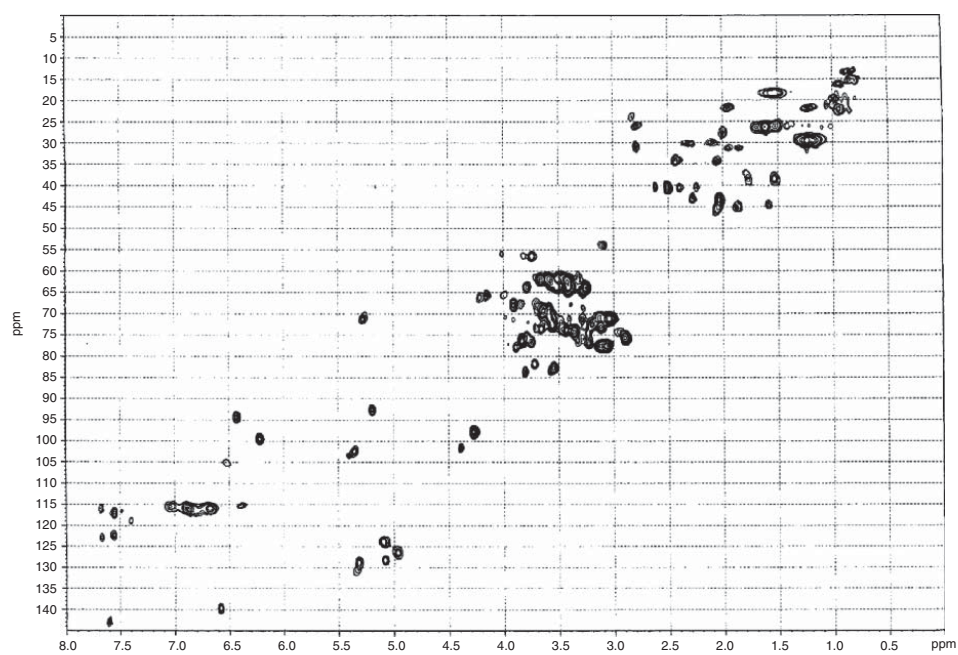
Figure 3 ^{1}H-^{13}C HMQC spectrum of St John's wort commercial extract. (*Source*: Reprinted with permission from Bilia AR, Bergonzi MC, Mazzi G, Vincieri FF. Analysis of plant complex matrices by use of nuclear magnetic resonance spectroscopy: St. John's wort extract. 2001. J Agric Food Chem 49: 2115–2124. Copyright (2001) American Chemical Society.)
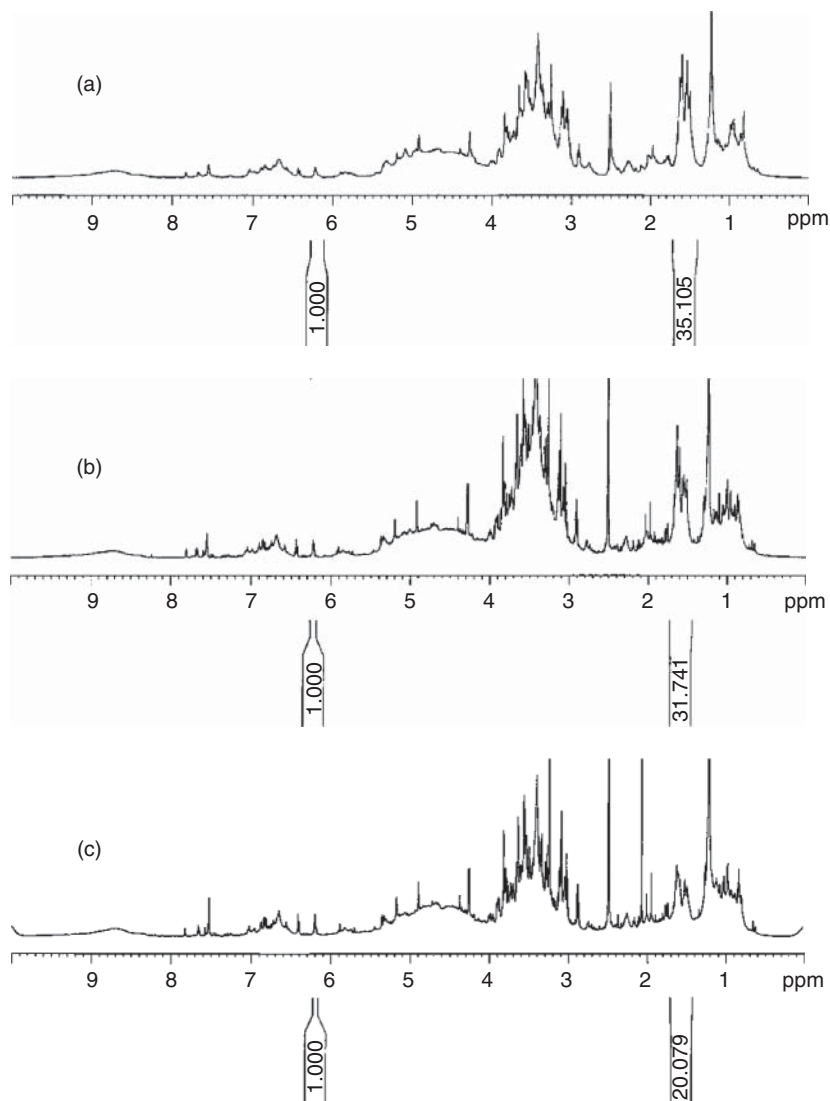
Figure 4 ^{1}H NMR spectra of St John's wort commercial extract (a) and extract submitted to thermal (b) and photostability (c) testing. (*Source*: Reprinted with permission from Bilia AR, Bergonzi MC, Mazzi G, Vincieri FF. Analysis of plant complex matrices by use of nuclear magnetic resonance spectroscopy: St. John's wort extract. 2001. J Agric Food Chem 49: 2115–2124. Copyright (2001) American Chemical Society.)

5,6 and/or 7,8 unsaturations, so it is very important to quantify the content of each kavalactone group.

The ^{1}H NMR spectrum (Figure 6) directly obtained by dissolving the commercial extracts in DMSO-d_6 can be quite easy and unambiguously analyzed to assign signals corresponding to H-3 of the different A–C systems. Thus, the chemical shift of H-3 is present in the region from 5.5 to 5.7 ppm for 5,6 unsaturated constituents (B system) and it is in the region from 5.1 to 5.3 ppm in 5,6 saturated ones (A and C systems). The signals corresponding to H-6 were observed in the range from 4.2 to 4.4 ppm in 7,8 saturated constituents (A system) and from 5.0 to 5.3 ppm in 7,8 unsaturated ones (C system). In the fully unsaturated compounds (B derivatives), signals of H-3 of two B-derivatives were identified at 5.61 ppm ($J_{3,5}$ = 2.2 Hz) and 5.65 ppm ($J_{3,5}$ = 2.2 Hz).
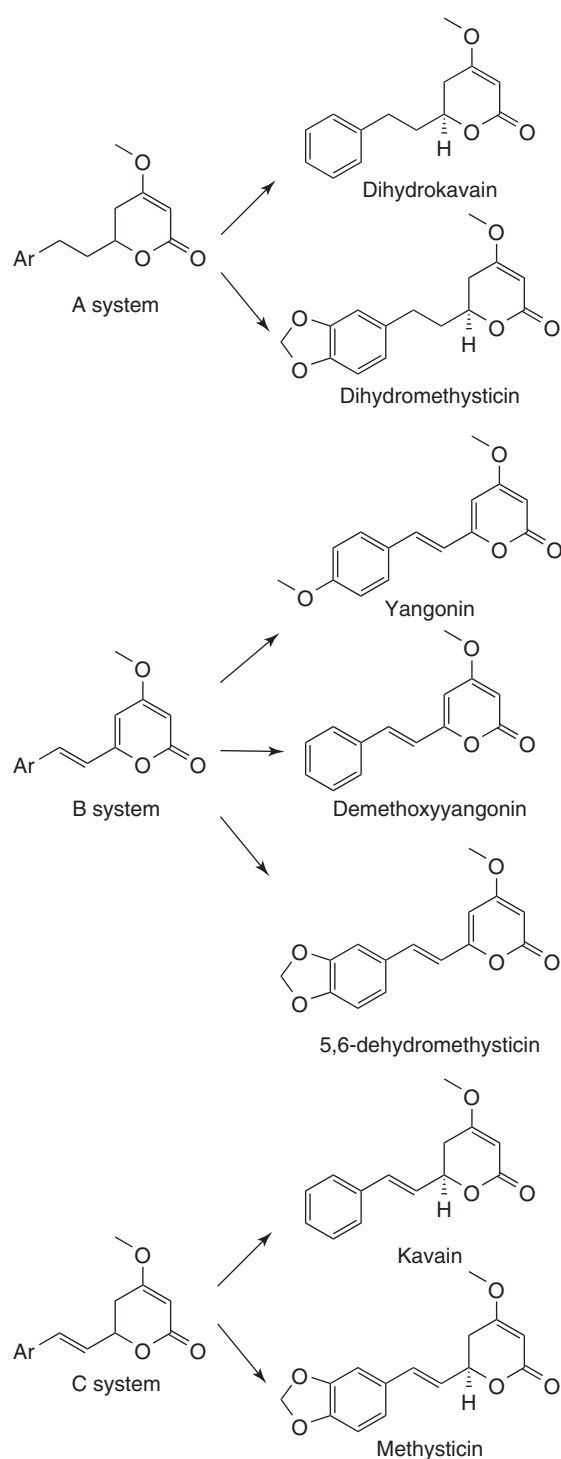
Figure 5 Kavalactones from kava kava.

COSY and HMQC experiments were recorded to assign the proton spin systems via scalar couplings and the ^{13}C chemical shifts with protons via one bond of kavalactones. In addition, 1D-TOCSY was revealed a very useful technique in the assignment of chemical shifts of each pyrone moiety. This experiment was obtained by selective irradiation of the different assigned H-3 and H-6 protons of A–C pyrone systems, and the resolution and unequivocal assignments of the signals have been achieved. In Figure 7, 1D-TOCSY is reported, which obtained from selective irradiation of signal at (a) 5.22 ppm (H-3 of C moiety), (b) 5.61 ppm (H-3 of B moiety), (c) 5.65 ppm (H-3 of a second B system), and (d) 6.42 ppm (H-7 of C moiety). In addition, TOCSY was also very useful for the unambiguous signal assignment of the styryl moieties and to assign the substitution pattern of aromatic rings. A semiquantitative approach, using maleic acid as internal standard, gives the relative ratio among the different derivatives (A–C), which are related to different activities. In the study, the suitability of NMR analysis in the evaluation of quality and identity of finely powdered herbal drug (HD) was successfully investigated through the direct NMR analysis of the solution obtained by the direct treatment of 50 mg HD in 0.8 mL DMSO-d_6, after filtration of the powder. This is the first report of its use in the direct analysis of an herbal drug. DMSO is a suitable and general solvent to completely dissolve metabolites from the herbal drugs, as demonstrated by the absence of metabolites in the chromatograms obtained by HPLC analysis of the residual powder (Bilia *et al*., 2002a).

The same authors (Bilia *et al*., 2002b) have also been disclosed the efficiency of COSY and HMQC experiments in obtaining a fingerprint of the constituents of an innovative extract of arnica, a supercritical carbon dioxide (CO_2) commercial extract. NMR allowed the identification of the signals corresponding to the characteristic constituents, sesquiterpenes (Figure 8), and all the other extracted metabolites.

The extract was directly dissolved in DMSO-d_6 and using a similar procedure described for the analysis of St John's wort and kava kava, and spectral assignments of the constituents were carried out according to the data (chemical shifts and coupling constants) found in the literature and by means of 1D- and 2D-NMR spectra. Typical resonances of helenalin and 11,13-dihydrohelenalin derivatives were easily attributed because of their characteristic chemical
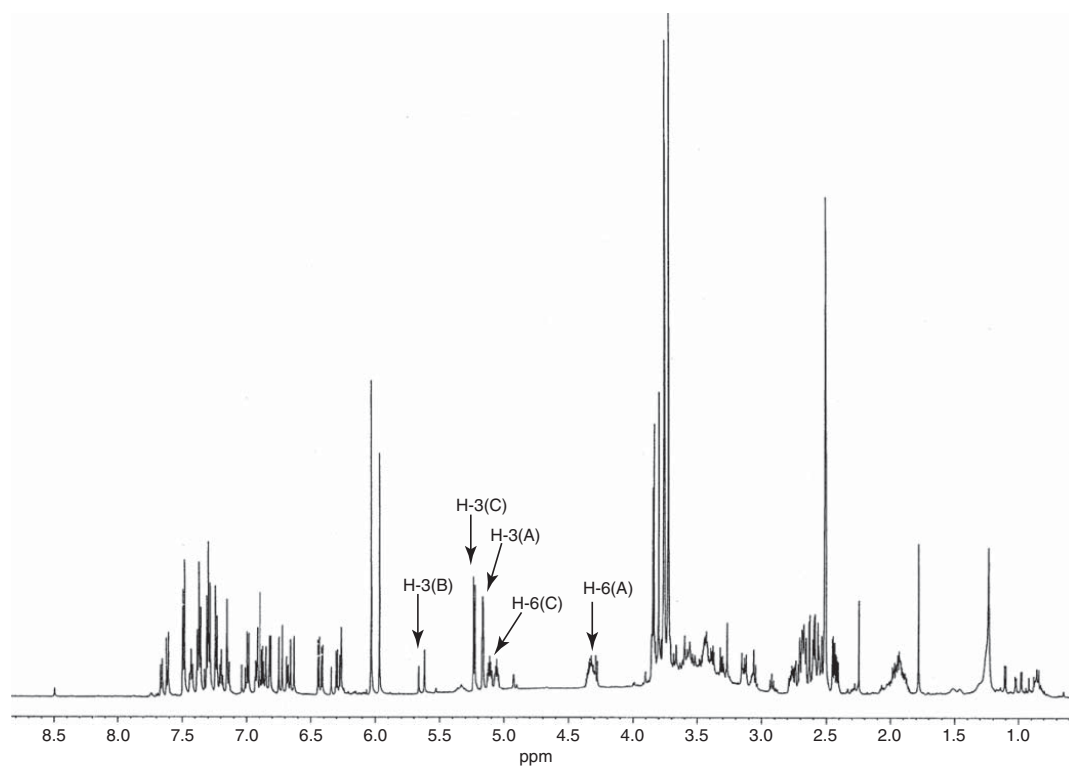
Figure 6 ^{1}H NMR spectra of kava kava commercial extract. (*Source*: Reprinted with permission from Bilia AR, Bergonzi MC, Mazzi G, Lazari D, Vincieri FF. 2002. Characterization of commercial Kava-Kava herbal drug and herbal drug preparations by means of nuclear magnetic resonance spectroscopy. J Agric Food Chem 50: 5016–5025. Copyright (2002) American Chemical Society.)
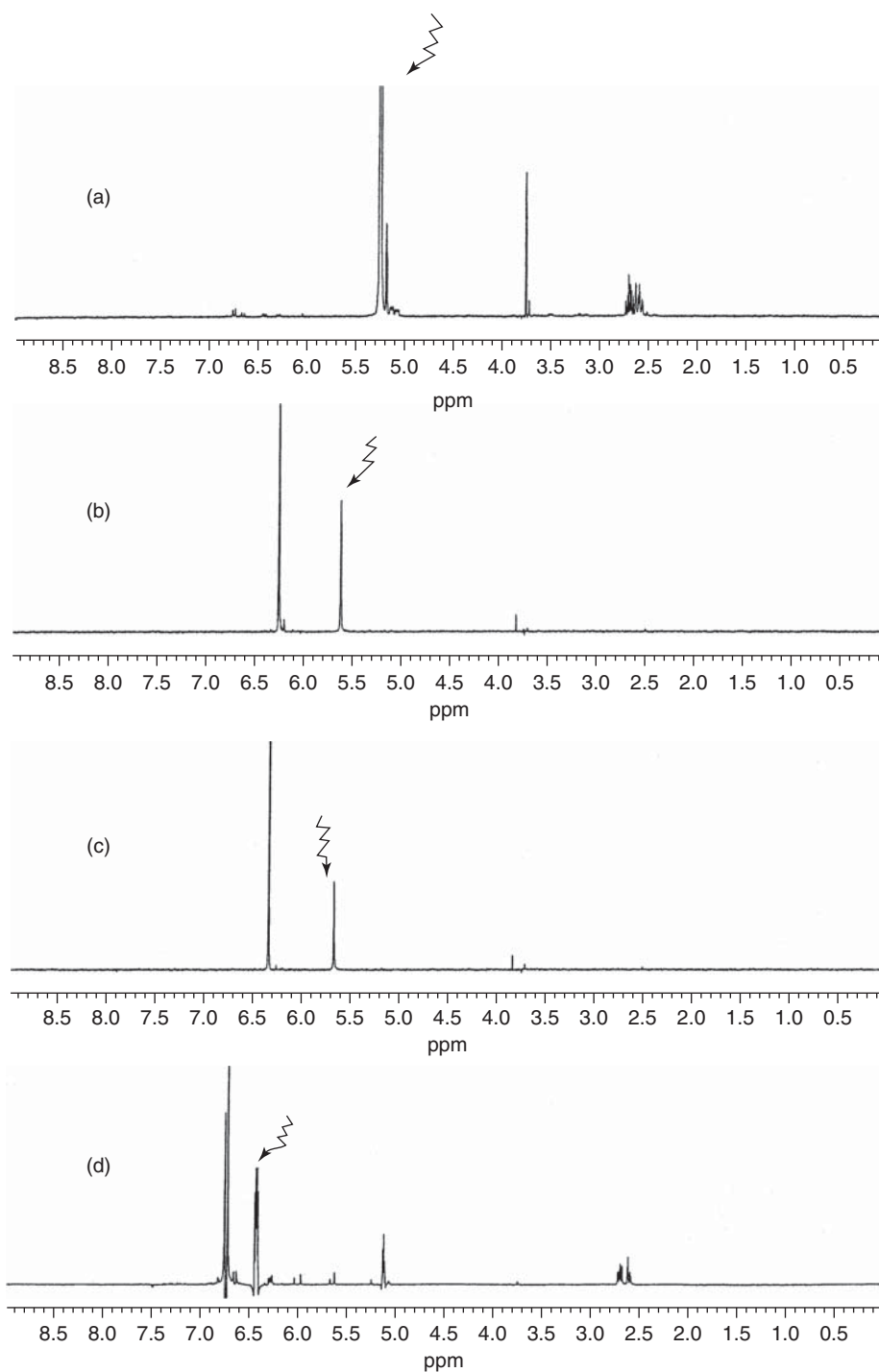
Figure 7 1D-TOCSY obtained from selective irradiation of signal at (a) 5.22 ppm (H-3 of C moiety), (b) 5.61 ppm (H-3 of B moiety), (c) 5.65 ppm (H-3 of a second B system), and (d) 6.42 ppm (H-7 of C moiety). (*Source*: Reprinted with permission from Bilia AR, Bergonzi MC, Mazzi G, Lazari D, Vincieri FF. 2002. Characterization of commercial Kava-Kava herbal drug and herbal drug preparations by means of nuclear magnetic resonance spectroscopy. J Agric Food Chem 50: 5016–5025. Copyright (2002) American Chemical Society.)

R = CH$_3$CO, -acetyl
R = CH$_2$=C(CH$_3$)CO, -α-metacryl
R = CH$_3$CH=C(CH$_3$)CO, tiglioyl
R = (CH$_3$)$_2$CHCO, isobutirryl
R = CH$_3$CH$_2$CH(CH$_3$)CO, -2-metylbutirryl
R = (CH$_3$)$_2$CHCH$_2$CO, -isovalerionyl
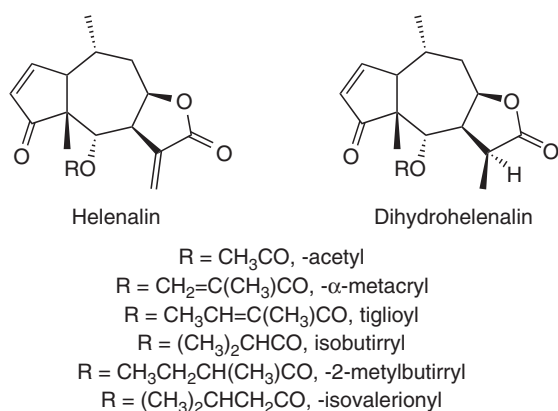
Figure 8   Sesquiterpene lactones from arnica.

shifts together with their splitting and couplings. The signals at 6.10 and 7.93 ppm were assigned to the

olefin protons of the typical ketofurane ring of helenalin derivatives, namely H-3 and H-2, respectively (Figure 9). Cross peaks in the COSY experiments led to the assignments of the signals of neighbor protons, unequivocally confirmed by HMQC experiments.

In the region between 5.5 and 6.5 ppm of the proton spectrum, two other signals were easily distinguished from the others, the resonances at 5.98 and 6.26 ppm (Figure 9). They were attributed to protons of α-methylene group exocyclic to the γ-lactone function of helenanin and its derivatives, namely H-13a and H-13b. These data were unequivocally confirmed through HMQC experiments by the connectivities of these two proton resonances with the carbon resonance at 124.2 ppm. COSY and HMQC experiments led to the complete assignment of all the protons and carbons of helenin. Furthermore, using a similar procedure, all the proton and carbon
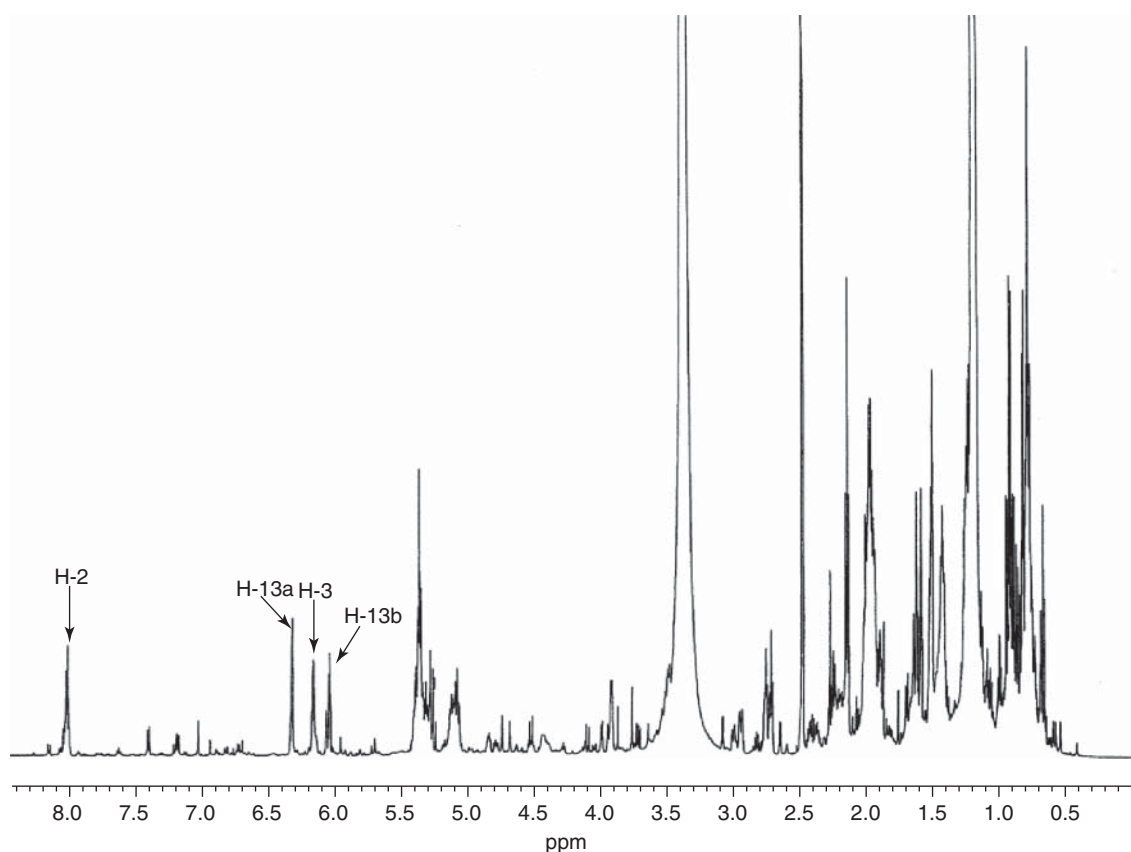
Figure 9   Full-resolution-enhanced 600-MHz ^{1}H NMR spectrum of arnica supercritical CO$_2$ extract. (*Source*: Reprinted from J Pharm Biomed Anal, 30, Bilia AR, Bergonzi MC, Mazzi G, Vincieri FF, NMR spectroscopy: a useful tool for characterisation of plant extracts, the case of supercritical CO$_2$ arnica extract, 321–330, Copyright (2002), with permission from Elsevier.)

resonances of dihydrohelenalin were assigned. In addition, signals of acetyl, metacrylic, tiglioyl, isobutirryl, 2-methyl-butirryl, and isovalerionyl moieties were identified by their chemical shift, splitting, and cross peaks in the COSY and HMQC experiments and confirmed by comparison with data reported in the literature. Moreover, NMR experiments were useful to identify signals of other noncharacteristic constituents, such as polyketides, aromatic derivatives, and esters of glycerol (Bilia *et al.*, 2002b).

In recent decade, a further NMR experiment, high resolution DOSY, has been used successfully to separate different compounds in a mixture (Barjat, Morris, and Swanson, 1998; McLachlan *et al.*, 2009). DOSY involves the synthesis of multidimensional NMR spectra in which one dimension displays chemical shift and the other the apparent diffusion coefficient, calculated by fitting of the measured decay of spin echo spectra recorded using pulsed field gradients.

An interesting application of DOSY in the analysis of plant extracts has been recently reported by León and coworkers (León, Chávez, and Delgado, 2011) to analyze the dimeric phthalides and other constituents (Figure 10) of *Ligusticum porteri* L. rhizome acetone extracts and to compare the constituents obtained from fresh and dried plant materials. It is well known a chemical variation in some constituents of *Ligusticum* species during storage and in some taxonomically related plant materials processed for preparations used in traditional Asian medicine.

In the first step of the study, a proton spectrum of the acetone extracts of *L. porteri* was acquired in order to identify the components, verifying the presence of the dimeric phthalides and establishing variations in the composition of the rhizome at different drying times. The assignments were made by comparisons with the resonances of authentic samples. The trace of proton spectrum shows the richness of the plant extract (of the fresh rhizomes) in the complexity of the profile. The spectrum was divided into four sections in order to facilitate signal assignment and the analysis led to the identification of the characteristic constituents and, in addition, also provided direct evidence for the decline of coniferyl ferulate with drying.

Additional evidence of the presence of the identified constituents was obtained from DOSY NMR experiments. DOSY provided both virtual separation (without physical separation) and structural information. It revealed four main diffusion rate levels: A, B, C, and D (Figure 11). Looking at the 7.00–4.3 ppm region, the signals that appear with a diffusion coefficient of 1.75×10^{-10} m^2 s^{-1} (highlighted as level A) correspond to a mixture of coniferyl ferulate and ferulic acid. On the next levels, B and C, which occurred at a diffusion coefficient range 2.20–2.45×10^{-10} m^2 s^{-1}, the most representative signals were found for diligustilide (H-7′ at δ 7.50, H-8 at δ 5.35, and H-8′ at δ 4.90) and tokinolide B (H-7′ at δ 7.64 and H-8′ at δ 4.45). This analysis corroborates the existence of the dimeric phthalides. The signals of the major compounds, which displayed a diffusion coefficient of 3.65×10^{-10} m^2 s^{-1} (level D), belong to the group of signals corresponding to *Z*-ligustilide.
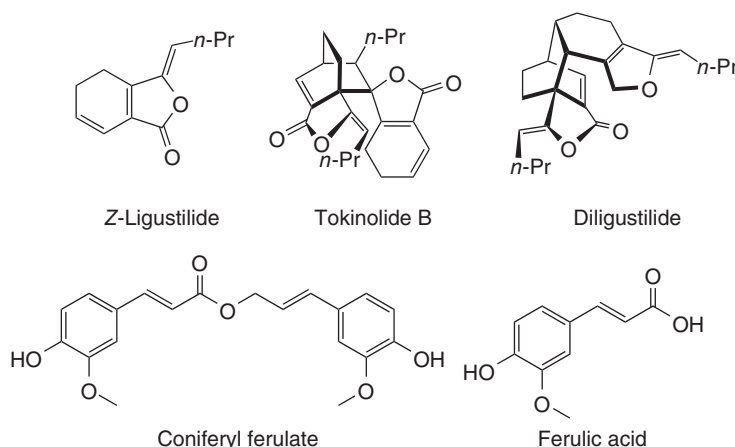
Figure 10 Characteristic constituents of *Ligusticum porteri*.
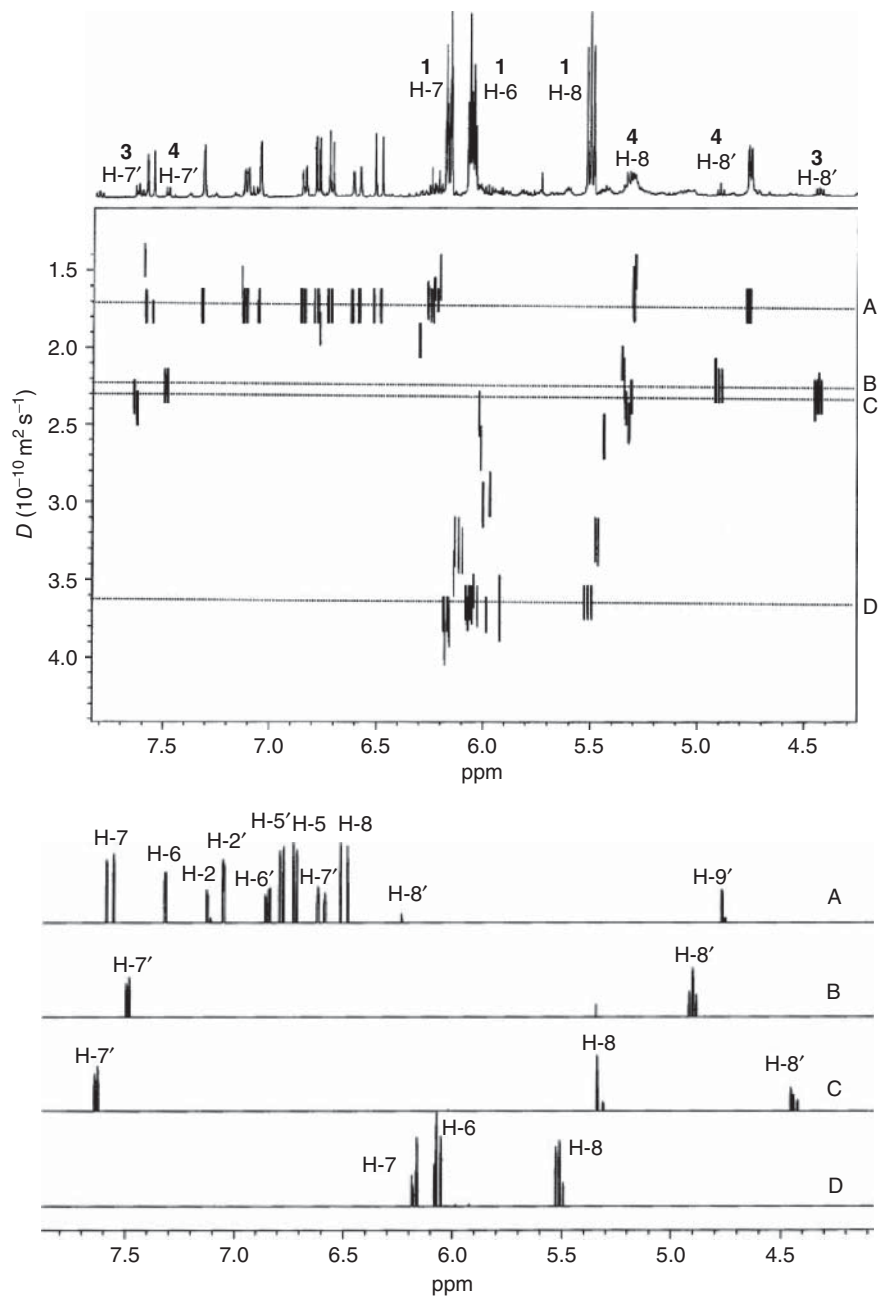
Figure 11 DOSY spectrum and DOSY slice spectrum of the acetone extract of *Ligusticum porteri*. Different diffusion coefficients: level A $(1.75 \times 10^{-10} \text{ m}^2 \text{ s}^{-1})$, mixture of coniferyl ferulate and ferulic acid; levels B and C $(2.20-2.45 \times 10^{-10} \text{ m}^2 \text{ s}^{-1})$ diligustilide and tokinolide B, respectively; and level D $(3.65 \times 10^{-10} \text{ m}^2 \text{ s}^{-1})$ *Z*-ligustilide. (*Source*: León, Chávez, and Delgado (2011). Reproduced from John Wiley and Sons, Ltd.)

DOSY was revealed a powerful tool for analyzing complex mixtures and confirmed that phthalides are natural products from this species and they are not formed as postharvest compounds (León, Chávez, and Delgado, 2011).

2.2 Quantitative NMR (qNMR) Analysis

qNMR is amenable to all NMR-sensitive nuclei and unrestricted in dimensions offering a unique and critical view of the analyte. Besides qNMR is almost as old as NMR itself, its application is greatly underestimated and literature has not emphasized this NMR prospect. Only recent developments in the field have provided evidence that NMR can be utilized as a precise quantitative tool, less time consuming (no equilibration time) with respect to the classical analytical methods, easy to perform, very specific and highly reproducible, and, in time, can even be a primary analytical tool (Pauli, Jaki, and Lankin, 2005).

The application of quantitative ^{1}H NMR (qHNMR) requires that at least one nonoverlapping signal for each molecule to be quantified is available for integration. Some experimental factors can interfere with quantitative determinations to make NMR a precise analytical tool, that is, "quantitative experimental conditions" such as relaxation delay, digitization, pulse sequence design, and instrumental parameters, together with the selection of appropriate postacquisition processing parameters for optimized spectral integration. All these aspects, including validation of the analytical method, have been reviewed and discussed in an outstanding manuscript by Pauli, Jaki, and Lankin (2005).

Proton NMR is the most suitable nucleus for quantitative studies and is preferred over the much more dispersive, but inherently less sensitive heteronuclei. Nowadays, typical errors fall in ^{1}H NMR the range 0.5–2%. ^{13}C NMR nucleus is considerably less sensitive (1.6% of ^{1}H sensitivity for an equal number of nuclei, augmented by a sensitivity loss due to the 1.1% natural abundance of ^{13}C) and affords quantitative information significantly more difficult to obtain, in particular, for small natural product samples.

The choice of reference compound should be exercised with great care, keeping in mind that reaction, complexation, acid–base, or any other type of chemical/physical transformation can occur. There is a tendency in the literature to propose internal reference standards with simple proton NMR spectra, preferably singlets. An ideal internal standard should be readily available in a highly pure form, inexpensive, stable and chemically inert, nonvolatile, not hygroscopic, and soluble in all (or most) of the NMR solvents. The most widely used is maleic acid, anthracene, trimesic acid, 3,4,5-trimethoxybenzaldehyde, and 1,3,5-trimethoxybenzene (Pauli, Jaki, and Lankin, 2005).

In the literature, there are many examples of successful application of qNMR to quantify the plant constituents, and in the majority of the studies, NMR results have been proved to be well in agreement with those obtained by HPLC.

Among medicinal plants, so far ginkgo has been the most attractive target for qHNMR analysis. *Gingko biloba* L. (gingko) extracts include different constituents, the most known are terpene trilactones, that is, ginkgolides A, B, C, D, and bilobalide, many polyphenols such as flavonol glycosides, biflavones, proanthocyanidins, alkylphenols, simple phenolic acids, 6-hydroxykynurenic acid, 4-*O*-methylpyridoxine, and polyprenols (Figure 12). The chemical analysis of ginkgo extracts is quite comprehensive, each class of constituents is evaluated with specific analytical methods and includes cleanup, purification, or derivatization steps. For instance, flavonoids that include a great variety of flavonol glycosides as mono-, di-, and triglycosides as well as their cinnamic acid esters, based on the aglycones kaempferol, quercetin, and isorhamnetin, are analyzed only after acid hydrolysis. Terpene trilactones are difficult to analyze because of their poor UV absorption due to the lack of a good chromophore, both for qualitative and quantitative purposes, and pure reference compounds are difficult to obtain commercially because of their limited purity and high price (van Beek, 2002).

The first qHNMR method proposed to determine the terpene trilactones in *G. biloba* dates back to 1993 (van Beek *et al.*, 1993). Using a preseparation method, van Beek *et al.* were able to quantitate five trilactones based on their H-12 olefinic proton signals using a 200-MHz spectrometer. The method is based on the comparison of the integral of each H-12 proton with that of the olefinic protons of the internal standard (maleic acid). These protons are all well separated at 200 MHz and occur in a less crowded region

R = H, kaempferol
R = OH, quercetin
R = OCH$_3$, isorhamnetin

R$_1$ = R$_2$ = H, ginkgolide A
R$_1$ = OH, R$_2$ = H, ginkgolide B
R$_1$ = OH, R$_2$ = OH, ginkgolide C
R$_1$ = H, R$_2$ = OH, ginkgolide J

Bilobalide
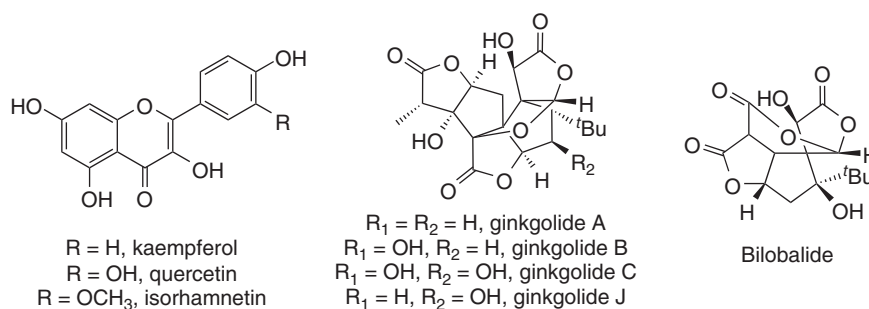
Figure 12 Characteristic constituents of *Ginkgo biloba* L.

of the NMR spectrum (6.15–6.50 ppm). The authors estimated that within a 30-min total experiment time they were able to quantitate circa 0.1 mg of each lactone and concluded that the method was as selective, reproducible, and sensitive as HPLC with refractometric detection, but with the advantage of not requiring pure trilactone reference substances. Regrettably, the method needs a chromatographic preparation step because some constituents (principally flavonoids) could interfere with the NMR analysis.

A decade after, Choi *et al*. (2003) improved the NMR method by the optimization of ^{1}H-NMR solvent for the analysis of the compound was selected through the evaluation of solvent effects on the resolution of the signals of H-6 and H-8 of the flavonoids that may interfere with the target proton H-12 signals of terpenes without any tedious chromatographic purification.

The addition of benzene-d_6 into acetone-d_6 was found to shift the 1H-NMR peaks of the H-6 or H-8 of the flavonoids to lower field and thus to be the optimum NMR solvent for the terpene analysis. This NMR solvent showed a good separation of the H-12 peaks of bilobalide and ginkgolides A, B, and C. The mixture of benzene-d_6 and acetone-d_6 (50 : 50) gave a good separation of all peaks originating from flavonoids and terpenoids. As an appropriate internal standard, phloroglucinol (1,3,5-trihydroxybenzene, 25 µg mL^{-1}) was used because it is stable, nonvolatile, and has a sharp singlet in the same areas of the target compounds but well separated. This method allows rapid and simple quantitation of underivatized bilobalide and ginkgolides in 5 min without any prepurification steps.

As the next step, the detection limit was measured for 64 scans of an extract of 500-mg plant material (acquisition time: 4 min). In this case, 10 ppm of each

compound was fully detected. Even 1 ppm could be detected by increasing to 1024 scans.

Li *et al*. (2004) extended qHNMR methodology to the simultaneous analysis of ginkgolides and flavonoids and have addressed problems because of limited extract solubility and the degradation of the internal standard phloroglucinol. A simple hydrolysis step was proposed as an alternative to the fraction step reported by van Beek and coworkers. Terpene trilactones were extraordinarily stable even in boiling HNO3. Utilizing this unique stability, Ginkgo extracts were hydrolyzed to simultaneously analyze the flavonol aglycones and terpenes (Figure 13). Each aglycone has its own characteristic H-2′ signal (or H-2′/6′ for kaempferol) in the ^{1}H NMR spectrum because of the different substitution patterns of ring B. In addition, these aglycon signals appear between 7.8 and 8.8 ppm, which generally is a not-crowded spectroscopic region and thus shows relatively little interference with other compounds, being the target signal to quantify the main aglycones: kaempferol, quercetin, and isorhamnetin. To improve solubility, a different solvent mixture (a 65 : 35 mixture of methanol-d_4 and benzene-d_6) and a different, more stable internal reference compound (1,3,5-trimethoxybenzene) were developed. The repeatability of the ^{1}H NMR method was found to be high with a maximum standard deviation of 3.2%. The results of recovery tests showed that hydrolysis did not affect the quantity of flavonol aglycones or terpene trilactones.

The determination of ginkgolic acids in ginkgo products represents another strategic analysis of the quality control because these constituents are indicated as possible hazardous compounds, being allergenic, limiting their content in *Ginkgo* extracts (5 or 10 ppm) (Blumenthal *et al*., 1998). However, a
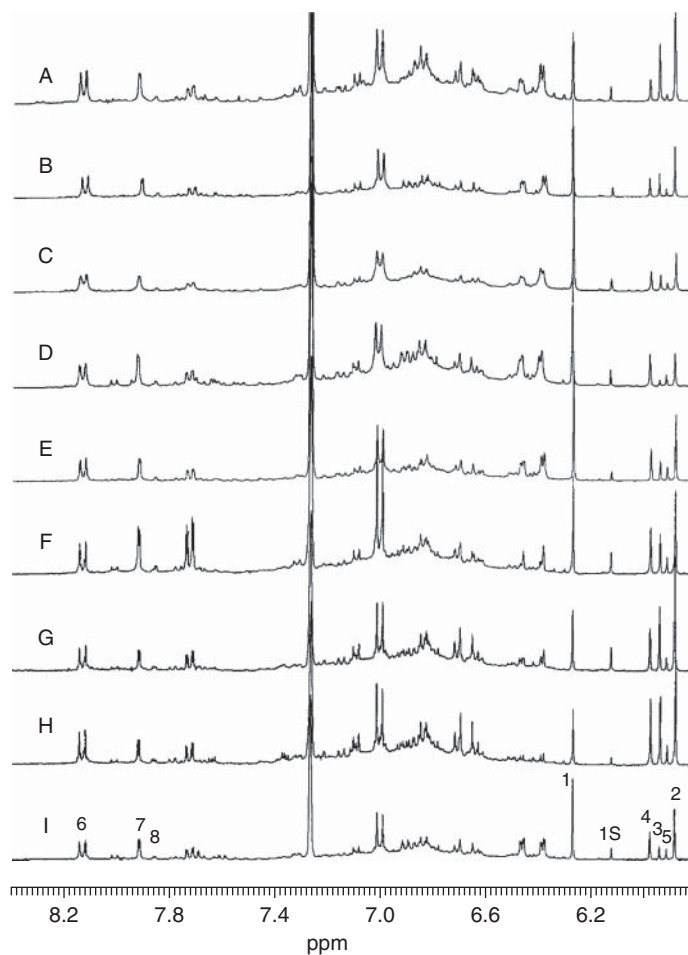
Figure 13 ^{1}H NMR spectra of commercial ginkgo extracts in the range δ 8.4–5.8. Sample numbers: (A) 1, (B) 2, (C) 3, (D) 4, (E) 5, (F) 6, (G) 7, (H) 8, and (I) 9. Peak numbers: 1, bilobalide; 2, ginkgolide A; 3, ginkgolide B; 4, ginkgolide C; 5, ginkgolide J; 6, kaempferol; 7, quercetin; 8, isorhamnetin; and IS, internal standard (1,3,5-trimethoxybenzene). (*Source*: Reprinted with permission from Li CY, Lin CH, Wu CC, Lee KH, Wu TS. 2004. Efficient ^{1}H nuclear magnetic resonance method for improved quality control analyses of ginkgo constituents. J Agric Food Chem 52: 3721–3725. Copyright (2004) American Chemical Society.)

number of ginkgolic acids with different side chains may be present and this makes their analysis by conventional chromatographic methods very complex (van Beek, 2002).

Choi and coworkers have reported the NMR analysis of this class of gingko constituents (Choi *et al*., 2004a), and a typical structure of ginkgolic acid, a C15:1 derivative, is shown in Figure 14. The proton spectrum of the chloroform extract (Figure 15) provides a simple pattern of peaks because all the ginkgolic acids possess the aromatic structure in common in which the proton signals are hardly affected by the difference in the alkyl side chain. In

fact, the chemical shifts of the characteristic aromatic signals of H-3 (δ 6.86, *d*, $J = 8.3$ Hz), H-4 (δ 7.35, *t*, $J = 8.1$ Hz), and H-5 (δ 6.76, *t*, $J = 7.5$ Hz) of the main ginkgolic acids are the same and are well separated. The quantity of the compounds was calculated from the relative ratio of the integral of each peak to the integral of the peaks of a known amount (100 μg) of anthracene used as an internal standard. H-9 and H-10 of anthracene were detected as a singlet at δ 8.42 and used as the reference peak of the internal standard for the analysis. The quantitative results obtained by ^{1}H-NMR analysis were comparable with those obtained by GC and ^{1}H-NMR method allows a
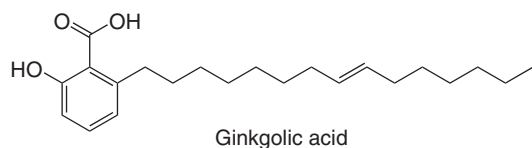
Figure 14   Structure of ginkgolic acid (C15:1).

simple quantification of total ginkgolic acids without any prepurification steps (Choi *et al*., 2004a).

In recent years, the ability of qHNMR to rapidly quantitate plant secondary metabolites without the need for prepurification or to obtain reference compounds was underlined by many groups of research. Hazekamp, Choi, and Verpoorte (2004) quantified cannabinoids from *Cannabis sativa* L.

The proton signals selected for this study were in the range δ 4.0–7.0, as this is the range where the ^{1}H-NMR spectra (Figure 16) are most distinguishable signals of each cannabinoid constituent. Part of the spectrum is enlarged to show the overlap of proton signals of Δ^9-tetrahydrocannabinolic acid A (THCA; Figure 17) with signals of minor compounds.

Anthracene was selected as internal standard because it is a very stable compound with a simple ^{1}H-NMR spectrum consisting of a singlet (δ 8.43) and two quartets (δ 8.01 and δ 7.48). For quantitation, the singlet of anthracene (*) and H-4 of THCA was used. These signals do not overlap with other signals of the cannabinoids. This method allows rapid and simple quantitation of THCA with
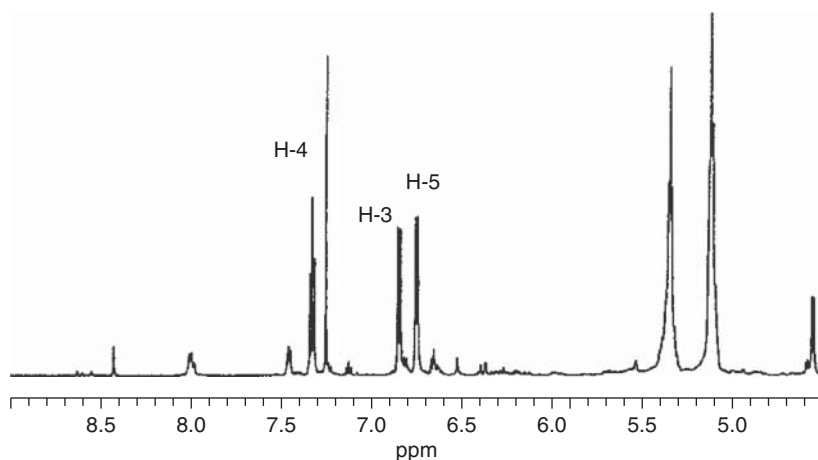
Figure 15   ^{1}H-NMR spectrum of the chloroform extract of leaves of *Gingko biloba*. (*Source*: Y.H. Choi *et al*. (2004). Reproduced from John Wiley and Sons, Ltd.)
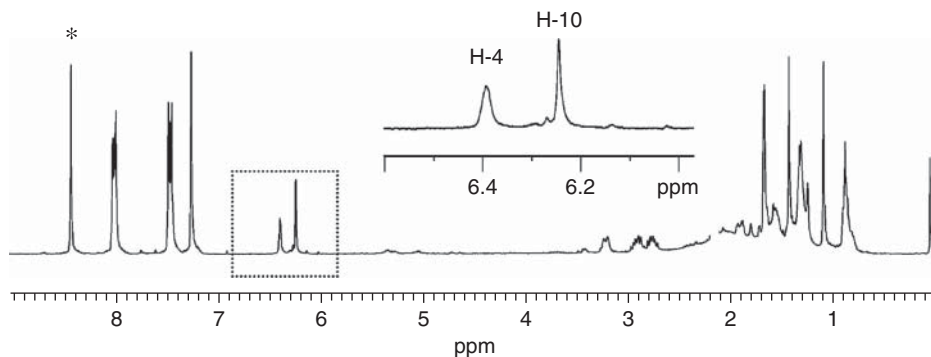
Figure 16   ^{1}H-NMR spectrum of a *Cannabis sativa* extract plus 1 mg of anthracene as Internal Standard. Part of the spectrum is enlarged to show the overlap of proton signals of THCA with signals of minor compounds. For quantitation, the singlet of anthracene (*) and H-4 of THCA were used. (*Source*: Reproduced with permission from Chemical & Pharmaceutical Bulletin Vol. 52 No. 6. Copyright (2004) The Pharmaceutical Society of Japan.)

Δ^9-Tetrahydrocannabinolic acid A
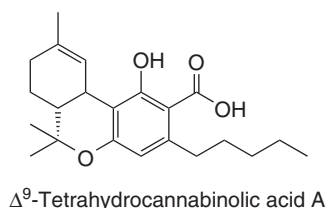
Figure 17 Structure of Δ^9-tetrahydrocannabinolic acid A (THCA).

Strychnine
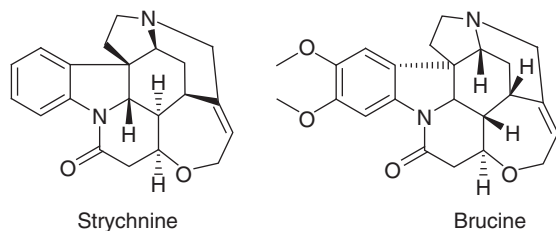
Brucine

Figure 18 Structure of brucine and strychnine.

Ephedrine

Figure 19 Structure of ephedrine.

a final analysis time of only 5 min, which is much shorter than conventional without the need for a prepurification step.

The same group of research also developed a qNMR analysis for the analysis of strychnine and brucine (Figure 18) in *Strychnos nux-vomica* L. seeds and stems (Frédérich, Choi, and Verpoorte, 2003). The analysis of the NMR spectra revealed that potentially the protons H-12 of strychnine (doublet) and brucine (singlet), resonating in a not-crowded region of the spectra, around δ 8.0, could be used for quantification. A study of the shift values of strychnine and brucine base, hydrochloride, sulphate, acetate, and nitrate in several NMR solvents revealed the need to add trifluoroacetic acid (1%, v/v) to the methanol-d_4, resulting in a controlled acidity and a good stability of shifts values. The results, showing relative standard deviations of <10%, are referring to both NMR method and extraction. The limit of quantification and the limit of detection have been determined, with a number of scans fixed at 256, as 0.05 and 0.005 mg mL^{-1} of alkaloid, respectively.

Among the advantages of this method, no reference alkaloids are needed for calibration curves, the quantification could be directly realized on a crude extract, strychnine and brucine could easily be distinguished, and less time consuming in comparison to conventional HPLC methods, for instance.

In a further study of the same group of research, ephedrine (Figure 19) and the analogs, pseudoephedrine, methylephedrine, and methylpseudoephedrine were simultaneously determined by ^{1}H-NMR (Figure 20) from *Ephedra* species without any precleaning steps, using anthracene as internal standard (Kim *et al.*, 2003). In the region of δ 5.0–4.0, the signals of H-1 attached to the same carbon with a hydroxyl were well separated from each other even though they are diastereomers. The chemical shifts of H-1 of ephedrine and pseudoephedrine are δ 4.77 (*d*, $J = 3.9$ Hz) and δ 4.17 (*d*, $J = 8.2$ Hz), respectively. Another pair of diastereomers and methylephedrine and methylpseudoephedrine are also well separated from each other and show up at δ 4.96 (*d*, $J = 3.8$ Hz) and δ 4.19 (*d*, $J = 6.4$ Hz), respectively. The amount of each alkaloid was calculated by the relative ratio of the intensity of H-1 signal to the known amount of internal standard, 200 µg of anthracene. This method allows rapid determination of the quantity of ephedrine alkaloids from *Ephedra* species.

Remarkable results on the application of qNMR to plant extracts are reported by Pieri *et al.* (2011). In a first study, ^{1}H NMR spectroscopy was successfully used for the characterization of *Stevia rebaudiana* Bertoni extracts. The quantitative determination of the major steviol glycosides (Figure 21) in purified extracts and fractions obtained from various stages of the purification process is reported and, in addition, NMR was a powerful tool to differentiate between glycosides that are naturally occurring in the plant and artifacts formed in the course of the manufacturing process. After the identification of steviol glycosides achieved principally by the use of 2D-NMR techniques, their quantification was based on qHNMR using anthracene as internal standard because its multiplet at δ 8.11 (4H) does not overlap with signals arising from constituents or solvents in the sample. The solvent mixture pyridine-d_5/DMSO-d_6 (6 : 1) enabled satisfactory separation of the signals to be integrated.
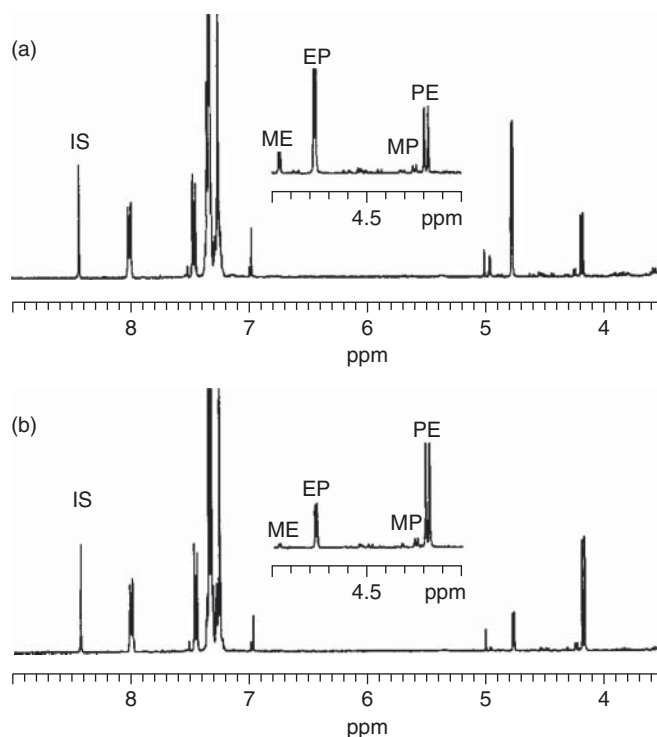
Figure 20 ^{1}H-NMR spectra *Ephedra sinica* (a) and *Ephedra intermedia* (b) in the range 3.5–8.5 ppm. (*Source*: Reproduced with permission from Chemical & Pharmaceutical Bulletin Vol. 51 No. 12. Copyright (2003) The Pharmaceutical Society of Japan.)

$R_1 = R_2 = \beta$-D-Glc, $R_3 = H$, stevioside
$R_1 = R_2 = R_3 = \beta$-D-Glc, rebaudioside A
$R_1 = H, R_2 = R_3 = \beta$-D-Glc, rebaudioside B
$R_1 = R_3 = \beta$-D-Glc, $R_2 = \alpha$-L-Rha, rebaudioside C
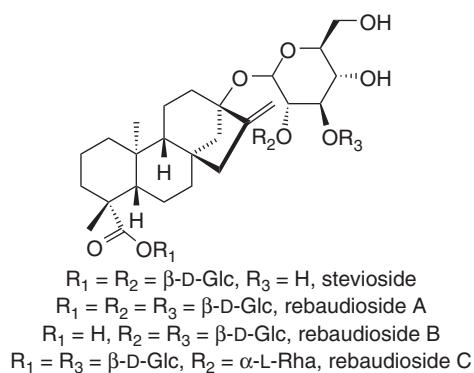
Figure 21   Major steviol glycosides from *Stevia rebaudiana*.

Validation of the method was performed in terms of specificity, precision, accuracy, linearity, robustness, and stability. Intraday precision, interday precision, and repeatability were respectively within 2.11%, 4.11%, and 4.11%, respectively, all expressed as relative standard deviations. Recovery rates for accuracy determination were within 101.9% and 95.5%.

Robustness was investigated by deliberate modification of shim, flip angle, and phase. A maximum deviation of 1.0% from the reference values was observed, which confirmed the robustness of the method. Linearity of the method could be confirmed in the concentration range 0.2–19.3 mg mL^{-1} for steviol glycoside RbA by constructing calibration curves using the signals of H-1″ and H-17a ($R^2 = 0.9995$ and $R^2 = 0.9997$, respectively).

The quantification of cynaropicrin (Figure 22), the major bitter principle of *Cynara scolymus* L. (artichoke) leaf extracts by means of quantitative ^{1}H NMR, was also performed and validated in terms of selectivity, linearity, precision, accuracy, and robustness (Pieri and Stuppner, 2011). The content of cynaropicrin is subjected to considerable variation according to the starting plant material, harvesting period, and drying conditions; therefore, the availability of validated methods for quantitative purposes is of great interest. The well-known instability of cynaropicrin is an important limitation for analytical methods, based on external calibration, such as HPLC-UV. In the proton NMR, the signals at δ 5.51
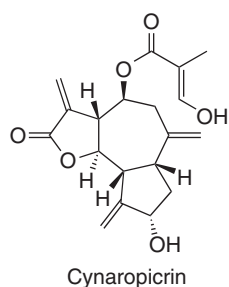
Figure 22 Structure of cynaropicrin.

(H-13b, d, $J = 3.0$ Hz) and δ 8.58 (2H, s) of cynaropicrin and anthracene used as internal standard were used for quantification.

The linearity range tested was between 5.94 and 475.20 μg mL^{-1} and the acquisition of one ^{1}H NMR spectrum was completed in about 17 min, with an LOQ (limit of quantitation) of 7.63 μg mL^{-1}. Relative standard deviations for intraday, interday precisions, and repeatability were, respectively, within 3.99%, 2.32%, and 1.50%. Recovery rates for accuracy determination were within 97.7% and 100.9%.

A further example of effectively targeted approach based on qHNMR is represented by the analysis of *Thymus vulgaris* L. extracts (Pieri *et al.*, 2012). Quality control of *Thymus* is generally achieved by the analysis of volatiles obtained by hydrodistillation, and this approach does not consider the nonvolatile fraction. Extracts were prepared by maceration of plant material with DMSO-d_6 containing anthracene as internal standard and analyzed by NMR after filtration.

Close inspection of the ^{1}H NMR spectra revealed that identified constituents thymol, *p*-cymene, γ-terpinene, linalool, and carvacrol (Figure 23) are characterized by at least one resonance, which is free of major interference from nearby signals, suggesting that such spectra are suitable for a targeted approach based on qHNMR. Anthracene was selected as internal standard because revealed a singlet at 8.58 ppm (2H,s) which showed no overlap with other resonances in the crude extracts (Figure 24). Validation was performed in terms of precision (intraday RSD B 4.51%, interday RSD B 4.18%), repeatability (RSD B 2.30%), accuracy (recovery rates within 93.4 and 103.4%), linearity (correlation coefficients 0.9990), robustness, and stability.

Quantitative ^{13}C spectroscopy is not as straightforward as quantitative ^{1}H NMR spectroscopy. In conventional ^{13}C NMR, the nuclear Overhauser effect (NOE) is applied to enhance signal intensity, but the achieved increment is not the same for all resonances; thus, signal intensity is not proportional to molar concentration. For quantitative determinations, a special pulse program with inverse gated decoupling is used, in which the NOE is suppressed. On the other hand, long delays have to be incorporated between experiments to ensure the relaxation of all ^{13}C resonances, which results in long acquisition times. To prevent this, a relaxation reagent to shorten the T1 (relaxation time) value is added to the samples.

A very interesting recent study (Palomino-Schätzlein *et al.*, 2011) examined the use of paramagnetic agent chromium(III) triacetylacetonate $(Cr(acac)_3)$ to quantify constituents using ^{13}C NMR. The addition of $Cr(acac)_3$ must be done cautiously, as excess will increase relaxation rates to the extent that resonances will broaden (fast T_2 relaxation) and will be difficult to detect. To optimize the $Cr(acac)_3$ concentration, an inversion/recovery NMR experiment was performed on a sample containing selected lipophilic secondary metabolites (model mixture A, nonacosane, lupeol, stigmasterol, phytol, and tocopherol) at known concentrations. The T_1 value was calculated for different $Cr(acac)_3$ concentrations optimizing the concentration at 60 mM, as T_1 was considerably reduced (0.7 s) and the spectral resolution still fell in a good range (line width at half height <1.5 Hz). Furthermore, no important chemical shift changes were observed.
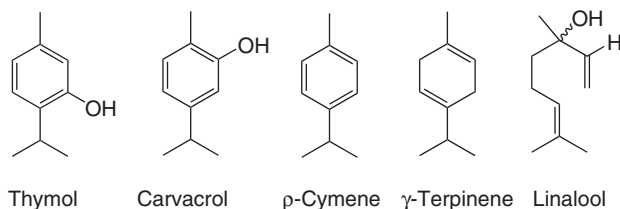
Figure 23 Structures of characteristic volatile constituents of *Thymus vulgaris* L.
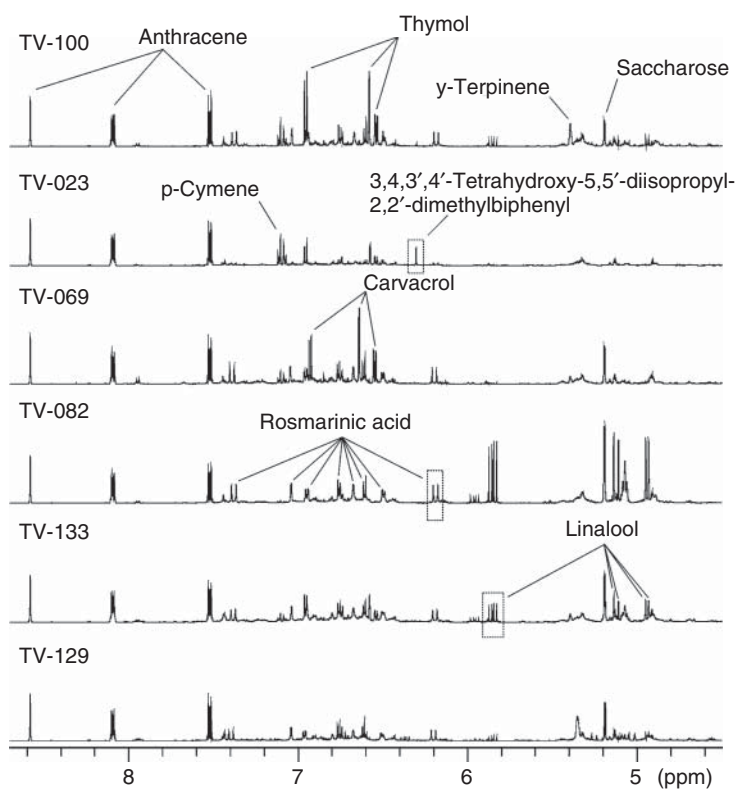
Figure 24 ^{1}H NMR spectra of selected *T. vulgaris* DMSO-d[6] extracts (4.6–8.6 ppm), (*Source*: Metabolomics, 8, 2012, 335, ^{1}H NMR-based metabolic profiling and target analysis: a combined approach for the quality control of Thymus vulgaris, Pieri V., Sturm S., Seger C., Franz C. and Stuppner S., Figure 3. With kind permission from Springer Science+Business Media B.V.)

Lipophilic extracts of white mulberry, bladder flower, and Mediterranean spurge were then used for the ^{13}C NMR analysis. In region of the spectrum ranging from 40 to 60 ppm, few signal overlaps take place and most signals are unambiguously identified, which was not possible for the ^{1}H spectrum.

The characteristic signals of alkanes, fatty acids, esters, and alcohols were clearly identified, even if the NMR spectra did not allow the determination of the exact chain length of the molecules. Furthermore, a characteristic region for triterpen-3-ols and triterpene esters at 78.4–79.2 and 80.4–81.2 ppm, respectively, was assigned, thus allowing the determination of the total quantity of this family of compounds in the extracts (Palomino-Schätzlein *et al*., 2011).

2.3 Chemometric Analysis of NMR Data

NMR generates either a metabolite profile, a quite difficult and time-consuming process achieving identification of the signals of the most important and/or specific metabolites, or a metabolite fingerprint, in which the analysis is based on the distribution of intensity in the NMR spectrum rather than the assignment of the signals (Krishnan *et al*., 2004). In this case, NMR-related techniques can be understood as "high throughput, rapid, global analysis of samples to provide sample classification" (Dunn and Ellis, 2005; Hall, 2006).

Consequently, signals in NMR spectra are identified for a number of accessions that correlate samples according to biological activity, quality control, toxicity prediction, biomarker screening, or geographical origin. Fingerprinting ignores the assignment problem presented by the multitude of signals in a high resolution ^{1}H NMR spectrum and, instead, uses multivariate analysis to compare sets of spectra and hence the samples from which the spectra were derived.

^{1}H- and ^{13}C-NMR analysis in combination with chemometrical methods as fingerprinting technologies is known since the 1980s, being the

first application of chemometrics to NMR spectra appeared by Johnels *et al*. (1983) In recent decades, sample preparation, NMR data generation, and statistical data exploration used for the assessment of plant extracts have strongly evolved as reported in details in some recent reviews (Holmes *et al*., 2006; Kim and Verpoorte, 2010; Verpoorte, Choi, and Kim, 2007; Verpoorte *et al*., 2008).

Multivariate statistical analysis can take several forms and linear discriminant analysis (LDA), hierarchical clustering trees (HCTs), and principal component analysis (PCA) are generally reported in the field of plant extract analysis. The simplest methods such as cluster analysis and PCA are termed unsupervised methods whereby the spectra are grouped according to their similarities without using the knowledge of sample class. Thus, these methods generate an unbiased overview of natural clustering among biochemically similar samples, generating a classification of plant extracts.

PCA is essentially a descriptive method. This method is, normally, the first step in data exploration, which allows the main variability aspects of a data set to be visualized, without the constraint of an initial hypothesis concerning the relationship within samples and between samples and responses (variables). The main goals of this procedure are to find relationships between the different parameters (objects and variables) and to detect possible clusters within objects and/or variables.

This method condenses the multivariate data into a reduced number of principal components that describe the greatest amount of variance. Each PC is a linear combination of the original variables whereby each successive PC explains the maximum amount of variance possible in the dataset and each PC is orthogonal to every other PC. By applying such a technique to the NMR data, similarities and differences between samples of interest can be visualized in simple two- or three-dimensional plots. Many factors including the intrinsic pH of the extract, NMR probe temperature stability, and quality of residual solvent suppression will all affect the robustness of the used multivariate models. Defernez and Colquhoun (2003) have summarized some possible pitfalls and solutions in an excellent review, including signal alignment methods that can be applied postacquisition to suboptimal data sets.

NMR-based metabolic profiling has been successfully employed in the analysis of food and food products, and early applications focused on fruit juices were aimed to validate the ability of these techniques to discriminate between samples derived from different varieties of the same fruit. The studies on plant extracts have aimed to authentication and quality control, prediction of biological and toxicolocal activities, chemotaxonomic analysis, and discrimination of cultivars.

Verpoorte and coworkers have reported a series of publication of successful application of multivariate data analysis to classify extracts from different species or cultivars, species authentication, and quality control.

^{1}H NMR spectrometry and multivariate analysis techniques have been applied for the metabolic profiling of three *Strychnos* species: *S. nux-vomica* L. (seeds, stem bark, and root bark), *Strychnos ignatii* Bergius (seeds), and *Strychnos icaja* Baill (leaves, stem bark, root bark, and collar bark) (Frédérich *et al*., 2004). The principal aim of the investigation was to explore this analytical tool to enable an efficient identification (metabolic fingerprinting) of the nature of *Strychnos* samples (species and organs). Secondarily, the method was also applied to a series of samples of "false angostura bark" tree and is used in traditional medicine to treat dysenteries, paralytic infections, and as a tonic and to make bitter liquors. The PCA of the ^{1}H NMR spectra showed a clear discrimination between all *Strychnos* species from various origins, using the key compounds responsible for the discrimination brucine, loganin, fatty acids, and *S. icaja* alkaloids such as icajine and sungucine. The method was then applied to the classification of several "false angostura" samples that are a falsification of angostura bark (*Galipea officinalis* Hancock). These samples were, as expected, identified as *S. nux-vomica* by PCA but could not be clearly discriminated as root bark or stem bark samples after further statistical analysis.

A study of 11 *Ilex* species, including *Ilex paraguariensis* A. St Hil., used to produce a herbal tea (mate) drunk was carried out by NMR spectroscopy and multivariate data analysis (PCA) to a clear discrimination of *I. paraguariensis* from the other *Ilex* species used as substitutes or adulterants (Choi *et al*., 2005). The analysis was carried out from both aqueous and organic extracts. The major metabolites that contribute to the discrimination were arbutin, caffeine, theobromine, and phenylpropanoids (Figure 25). In the classification based on the metabolites obtained from organic fractions, most of the species do not overlap and showed unique metabolomic profiles as
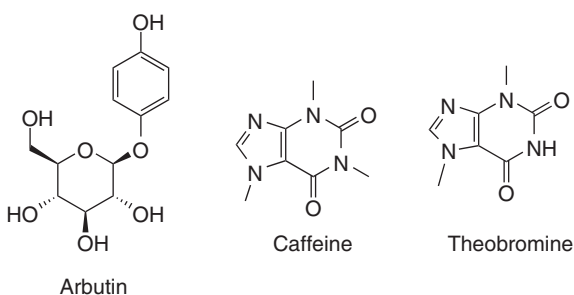
Figure 25 Structures of some characteristic constituents of *Ilex paraguariensis*.

Cannabidiolic acid

Figure 26 Structure of cannabidiolic acid.

an example of *I. paraguariensis* with the exception of *I. pseudobuxus* Reissek, *I. brasiliensis* (Spreng.) Loes., and *I. theezans* C. Martius ex Reisseck. In the classification of metabolites obtained from aqueous fractions, no species overlapped and the metabolite profiles were all clearly distinguished from each other. The study proves that it is possible to discriminate the 11 *Ilex* species by multivariate analysis of their metabolite fingerprints obtained by ^{1}H NMR spectra of crude extracts of the plant materials.

The metabolomic analysis of *Ephedra* species using ^{1}H-NMR spectroscopy and multivariate data analysis was another example of successful tool for chemotaxonomic analysis and authentication and quality control of plant extracts (Kim *et al.*, 2005). The PCA used to reduce the huge data set obtained from the ^{1}H NMR spectra of the plant extracts. This analysis clearly and successfully discriminated three different *Ephedra* species, namely *E. sinica* Stapf., *E. intermedia* Schernk ex C. A. Mey., and *E. distachya* L. var. *distachya*. Benzoic acid analogs in the aqueous fraction and ephedrine-type alkaloids in the organic fraction represented the major metabolites that contributed to the differentiation. On the basis of this metabolomic recognition, one of nine commercial *Ephedra* materials evaluated was shown to be a mixture of *Ephedra* species.

A similar application was used to analyze 12 *C. sativa* L. cultivars allowed the discrimination of the investigated samples without any prepurification steps. Both Δ^9-tetrahydrocannabinolic acid (Figure 17) and cannabidiolic acid (Figure 26) exerted a strong influence on the segregation of the sample groups in the PCA scores plot, leading to distinguish the cultivars. The discrimination of the cultivars could also be obtained from a water extract containing carbohydrates and amino acids: sucrose,
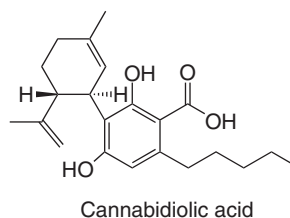
glucose, asparagine, and glutamic acid are found to be major discriminating metabolites of these cultivars (Choi *et al.*, 2004b).

Some outstanding examples of the use of multivariate analysis in quality control and prediction of biological activity have been reported by Holmes group research.

A combination of high resolution NMR spectroscopy with pattern recognition has been employed to investigate commercially available *Tanacetum parthenium* (L.) Schultz Bip samples originating from several different manufacturers (Bailey *et al.*, 2002). Initial observation of the proton spectra from 14 different suppliers (A–N) proved that spectra were reasonably similar (Figure 27). Exceptions to this are classes I and K that appear very different from the remaining 12 classes. Even in products originating from the same manufacturer, batch-to-batch variation was evident, emphasizing the need for better standardization of such products. PCA following multivariate analysis of the ^{1}H NMR spectra for the different feverfew samples (classes) showed that while two classes (I and K) were well separated with respect to the other samples in the first three PCs, the remaining 12 classes were difficult to differentiate.

NMR-based metabolomic strategy has been fruitfully employed to determine the composition of three chamomile (*Matricaria recutita* L.) samples obtained from three geographical regions with the aim to differentiate the origin and herb quality (Wang *et al.*, 2004).

A typical ^{1}H NMR spectrum of a chamomile extract is shown in Figure 28. The resonance assignments were made with the assistance of COSY and TOCSY NMR spectroscopy. The spectra were dominated by glucose, sucrose, and a number of amino acids including alanine, threonine, leucine, isoleucine, glutamine, asparagine, glutamate, and valine. Identification of chlorogenic acid and α-bisabolol (Figure 29) was confirmed by
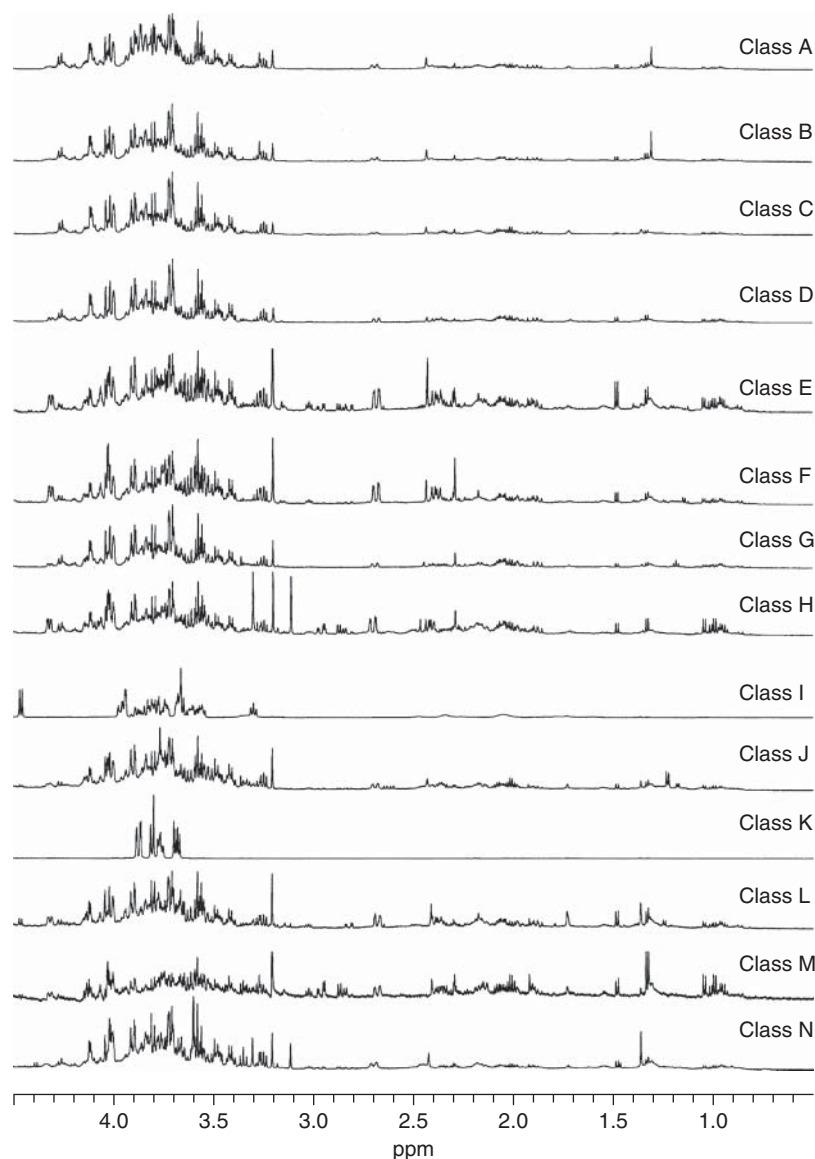
Figure 27   ^{1}H NMR spectra of the 14 feverfew sample classes. (*Source*: Bailey NJC, Sampson J, Hylands PJ, Nichlson JK, Holmes E, Multi-Component Metabolic Classification of Commercial Feverfew Preparations via High-Field 1HNMR Spectroscopy and Chemometrics, Planta Med, 2002, 68, 734-738, © Georg Thieme Verlag KG · Stuttgart · New York.)

spiking with the authentic samples of chlorogenic acid and α-bisabolol.

It has been demonstrated that this method is capable of identifying samples of chamomile from three different locations and capable of discriminating between samples with different preparation procedures, thus facilitating the development of a sensitive methodology for the quality control of plant extracts (Wang *et al*., 2004).

The same group of research has been successfully carried out studies on the extracts of *Artemisia annua* L. to discriminate samples from different sources and classify them according to their antiplasmodial activity and their toxicity to cell cultures without prior knowledge of this activity (Bailey *et al*., 2004).

Representative ^{1}H NMR spectra are shown in Figure 30. These spectra (representing extracts of
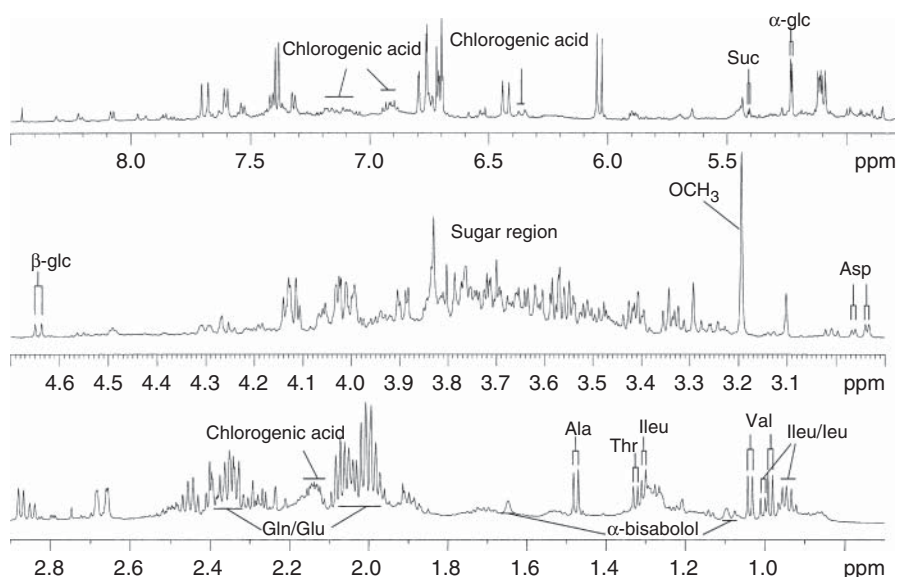
Figure 28 ^{1}H NMR spectrum of chamomile extract. (*Source*: Wang Y, Tang H, Nicholson JK, Hylands PJ, Sampson J, Whitcombe I, Stewart CG, Caiger S, Oru I, Holmes E, Metabolomic Strategy for the Classification and Quality Control of Phytomedicine: A Case Study of Chamomile Flower (Matricaria recutita L.), Planta Med, 2004, 70, 250-255, © Georg Thieme Verlag KG · Stuttgart · New York.)
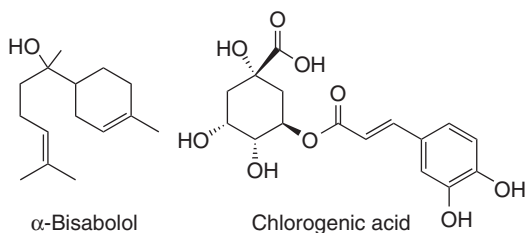
Figure 29 Structures of chlorogenic acid and α-bisabolol.

different IC$_{50}$ values) show that while the artemisinin (Figure 31) resonances themselves are readily apparent, particularly in the more potent extract 7, there are many other regions of the spectrum where large differences occur between the different extracts. In particular, the region between 6.8 and 8.0 ppm has marked differences in the spectra.

The PCA algorithm was used to reflect the biological activities of each sample. A coding was performed whereby the data were split into three classes based on their IC$_{50}$ value, with the cutoff values being 0.1 and 1 µg mL^{-1}. The resulting PCA scores plot can be seen in Figure 32. It is apparent that this model (containing 78% variance in the first two PCs) is able to discriminate between

the three classes used. This suggests that ^{1}H NMR spectra contain sufficient information relating to the physicochemical properties of the extract to be able to predict the potential magnitude of antiplasmodial activity found in a plant extract.

While PCA clearly demonstrates the potential of this technique, a more robust approach to obtaining predictive data was employed, the supervised methods. These methods involve providing the model with the values for the variable to be predicted (i.e., IC$_{50}$ value) for part of the data set (the training set), with the model then being optimized based on those values. Because the algorithm in effect uses the answers to create the model, it is then necessary to validate this model using the remaining unused samples (the test set). These samples with an IC$_{50}$ value >1 µg mL^{-1} were excluded from this analysis, for two reasons. This class is the smallest of the three and having larger values means that the model is likely to be skewed in order to consider them. In addition, the higher values mean that these samples are not of interest anyway, as they essentially have no activity. Using the remaining two classes as above (IC$_{50}$ < 0.1 µg mL^{-1} and IC$_{50}$ > 0.1 µg mL^{-1}, respectively), it is possible to construct a "dummy" *y*-matrix whereby the two classes are represented by a 1 or a 0. Partial least squares discriminant analysis (PLS-DA)
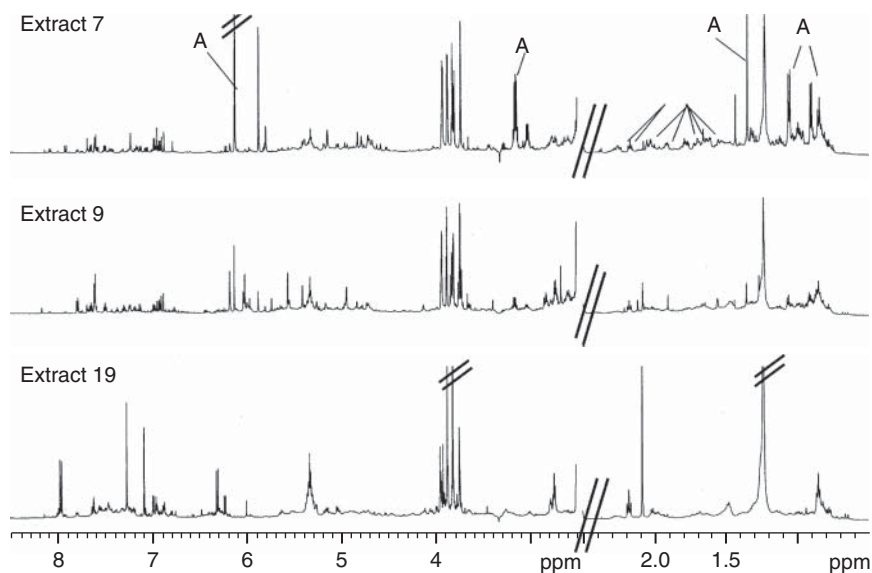
Figure 30 Representative ^{1}H NMR spectra for the three IC_{50} classes, $IC_{50} < 0.1\,\mu g\,mL^{-1}$ (extract 7), $IC_{50} > 0.1\,\mu g\,mL^{-1}$, $<1\,\mu g\,mL^{-1}$ (extract 9), and $IC_{50} > 1\,\mu g\,mL^{-1}$ (extract 19). Region 2.5–8.5 ppm is expanded vertically by a factor of 6 to allow observation of lower level aromatic resonances. Resonances attributable to artemisinin are indicated with an "A." (*Source*: Reprinted from J Pharm Biomed Anal, 35, Bailey NJC, Wang Y, Sampsonc J, Davis W, Whitcombec I, Hylands PJ, Croft SL, Holmes E, Prediction of anti-plasmodial activity of Artemisia annua extracts: application of H NMR spectroscopy and chemometrics, 117-126, Copyright (2004), with permission from Elsevier.)
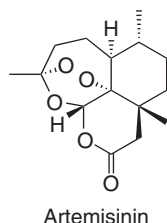
Artemisinin

Figure 31 Structure of artemisinin.

can then be performed on the data to construct a new model using this additional data. This analysis allowed the prediction of actual values for antiplasmodial activities for independent samples not used in producing the models. PLS-DA was useful to maximize the separation between two or more sample classes based on prior knowledge of class membership and uses a discrete class identifier in the *y*-matrix rather than a continuous measure of response. The models were constructed using approximately 70% of the samples, with 30% used as a validation set for which predictions were made (Bailey *et al.*, 2004).

In a further study, proton NMR spectra of various extracts of St John's wort (*H. perforatum*) samples derived from four different accessions obtained using six different extraction solvent mixtures were chemometrically evaluated to predict the pharmacological efficacy. In a first approach, a consistent correlation for the spectroscopic pattern of the extracts and the corresponding IC_{50} values derived from nonselective binding to opioid receptors was found.

In a second approach, a partial least squares (PLS-1 and PLS-2) regression model was used to predict the biological activity of eight St John's wort extracts based on two pharmacological data sets: nonselective binding to opioid receptors and antagonist effect at corticotrophin-releasing factor type 1 (CRF_1) receptors. The PLS 2 model, using the two pharmacological targets, confirmed the useful application of this approach to assess the quality of St John's wort extracts.

However, no details about the spectral regions responsible for the correlation with the pharmacological effects have been disclosed and the identification of putative molecular carriers of the bioactivity has not been achieved (Roos *et al.*, 2004).

Le Gall, Colquhoun, and Defernez (2004) analyzed by ^{1}H NMR a set of 190 samples of *Camellia chinensis* (L.) Kuntze (green tea) originating from
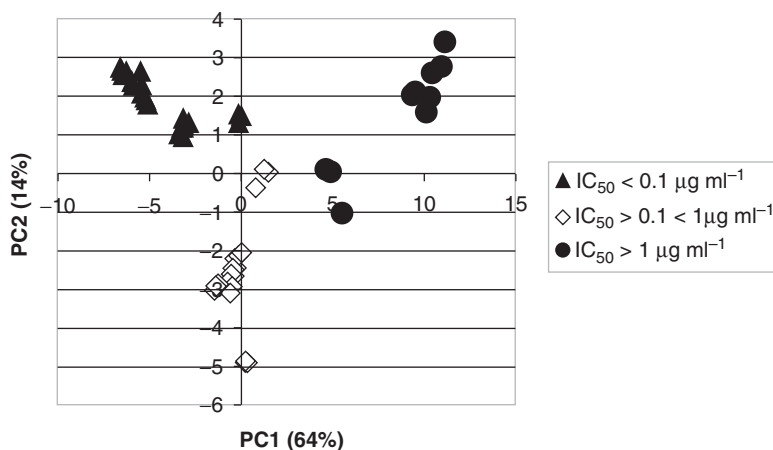
Figure 32 PCA scores plot for *Artemisia annua* plant extracts. Samples are separated into three groups: $IC_{50} < 0.1$ μg mL^{-1} (▲), $IC_{50} >$ 0.1 and <1 μg mL^{-1} (◇) and $IC_{50} > 1$ μg mL^{-1} (●). (*Source*: Reprinted from J Pharm Biomed Anal, 35, Bailey NJC, Wang Y, Sampsonc J, Davis W, Whitcombec I, Hylands PJ, Croft SL, Holmes E, Prediction of anti-plasmodial activity of Artemisia annua extracts: application of H NMR spectroscopy and chemometrics, 117–126, Copyright (2004), with permission from Elsevier.)

different regions to establish if the teas could be discriminated according to the country of origin or with respect to quality. Both PCA and cluster analysis were applied to the data. The quality of green tea is judged on a combination of its appearance, flavor, and aroma with amino acids, catechins, and caffeine (Figure 33) contributing strongly to the overall quality.

About 30 compounds (Figure 34) were identified in the 1D- and 2D-spectra, and more than 50 signals or groups of signals were indexed overall. Apart from ubiquitous compounds such as amino and fatty acids and common sugars such as sucrose and glucose that are assigned in the literature compilations, the signals of phenolics, flavonoids (catechins and flavonols), xanthines, and minor sugars can be

observed. Theanine, unique to tea, is the predominant amino acid present.

Using PCA, it was possible to separate Chinese teas from the non-Chinese teas in the first component. Moreover, "high quality" Chinese teas were also well clustered. By analysis of the PCA loading plots, higher levels of theanine, theogallin, epicatechin gallate, gallic acid, caffeine, and theobromine were assigned to these samples, whereas the relative concentrations of fatty acids, quinic acid, sucrose, and epigallocatechin were higher in lower quality teas.

NMR analysis was able to generate a more comprehensive overview and better classification of teas than chromatography-based methods.

$R_1 = OH$, $R_2 = $ gallate, (−)-epigallo-catechin-3-gallate
$R_1 = H$, $R_2 = $ gallate, (−)-epicatechin-3-gallate
$R_1 = OH$, $R_2 = H$, (−)-epigallo-catechin
$R_1 = R_2 = H$, (−)-epicatechin

Caffeine   Theobromine
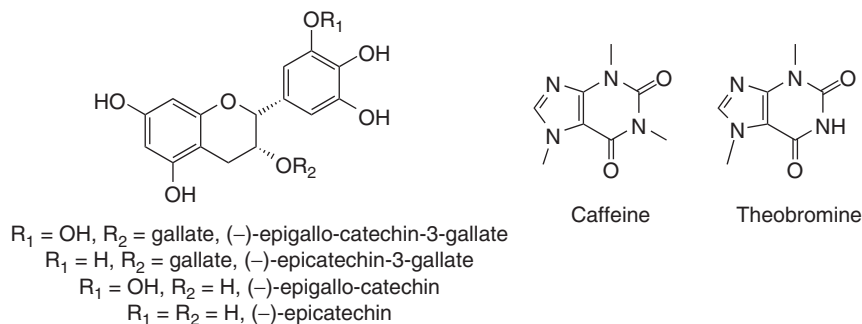
Figure 33 Characteristic constituents of green tea.
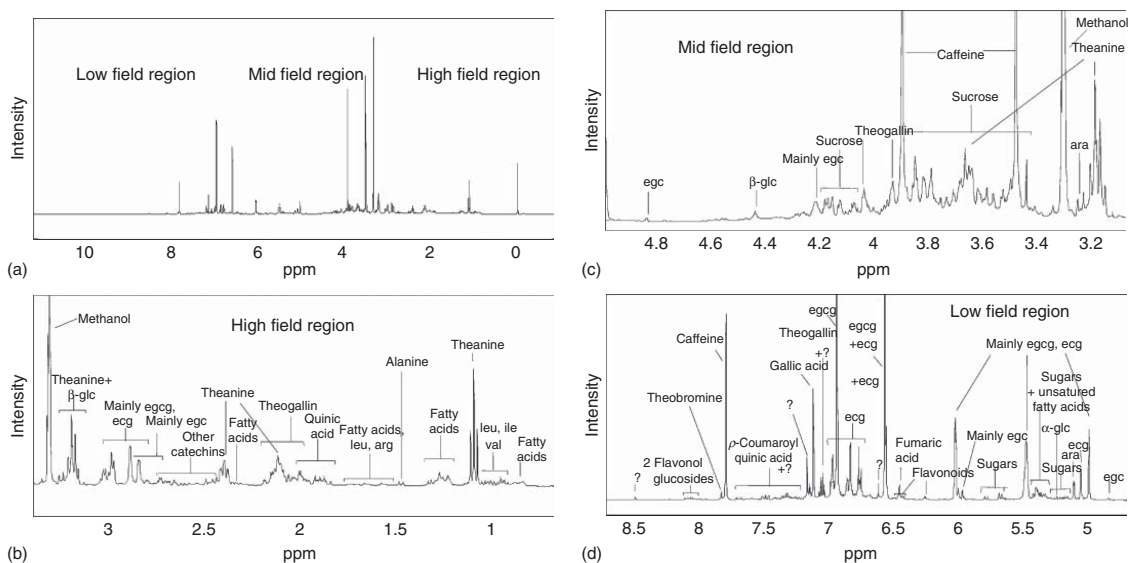
Figure 34 Details of ^{1}H NMR spectrum of a high grade Longjing green tea extract. Keys: leu, leucine; ile, isoleucine; val, valine; arg, arginine; glc, glucose; and ara, 2-*O*-(α-L-arabinopyranosyl)-myo-inositol. (*Source*: Reprinted with permission from Le Gall G, Colquhoun IJ, Defernez M. 2004. Metabolite profiling using ^{1}H NMR spectroscopy for quality assessment of green tea, Camellia sinensis (L.). J Agric Food Chem 52: 692–700. Copyright (2004) American Chemical Society.)

3 CONCLUSIONS

NMR experiments can be considered a simple, widely applicable and rapid analytical instrument in determining the quality and the stability of plant extract constituents without requiring fractionation or isolation steps. The major advantage over conventional analytical methods is the tremendous versatility, not depending on the nature of the extract. NMR experiments can provide a real and complete fingerprint of the extract, as required especially for the innovative ones. Thus, a global vision and a "separation" of all constituents (by the aid of 2D classical experiments and also DOSY) can be achieved even better than conventional chromatographic methods because all the molecules present, also eventually unknown or unexpected compound can be detected.

NMR spectroscopy is a very useful technique, and it is not destructive, requires a very small amount of the sample, and provides signals whose intensity reflects the proportion of nuclei with the same local field or, in other words, with the same chemical environment. The assignment of the resonances of the different functional groups and metabolites is, in general, obtained by comparing NMR data with references or by structure elucidation using 2D NMR.

NMR spectroscopy and, in particular, proton NMR is an effective technique for quantitative constituent analysis in plant extracts. Advantages are the nondestructive nature, and the lack of authentic and identical references, using simply a single, universal reference standard, with a direct proportionality of the integrated resonance intensity and concentration of nuclei giving the resonance. qHNMR possesses the required accuracy and precision to readily become a routine quantitative tool, being the only crucial prerequisite the application of appropriate data acquisition parameters and similarly appropriate postacquisition processing treatment (quantitative conditions). Besides the reason of the cost of expensive equipment, the other important limitation is that the determined signal by NMR should be well separated from the others.

NMR has the disadvantage of inherently low sensitivity [LOD (mol): 10^{-9} to 10^{-11}] relative to other spectroscopic and chromatographic methods, such as fluorescence [LOD (mol): 10^{-18} to 10^{-23}] and mass spectrometry [LOD (mol): 10^{-13} to 10^{-21}], but in the same range of UV–vis absorbance [LOD (mol): 10^{-13} to 10^{-16}] and adequate for quality control of plant extracts.

Finally, chemometric analysis of NMR data can sort data sets into categories. Typically, the starting point is a PCA of the digitized spectrum and this in itself may be sufficient to divide the sample set into a number of categories. Subsequently, it may be informative to investigate the variables that are most important in discriminating between the samples and this leads back to the NMR signals and the metabolites that they represent. This approach has the great merit of avoiding the often time-consuming process of signal assignment before it is necessary, focusing the attention on those parts of the spectrum that are most relevant to the question being addressed. Successfully studies are related to the chemometric analysis of plant extracts for evaluation of quality control, authentication, in determining geographical origin, and for detecting adulteration of extracts or herbal drugs.

REFERENCES

Bailey, N. J. C., Sampson, J., Hylands, P. J., *et al.* (2002) *Planta Med.*, **68**, 734–738.

Bailey, N. J. C., Wang, Y., Sampsonc, J., *et al.* (2004) *J. Pharm. Biomed. Anal.*, **35**, 117–126.

Barjat, H., Morris, G. A. and Swanson, A. G. (1998) *J. Magn. Reson.*, **131**, 131–138.

van Beek, T. A. (2002) *J. Chromatogr. A*, **967**, 21–25.

van Beek, T. A., van Veldhuizen, A., Lelyveld, G. P., *et al.* (1993) *Phytochem. Anal.*, **4**, 261–268.

Bilia, A. R., Bergonzi, M. C., Mazzi, G., *et al.* (2001a) *J. Agric. Food Chem.*, **49**, 2115–2124.

Bilia, A. R., Bergonzi, M. C., Morgenni, F., *et al.* (2001b) *Int. J. Pharm.*, **213**, 199–208.

Bilia, A. R., Bergonzi, M. C., Mazzi, G., *et al.* (2002a) *J. Agric. Food Chem.*, **50**, 5016–5025.

Bilia, A. R., Bergonzi, M. C., Mazzi, G., *et al.* (2002b) *J. Pharm. Biomed. Anal.*, **30**, 321–330.

Bilia, A. R., Bergonzi, M. C., Morris, G. A., *et al.* (2002c) *J. Pharm. Sci.*, **91**, 2265–70.

Blumenthal, M., Busse, W. R., Goldberg, A., *et al.* (1998) *The Complete German Commission E Monographs-Therapeutic Guide to Herbal Medicines*, Austin, TX, American Botanical Council.

Choi, Y. H., Choi, H. K., Hazekamp, A., *et al.* (2003) *Chem. Pharm. Bull.*, **51**, 158–161.

Choi, Y. H., Choi, H.-K., Peltenburg-Looman, A. M. G., *et al.* (2004a) *Phytochem. Anal.*, **15**, 325–330.

Choi, Y. H., Kim, H. K., Hazekamp, A., *et al.* (2004b) *J. Nat. Prod.*, **67**, 953–957.

Choi, Y. H., Sertic, S., Kim, H. K., *et al*. (2005) *Agric. Food Chem.*, **53**, 1237–1245.

Cordella, C., Moussa, I., Martel, A. C., Sbirrazzuoli, N., Lizzani-Cuvelier, L. (2002) Recent developments in food characterization and adulteration detection: Technique-oriented perspectives. *J Agric Food Chem* **50**: 1751-1764:

Defernez, M. and Colquhoun, I. J. (2003) *Phytochemistry*, **62**, 1009–1017.

Dunn, W. B. and Ellis, D. I. (2005) *Trends Anal. Chem.*, **24**, 285–294.

Fan, T. W. M. (1996) *Prog. Nucl. Magn. Reson. Spectrosc.*, **28**, 161–219.

Formäcek, V. and Kubeczka, Ê.-H. (1982) ^{13}C NMR analysis of essential oils, in *Aromatic Plants: Basic and Applied Aspects*, eds. Í. Margaris, A. Koedam, and D. Vokou,, Martinus Nijhoff, London.

Frédérich, M., Choi, Y. H. and Verpoorte, R. (2003) *Planta Med.*, **69**, 1169–1171.

Frédérich, M., Choi, Y. H., Angenot, L., *et al*. (2004) *Phytochemistry*, **65**, 1993–2001.

Hall, R. D. (2006) *New Phytol.*, **169**, 453–68.

Hazekamp, A., Choi, Y. H. and Verpoorte, R. (2004) *Chem. Pharm. Bull.*, **52**, 718–721.

Holmes, E., Tang, H., Wang, Y., *et al*. (2006) *Planta Med.*, **72**, 771–785.

Johnels, D., Edlund, U., Grahn, H., *et al*. (1983) *J. Chem. Soc. Perkin. Trans.*, **2**, 863–871.

Kim, H. K. , Choi, Y. H., Erkelens, C., *et al*. (2005) *Chem. Pharm. Bull.*, **53**, 105–109.

Kim, H. K. and Verpoorte, R. (2010) *Phytochem. Anal.*, **21**, 4–13.

Kim, H. K., Choi, Y. H., Chang, W. T., *et al*. (2003) *Chem. Pharm. Bull.*, **51**, 1382–1385.

Krishnan, P., Kruger, N. J. and Ratcliffe, R. G. (2004) *J. Exp. Bot.*, **56**, 255–265.

Le Gall, G., Colquhoun, I. J. and Defernez, M. (2004) *J. Agric. Food Chem.*, **52**, 692–700.

León, A., Chávez, M. I. and Delgado, G. (2011) *Magn. Reson. Chem.*, **49**, 469–476.

Li, C. Y., Lin, C. H., Wu, C. C., *et al*. (2004) *J. Agric. Food Chem.*, **52**, 3721–3725.

Lindon, J. C. and Nicholson, J. K. (1997) *Trends Anal. Chem.*, **16**, 190–200.

Lindon, J. C. and Nicholson, J. K. (2008) *Trends Anal. Chem.*, **27**, 194–204.

Marston, A. (2007) *Phytochemistry*, **68**, 2785–2797.

McLachlan, A. S., Richards, J. J., Bilia, A. R., *et al*. (2009) *Magn. Reson. Chem.*, **47**, 1081–1085.

Palomino-Schätzlein, M., Escrig, P. V., Boira, H., *et al*. (2011) *J. Agric. Food Chem.*, **59**, 11407–11416.

Pauli, G. F., Jaki, B. U. and Lankin, D. C. (2005) *J. Nat. Prod.*, **68**, 133–149.

Pieri, V. and Stuppner, H. (2011) *Planta Med.*, **77**, 1756–1758.

Pieri, V., Belancic, A., Morales, S., *et al*. (2011) *J. Agric. Food Chem.*, **59**, 4378–4384.

Pieri, V., Sturm, S., Seger, C., *et al*. (2012) *Metabolomics*, **8**, 335–346.

Reynolds, W. F. and Enríquez, R. G. (2002) *J. Nat. Prod.*, **65**, 221–244.

Roos, G., Röseler, C., Berger-Büter, K., *et al*. (2004) *Planta Med.*, **70**, 771–777.

Verpoorte, R., Choi, Y. H. and Kim, H. K. (2007) *Phytochem. Rev.*, **6**, 3–14.

Verpoorte, R., Choi, Y. H., Mustafa, N. R., *et al*. (2008) *Phytochem. Rev.*, **7**, 525–537.

Wang, Y., Tang, H., Nicholson, J. K., *et al*. (2004) *Planta Med.*, **70**, 250–255.

NMR of Small Molecules*

Christoph Seger

University of Innsbruck, Institute of Pharmacy/Pharmacognosy, CCB—Centrum of Chemistry and Biomedicine, Innsbruck, Austria and *Institute of Medical and Chemical Laboratory Diagnostics (ZIMCL), University Hospital/Landeskrankenhaus Innsbruck, Innsbruck, Austria*

We are optimistic enough to think that NMR will be an evergreen for at least a few more years, that in the future it will continue to surprise the chemical community with new fundamental developments, new instrumentation, and new applications. We expect from the past and current trends in this remarkable field that it will continue to help solve an ever growing number of increasingly diverse chemical problems.

(Jonas and Gutowsky, 1980)

1 INTRODUCTION

This overview should provide basic guidance to understand nuclear magnetic resonance (NMR) spectroscopy. It is intended to provide the reader an introduction to the field and also to the following chapters of this handbook (**NMR as Analytical Tool for Crude Plant Extracts**, and **On-line and At-line LC-NMR and Related Micro NMR Methods**). For a deeper insight into NMR, that is, its physical basics including the quantum-mechanical model of NMR experiments and the technical realization of NMR measurements, the reader is referred to the abundant textbook and review literature (Abragam, 1961;

Berger and Braun, 2004; Claridge, 2009; Croasmun and Carlson, 1994; Ernst, Bodenhausen, and Wokaun, 1990; Freeman, 1998; Keeler, 2010; Martin and Zektzer, 1988; Morris and Emsley, 2010).

Since more than four decades (Voelter, 1978), NMR spectroscopy is definitely a key instrumental analysis technique in natural product chemistry, especially if small (<2000 Da) organic molecules are involved. Generally, identification and characterization of such analytical entities is a pivotal prerequisite in many analysis approaches used in natural product research setups. The structural identification of an analyte is a must for quantitative assays as well as for the characterization of discriminators in metabolic profiling approaches and fingerprint type of assays.

Too often it is taken granted that the structural identification of an analyte can be put into the hand of a third party, e.g., by purchasing certified materials with declared purities and guaranteed content. It should however not be overlooked that even in the case of rather simple organic molecules as the tyrosine kinase inhibitor (TKI) bosutinib, chemical synthesis can be such challenging that two regio-isomers with pronounced differences in their NMR spectra are produced and marketed more the less unnoticed as identical substances (Levinson and Boxer, 2012).

In natural product structure analysis efforts, with more complex molecular scaffolds to cover, incorrect identification of an analyte is the constant companion of confirmative natural product synthesis frequently

---

*This chapter is dedicated to the memory of Professor Otmar Hofer (1942–2009).

unveiling such errors (Nicolaou and Snyder, 2005; Suyama, Gerwick, and McPhail, 2011). Consequently, a state-of-the-art natural product laboratory devoted to analytical chemistry research approaches involving purified small organic molecules from natural sources, as novel secondary metabolites in uncharted species or hit candidates in bioactivity screens, should be definitively in the position to unequivocally characterize these analytes on their own (Berger and Sicker, 2009). The two major key technologies to be used in this context are NMR spectroscopy providing carbon–carbon and carbon–proton connectivity data to establish the molecular scaffold of an investigated entity and chromatography hyphenated to high resolution mass spectrometry providing molecular mass information resulting in one or few molecular formulae of the target. If crystals can be obtained, X-ray-based structure analysis may pose an alternative and complementary approach to gain information about the molecular skeleton of an investigated analyte. Provided that at least one well-scattering crystal can be obtained, which is not guaranteed for many secondary natural products (especially if large flexible structure elements are present), three-dimensional electron density maps allow the calculation of atom positions and bonds connecting these electron-rich centers to form a molecule. Care must be taken that from X-ray diffraction data analysis, the exact spatial position of hydrogen atoms (protons) usually cannot be deduced. In addition, the molecule conformation in the crystal does not reflect the liquid state average conformation obtained from NMR data. Hence, if needed, liquid state conformation analysis by NMR spectroscopy should accompany solid-state crystal analysis-based structure elucidation protocols.

Spectroscopic methods do aid the structure elucidation process, with infrared (IR) spectroscopy contributing mostly to the identification of functional groups and ultraviolet–visible (UV–vis) spectroscopy data supporting the identification of aromatic ring systems. Finally, if chirality centers are present and the absolute configuration of an analyte has to be proven beyond doubt, chiroptical methods such as optical rotation dispersion (ORD) or circular dichroism (CD) analysis (e.g., combined with quantum chemical CD calculations) are aiding the structure elucidation process (Allenmark, 2000; Bringmann *et al.*, 2008).

2 HISTORIC DIMENSION

The past generations have seen many scientific revolutions and turning points based on breakthrough discoveries and technological developments profoundly changing the understanding of our microscopic and macroscopic environments. While the technological concept "polymerase chain reaction" (PCR) (Mulis, 1994) changed our understanding of the genetic control of life and enabled us to dereplicate our genome and to create tailored life forms as well as tailored protein-based drugs, analytical approaches based on the technology platform "X-ray crystal structure analysis" set the basis for understanding the mechanism and structure of proteins – the "machines of life" to be targeted by tailored drugs – as proven beyond doubt in the case of the highly selective and potent antileukemia drug imatinib (Capdeville *et al.*, 2002). In this canon, modern NMR spectroscopy (Ernst, 1992; Wüthrich, 2003) can be seen as technology allowing a detailed insight into the scaffold and 3D structures of metabolites not only being the products of gene expression and enzyme-based biosynthesis but often triggering biological reactions in complex environments, for example, as semiochemicals (pheromones, kairomones, etc.), hormones (endocrine communication), or chemical defense entities (antibiotics, toxins, etc.).

3 THE INFORMATION CONTENT OF THE NMR SIGNAL

NMR signals are – compared to other techniques such as mass spectrometry or UV–vis spectroscopy – rather information rich as each signal originates from a specific nucleus in a complex molecule. The signal features are a function of their position in this molecule; their analysis allows reconstructing the molecular scaffold. As for any other spectroscopic technique, we have to discuss its position in the spectrum, its shape and fine structure, and its relative and absolute intensities.

The position of an NMR signal in an NMR spectrum is called *chemical shift* (δ), as it is mainly influenced by the electron density of its chemical environment. Neighboring atoms in the molecule – regardless if they are connected via chemical bonds or if they are only spatial close do influence the resonance frequency of the nuclear

spin. It is measured relatively to a reference substance – in organic solvents usually tetramethylsilane (TMS), in water usually trimethylsilylpropoinate (TSP), or 4,4-dimethyl-4-silapentane-1-sulfonic (DSS) acid. The chemical shift is defined as the resonance frequency difference of a nucleus relative to the resonance frequency of the methylsilyl moiety of the reference compound (e.g., 600 MHz for ^{1}H and 150 MHz for ^{13}C in a 14.1 T magnet). To be independent of the field strengths of the magnet, this difference (usually some hundred to some thousand Hertz) is scaled by the operating frequency of the magnet (e.g., 600 MHz for ^{1}H in a 14.1-T magnet) and expressed in parts per million for the sake of clearness in the presentation of the shift values. The chemical shift is closely related to the electron density/electronegativity (EN) of the chemical environment of an observed atom. More electronegative atoms reduce the electron density of the adjacent nucleus (e.g., oxygen as neighbor to carbon in CH_3OH), hence reducing the shielding effect (=deshielding) the atomic shell provides against external fields. Consequently, at a fixed magnetic field, in such atoms, a higher resonance frequency is needed to record an NMR signal. For example, the monosubstituted halomethanes CH_3Cl, CH_3Br, and CH_3I show chemical shifts increasing with the EN of the heteroatom: CH_3Br EN = 2.8, δ_H = 2.7 ppm; CH_3Cl EN = 3.0, δ_H = 3.1 ppm; and CH_3I EN = 4.0, δ_H = 4.3 ppm.

In addition to the ΔEN caused electron density anisotropy, secondary induced magnetic fields can contribute to the overall local electron density. Such fields are caused by the external magnetic field inducing a current in π-electron systems. Hence, in structure elements bearing π-electrons, for example, an aromatic ring system, an induced magnetic field is present that adds to the external field, such that a higher resonance frequency is needed for NMR signal detection. For example, a CH proton on an sp^2 hybridized carbon center with delocalized π-electron system resonates at δ_H = 7.2 ppm (benzene), whereas a CH proton on an sp^2 hybridized carbon center without a delocalized π-electron system resonates at δ_H = 5.6 ppm (cyclohexene). For comparison, CH_2 protons on an sp^3 hybridized carbon center resonate at δ_H = 1.4 ppm (cyclohexane). For the sake of completeness, it should not be overlooked that the opposite effect is found in the center of the ring – here lower resonance frequencies are found.

Under isotropic conditions in an undisturbed homogeneous magnetic field, NMR signals show Lorentzian line shapes. Any deviation from this symmetrical shape is a sign for a disturbed environment of the investigated nucleus. Errors in the experimental design as inhomogeneous magnetic fields (e.g., imperfect shim, imperfect NMR tubes, etc.) may cause distorted line shapes (Chmurny and Hoult, 1990) or are, in the case of paramagnetic impurities (e.g., from the sample matrix, an impure solvent or an inefficient purification protocol), associated with peak broadening because of a quick loss of the NMR signal in its observation period. Dynamic processes (e.g., exchange of acidic protons, conformation switches, etc.) may also cause NMR lines to be more broad humps than well-defined narrow lines. In contrast to experimental errors, these signal shapes are temperature dependent and may be restricted to only part of the investigated molecule (e.g., if a keto-enol tautomerism is possible).

The fine structure of an NMR signal – the splitting of a resonance signal in two or more lines forming the often complex "coupling pattern" – arises from the interaction of NMR active nuclei and is mediated through chemical bonds. The number of lines (multiplicity) is – in the first order – a function of the number of chemical equivalent neighboring atoms. Hence the chemical equivalent ^{1}H atoms of a methyl group in the close vicinity of a ^{1}H atom of a methine moiety split this resonance signal to a four-line group (quartet) with the relative intensities 1 : 3 : 3 : 1 and the methine ^{1}H atom splits the ^{1}H atoms of a methyl group into a two-line group (doublet) with the relative line intensities 1 : 1. The size of the resonance difference between these lines, the coupling constant J (expressed in hertz), is independent from the applied magnetic field strength and equivalent in both signal groups of the interacting nuclei. The size of the "J-coupling" constant depends on the nuclei involved (e.g., $J_{CH} > J_{HH}$), the distance between coupling partners and – if geometrically constrained systems (e.g., saturated ring systems in steroids) are present – the dihedral angle between the interacting nuclei. The relationship between dihedral angel and coupling constant is described by the Karplus equation with a minimal coupling constant at an angle of 90° and maxima at 0° (gauche-configuration) and 180° (anti-configuration) (Minch, 1994). Consequently, if analyzed in detail, coupling information allows a deep insight into the connectivity network of atoms making up a molecule.

In contrast to any other spectroscopic method/mass spectrometry, the magnitude of an NMR signal is directly proportional to the number of nuclei causing the signal (the theoretical signal ratio of $CH : CH_2 : CH_3$ is $1 : 2 : 3$), if some experimental framework conditions such as a sufficient signal to noise ratio, a good Lorenzian line shape, and a sufficient high data density (digital resolution) are met. In addition, saturation effects due to too short relaxation times between single experiments must be avoided as well as peak intensity deviations caused by nuclear Overhauser effect (NOE) buildup in experiments using decoupling.

If standard experimental conditions are used in ^1H NMR, the impact of saturation effects, which affects slower relaxing nuclei (e.g., an isolated aromatic proton in rutin) more than fast relaxing ones (e.g., the methyl group in rutin) on the accuracy of the measured signal integral, is of limited significance, although in the given methin / methyl group example the signal ratio might be altered from the theoretical $1 : 3$ to about $0.8 : 3$. Hence, if exact integral information is needed (e.g., to estimate the number of protons in a very crowded regions of a spectrum, in quantitative metabonomics applications, or in quantitative NMR), care must be taken that complete relaxation of the observed nuclei is made possible by increasing the delay between excitation pulses to five times the longest T_1 relaxation time in the sample (Pauli, Jaki, and Lankin, 2007). In the case of standard ^{13}C NMR, the NOE caused by the ^1H decoupling makes meaningful signal integration generally impossible, although this effect is superimposed by saturation effects caused by very long T_1 relaxation times in the minute time range (consequently, quaternary carbon atoms show very weak signals compared to proton-bearing carbon atoms). Several approaches can be pursued to allow analyte quantification from ^{13}C NMR data (Pieters and Vlietinck, 1989). Reducing the flip angle of the excitation pulse from 90° to 20° – 30° (Ernst angle relationship) reduces the relaxation delay needed for $5 \times T_1$ by orders of magnitude (Rabenstein, 1984; Traficante and Steward, 1994). Excluding carbon atoms with extremely long relaxation times from the quantitative analysis is also a feasible approach. Finally, the influence of the decoupling NOE can be reduced by liming decoupling to the acquisition time that is usually significantly shorter than the relaxation delay. Finally, paramagnetic shift reagents such as $Cr(acac)_3$ can be added to the sample to enhance the spin–lattice

relaxation (Berger and Braun, 2004). Major drawback of this approach is that the sample is usually lost due to the contamination with the complexing agent.

4   NMR EXPERIMENTS

Owing to different modes of acquisition, NMR experiments can be classified by their dimensionality – differentiating 1D (one-dimensional) and 2D (two-dimensional) NMR spectra (Figures 1 and 2) (Brenton and Reynolds, 2013; Bross-Walch *et al.*, 2005; Reynolds and Enriquez, 2002), although in biomolecular applications, 3D NMR is also frequently applied (Wüthrich, 2003). Within the 2D NMR that allows plotting spin–spin correlations in a 2D contour plot, homonuclear and heteronuclear techniques are discriminated (Table 1). While homonuclear NMR spectra are devoted to deduce correlation signals between atoms of the identical isotope (e.g., ^1H–^1H or ^{13}C–^{13}C), heteronuclear NMR spectra allow to track correlation signals between different isotope types with the bioanalytically most important ones being ^1H–^{13}C or ^1H–^{15}N shift correlations. The mode of magnetization transfer is another possibility to classify NMR experiments. If scalar coupling pathways are used as selection principle, strictly spin–spin coupling through chemical bonds is observed. Consequently, homonuclear or heteronuclear spin networks (=atom networks) can be established. Here several experiments are available (for details, see further below), giving a detailed
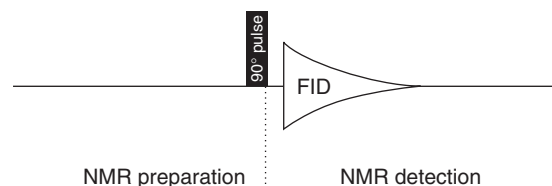
Figure 1 Generalized technical sketch of a one-dimensional (1D) NMR experiment. In a preparation phase, the spin system is set into a defined state. Unwanted signals characteristics, as, coupling to heteroatoms, or unwanted signals, as, from macromolecules or solvent in the sample, can be eradicated by decoupling, spin–echo sequences or presaturation, respectively. Excitation of the nuclei under investigation is carried out by a rectangular (hard) pulse covering the desired chemical shift range. The "free induction decay" (FID) is recorded in the detection period of the experiment followed by a waiting time allowing to reestablish the spin equilibrium.
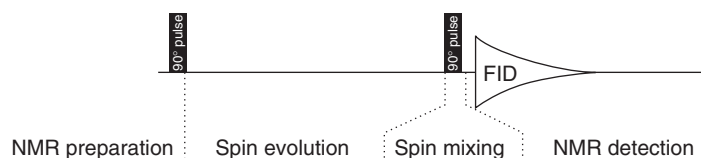
Figure 2 Generalized technical sketch of a two-dimensional (2D) NMR experiment. In addition to the preparation and detection period, an additional spin evolution and spin mixing time is added. Magnetization is labeled in the evolution time with the chemical shift of the first nucleus. In the mixing period, this information is transferred to a second nucleus via scalar coupling or dipolar interactions. Recorded NMR data (FID) carries the shift information of the first nucleus and is modulated with the shift information of the second nucleus. 2D Fourier transformation allows decoding the involved resonance frequencies.

insight into the coupling network of an observed nucleus – its "atomic neighborhood" so to say-within a molecule. If dipolar coupling pathways are selected, spin–spin coupling through the space is observed; spatially close atoms will correlate even if they are separated by many chemical bonds. Hence, such experiments allow remodeling a 2D spin correlation network to a 3D model of a molecule or to connect spin networks separated by NMR inactive heteroatoms (e.g., oxygen functions).

A prerequisite of "classical" 2D NMR experiments is the excitation of the complete chemical shift range expected in a molecule by applying a "hard pulse." If only a certain peak or peak group has to be observed, for example, because the complete structure of an analyte is already known and only the stereochemical assignment on one chiral carbon center remains unclear or if the homonuclear coupling partner of an certain ^{1}H atom is hidden under several other ^{1}H signals, a selective excitation of the resonance frequency of the investigated nucleus can be achieved by applying a "soft pulse." Consequently, the 2D spectrum is reduced to a selective 1D spectrum, an approach which may save a lot of measurement time and even – especially in ^{1}H–^{1}H shift correlations – may give insight into the proton spin system experimentally not achievable by other means.

Since NMR experiments in natural product analysis are mostly devoted to unravel the C–H coupling network, the following paragraphs will be restricted to the major techniques utilized in this context. For a comprehensive overview on small molecule NMR experiments, see the literature (Berger and Braun, 2004) and pulse program catalogs of the NMR instrument vendors.

4.1 1D-NMR Experiments

4.1.1 ^{1}H 1D-NMR

The most sensitive and therefore fastest NMR experiment is the simple ^{1}H NMR resonance scan. In small molecule research, it is the basis of any NMR-based investigation. Usually, chemical shift information and a rough estimate of the number of ^{1}H nuclei in the molecule as well as first clues about the coupling partners of isolated signal groups can be obtained. In addition, the absence or presence of structural elements (e.g., aromatic moieties, methoxy groups, methyl groups without coupling partner, and glycosides), which might be (in combination with knowledge of the secondary metabolite composition of the investigated species) of diagnostic value for the identification of the analyte compound class (e.g., several methyl groups without coupling partner make the presence of a terpenoid likely). For the more detailed insight into the ^{1}H–^{1}H coupling network, homonuclear 2D shift correlation spectra (see below) have to be aquired.

4.1.2 ^{13}C 1D-NMR

While recording of ^{1}H NMR data is a question of minutes, ^{13}C NMR spectra acquisition is quite time consuming. This is not only because the inherent sensitivity of ^{13}C is remarkably lower (1.6% of ^{1}H), but also because the natural abundance of ^{13}C is a minute fraction (1.1%) of the ^{1}H abundance. Hence, measuring ^{13}C means measuring less populated nuclei with lower sensitivity. Consequently, the direct observation of ^{13}C NMR resonances is, considering the actual sensitivity of 0.017% (1/5700 of the ^{1}H sensitivity), cumbersome. As

Table 1   Overview of frequently used 2D-NMR experiments for the structural characterization of small organic molecules.

Abbreviation	Nuclei	Coupling type	Correlation range	Remarks	Spectrometer time typically needed[a]
DQF-COSY	^{1}H	Bond-mediated scalar coupling	Usually over 2–4 bonds	COSY spectra are often crowded and hard to interpret. Cross-peak strengths can vary, depending on the ^{1}H/^{1}H coupling constant	<1 h
TOCSY	^{1}H	Relayed bond-mediated scalar coupling	All ^{1}H nuclei within a scalar coupling network	TOCSY spectra are often crowded and hard to interpret. Coupling propagation can be limited if only small three-bond ^{1}H/^{1}H coupling constants are present	<1 h
NOESY	^{1}H	Dipolar coupling through space	Typically <5 Å.	NOESY works best with small molecules. It is replaced with ROESY for higher molecular weights	Several hours
ROESY			Absent correlations must not be understood as distances exceeding this threshold		
HSQC	^{1}H, ^{13}C	Bond-mediated scalar coupling	Strictly one ^{1}H–^{13}C bond	Correlates the ^{1}H network with the ^{13}C scaffold. If multiplicity editing is used in the ^{13}C, signal loss may occur. HSQC spectra are often of "cleaner" appearance than HMQC spectra (Reynolds and Enriquez, 2002)	Several hours
HMQC					
HMBC	^{1}H, ^{13}C	Bond-mediated scalar coupling	Usually over 2–4 bonds	Correlates the ^{1}H network with the ^{13}C scaffold. Signal strength does not correlate with bond number between correlated signals	Over night
HSQC-TOCSY	^{1}H, ^{13}C	Bond-mediated scalar coupling	All ^{1}H and ^{13}C nuclei within a scalar coupling network	Combines the ^{1}H/^{13}C correlation with the ^{1}H coupling network. Coupling propagation can be limited if only small three-bond ^{1}H/^{1}H coupling constants are present	Over night
INADEQUATE	^{13}C, ^{13}C	Bond-mediated scalar coupling	One bond	Correlates the ^{13}C scaffold	1–2 days
ADEQUATE					

[a] For any NMR experiment, the data acquisition time depends on the concentration and complexity of the analyte as well as on and the field strengths of the employed magnet. Given figures are approximations for about 3 mg of a 500-Da compound (∼10 mM solution in 600 µl NMR solvent) measured on a 600-MHz spectrometer in conventional 5-mm NMR tubes.

a workaround the distortionless enhancement by polarization transfer (DEPT) and insensitive nuclei enhanced by polarization transfer (INEPT) pulse sequence was introduced. Here – as in all "inverse detected" heteronuclear 2D-NMR experiments (see below) – ^{13}C shift information is recorded through the more sensitive directly attached ^{1}H atoms. The experiment starts with the proton excitation followed by the magnetization transfer (polarization transfer) onto the carbon. On the basis of the rather uniform

$^2J_{CH}$ coupling constant (∼130 Hz), "spectral editing" according to the number of protons attached on a carbon center can be performed. Three DEPT spectra for CH (DEPT 90), CH/CH$_3$/CH$_2$ (DEPT 45), and CH/CH$_3$/CH$_2$ (inverse sign) (DEPT 135) can be obtained, allowing a discrimination of methine, methylene, and methyl moieties. Carbon centers not bearing protons are however not detectable with DEPT experiments, making alternative ^{13}C 1D NMR experiments as the attached proton test (APT) or the

J-modulated (J-MOD) spin–echo necessary. These two pulse sequences are directly recording ^{13}C NMR data. Hence, they are inherently less sensitive than the DEPT experiment as they do not use the sensitivity gain by the polarization transfer step. On the other hand, both experiments show resonances for quaternary carbon centers, an advantage not to be underestimated. The APT pulse sequence is understood as a variant of the J-MOD experiment giving slightly stronger quaternary carbon signals using a smaller initial excitation pulse to reduce saturation: multiplicity editing is also based on the $^2J_{CH}$ coupling constant, in contrast to the DEPT experiments, only one experiment has to be performed with methine and methyl signals having the opposite phase (and therefore sign) than methylene and quaternary resonances. Care must be taken that for all spectral editing pulse sequences that rely on building blocks using the $^2J_{CH}$ coupling constant, substructures with $^2J_{CH}$ deviating from the usually used ~130 Hz (e.g., polyine CH with ~250 Hz, strained systems with ~170 Hz) may give weaker signals, no signals, or signals with wrong signs.

4.2   2D-NMR Experiments

4.2.1   Homonuclear ^1H 2D-NMR Experiments

The ^1H coupling network of natural products can be complex. Structural elements of high similarity, for example, several glycoside moieties or several saturated ring systems as in triterpenoides are common. Hence, the mere analysis of the coupling pattern of single proton signals by extracting and comparing coupling constants of different multiplets, which might be feasible in rather low molecular weight compounds with isolated spin systems bearing only some protons (e.g., aromatic compound classes as some alkaloids, flavonoids, and coumarins) is not always a feasible dereplication strategy. Here, the homonuclear ^1H–^1H correlation experiment COSY (correlated spectroscopy) is a fast and valuable tool to overcome spectral overlap and the ambiguity of coupling constant-based correlation partner analysis. In contrast to heteronuclear correlation experiments, the H,H-COSY, usually experimentally realized as the significantly artifact reduced DQF-COSY (double quantum filtered correlated spectroscopy), can be obtained within a couple of minutes (Table 1).

Interpretation of COSY spectra is straightforward, as correlation signals (cross-peaks) are found symmetrically above and below the diagonal because of the different chemical shifts of the involved nuclei. The intensity of the cross-peak correlates with the magnitude of the coupling constant – geminal protons give more intense peaks ($J_{HH} \sim 12$–14 Hz) than vicinal ones, and then syn- or anticonfigured vicinal protons (e.g., the axial protons in the –CH_2CH_2– or –CH_2CH– elements of triterpenes with $J_{HH} \sim 10$–12 Hz) give stronger and hence easier detectable correlation signals than equatorial oriented protons.

If an experiment is performed under "spin lock" conditions, that is, the application of a weak additional radiofrequency keeping the nuclei in a state enabling the buildup of spin–spin correlations over the whole coupling network and not – as in DQF-COSY experiments – only between "next neighbors," additional spin–spin correlations can be detected (Figure 3). Consequently, in this experiment, the TOCSY (total correlation spectroscopy), also known as HOHAHA (homonuclear Hartmann–Hahn) experiment, a more complex cross-peak pattern is found. In addition to DQF-COSY detectable direct correlations, indirect "relayed" correlations are found. Usually, the complete spin system defined as the sum of all protons in a substructure connected by scalar coupling can be detected in a TOCSY experiment. It finds its greatest appreciation in the structural characterization of compounds with repetitive structural units with isolated proton spin systems, for example, glycoside moieties in oligosaccharides as realized in saponins or amino acids in a peptide chain.

Alternatively to scalar coupling pathways, magnetization transfer from one spin to the other through space (cross-relaxation) can be measured. Experimentally, this transfer is assessed by perturbing the equilibrium state of one spin and observing the intensity change on correlating spins (NOE). As the NOE is proportional to the inverse sixth power of the distance between the spins, only spatially close atom pairs can give rise of a correlation signal. 5 Å, the lengths of approximate 4C–H bonds, is usually seen as detection limit of the NOE. Experimentally, the NOE effect can be either measured by a 1D experiment ("steady state NOE" or "transient NOE," both interpreted in the difference mode), or – more often applied by the 2D-NOESY (two-dimensional nuclear Overhauser effect spectroscopy) pulse sequence (Keeler, 2010). Owing
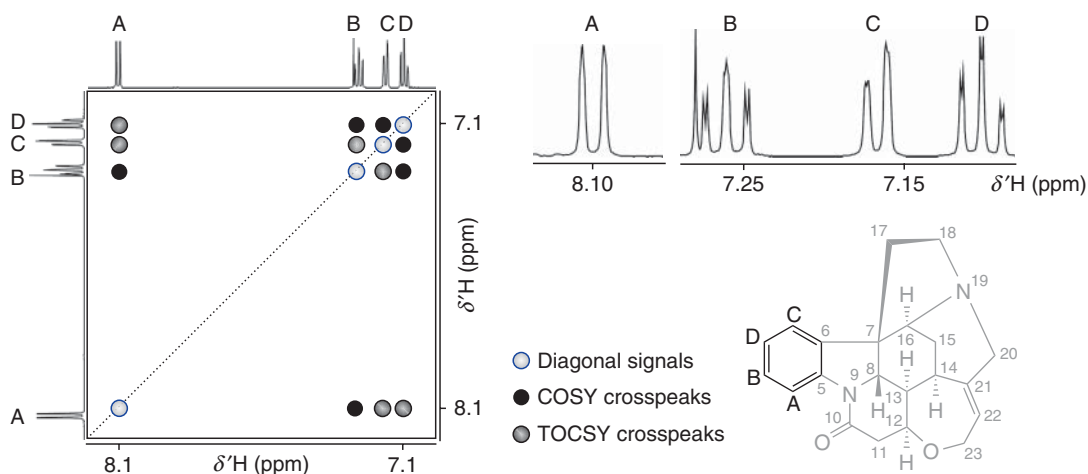
Figure 3 Schematic COSY/TOCSY comparison exemplified on the aromatic substructure of strychnine. In this four-proton spin system, the neighborhood relationship of the *para* substituents H_A and H_C could not be solved from the mere analysis of the proton spectrum because of the similarity of the $^3J_{H,H}$ coupling constants to H_B and H_D. In the COSY spectrum, however, the correlation of H_A with H_B and H_D with H_C is demonstrated by the respective off-diagonal cross peaks (black bullets). In the TOCSY spectrum, each proton correlates with all other protons in the spin system – the complete spin network is presented at each proton resonance (black and gray bullets).

to experimental problems, the "steady state NOE" experiment has almost vanished, and pulse sequences found in the software package of modern NMR spectrometer installations are usually utilizing "transient NOE" building blocks (Berger and Braun, 2004).

NOESY cross-peaks indicate that the two spins involved are spatially close. Application strongholds of NOESY experiments in natural product structure identification efforts are, for example, the elucidation of the relative configuration of saturated ring systems as realized in terpenoids and the discrimination of regio-isomers by proving the spatial closeness of isolated proton coupling networks, which is of special importance in the elucidation of derivatives with one or more aromatic ring or conjugated double bond moieties or – beside heteronuclear long-range coupling experiments (see below) – the use as independent prove of bond elements involving heteroatoms (e.g., amides, glycosides, lactones, epoxides, etc.). It must not be forgotten that absolute configuration determination is not possible with NOESY experiments. Only in combination with the independant absolute configuration elucidation of at least one chiral carbon center [for example by the introduction of a chiral selector into the sample (Hoye, Jeffrey, and Shao, 2007; Wenzel and Chisholm, 2011) or by the application of chiroptical methods (Bringmann *et al.*, 2008; Pescitelli *et al.*, 2009)], the relative configuration of an analyte can be turned into an absolute one.

Spatial closeness of spins holds of course also true for geminal protons in CH_2 moieties or other vicinal structure elements (e.g., CH–CH) giving rise of DQF-COSY signals. Consequently, COSY-like crosspeak pattern in NOESY spectra may occur, making the interpretation of NOE data from a NOESY spectrum cumbersome. In addition, chemical exchange crosspeaks (e.g., between bulk water and an OH group) can be observed in NOESY spectra. Fortunately, such signals have the opposite phase to the NOE cross peaks and the identical phase as the diagonal signals – hence, they cannot be mistaken as "real" NOE-based crosspeaks. If additional COSY–like cross-peaks or cross-peaks from 1D-^{1}H NMR spectral overlap are hampering the structural characterization of an analyte, the 1D NOE experiment is a valuable alternative to 2D measurements. With the aid of selective (soft) excitation pulses, only a small spectral region – usually stemming from a single ^{1}H spin – is investigated. Consequently, the resulting 1D NMR spectra are much faster to obtain and are simpler to interpret than the corresponding 2D NMR spectra.

One major disadvantage of the NOESY experiment is that owing to a change in sign of the cross-relaxation rate from positive to negative, in medium-size molecules, the NOE effect is weak or even zero. With higher magnetic fields used, this range shifts to lower molecular weights. Hence,

as an alternative, the observation of the NOE can be carried out under spin lock conditions as utilized in the TOCSY experiment. Consequently, the resulting NOE experiment going by the acronym ROESY (rotating frame nuclear Overhauser effect spectroscopy) is a valuable alternative to NOESY pulse sequences always to be considered.

4.2.2 Homonuclear ^{13}C 2D-NMR Experiments

It would be of course ideal, if the carbon backbone of an analyte could be traced with a homonuclear carbon shift correlation experiment. Owing to the low natural abundance of ^{13}C, the ^{13}C/^{13}C pairs necessary to establish a correlation signal are hardly present (10^{-4} compared to ^{1}H/^{1}H pairs). Consequently, such experiments are out of reach in average mass limited natural product samples and have not found their way into routine. If however ^{13}C-enriched analytes, for example, from biosynthesis studies are used, this limits are overcome and the incredible natural abundance double quantum transfer experiment (INADEQUATE) and adequate double quantum transfer experiment (ADEQUATE) can be utilized (Adam *et al.*, 2002; Arigoni *et al.*, 1997; Bringmann *et al.*, 2007; Ratcliffe and Shachar-Hill, 2006; Schneider, 2007).

4.2.3 Heteronuclear ^{1}H/^{13}C 2D-NMR Experiments

As in 1D NMR, the low abundance and low sensitivity of ^{13}C are the major obstacles to obtain meaningful C–H correlation data. Consequently, the DEPT/INEPT pulse sequence building blocks successfully used in ^{13}C 1D NMR are key to record ^{1}H/^{13}C 2D-NMR experiments in decent time frames. The sensitivity gain comes from transferring the magnetization to the less sensitive carbons after starting the experiment with proton excitation (polarization transfer). After spin evolution, the magnetization now coded with both the proton and the carbon frequencies is transferred back to the proton for recording the FID (free induction decay). Consequently, these experiments are named "inverse detected" shift correlations in contrast to the more time-consuming "direct detected" shift correlations.

^{1}H/^{13}C–heteronuclear correlation experiments based on chemical bond-mediated magnetization transfer (scalar coupling) are the centerpiece of NMR-based structure elucidation in organic substances (Figure 4). By selecting C–H two-bond correlations in heteronuclear single quantum correlation (HSQC) or heteronuclear multi quantum correlation (HMQC) experiments, the proton scaffold of an analyte investigated by DQF-COSY and TOCSY experiments can be connected with the carbon backbone not so easily accessible by homonuclear correlation experiments (see above). HSQC experiments are available in different variations, an especially attractive one facilitates – similar to the 1D-^{13}C APT experiment – the discrimination of CH/CH$_3$ correlation signals from CH$_2$ correlation signals.

To connect individual C-H pairs correlation data which might lead to ambiguous results because of signal overlap, long-range ^{1}H/^{13}C correlation data derived from experiments as the heteronuclear multi-bond correlation (HMBC) is utilized. This experiment is based on the selection of rather small "long-range" C/H coupling constants, hence it not only picks three- to four-bond correlations in the C/H framework on an analyte but allows to connect spin systems isolated from each other by heteroatoms such as oxygen, nitrogen, or sulfur. Hence, HMBC spectra enable the structure elucidator to assemble molecular building-blocks identified by single-bond correlation experiments – for example, to connect a terpenoid substructure with glycoside moieties to a saponine.

Technically spoken HSQC, HMQC, and HMBC are closely related. In all cases, the experiment starts by the excitation of the protons, transfers magnetization to the carbon atoms (with an "INEPT" pulse block), evolves it (here the experiments do differ), and transfers the magnetization back (with an "INEPT" pulse block) to the more sensitive and almost 100% abundant protons for signal recording. To select the long-range coupling constants in the HMBC, an additional waiting time of $1/2J_{C,H}$ (typical $^{2-3}J_{C,H} = 5-8$ Hz) is introduced into the evolution phase. This makes the HMBC inherently less sensitive, as spin–spin relaxation during this rather long period reduces the detectable spin population significantly.

As the spatial range of HMBC correlations is not well defined in terms of bonds covered and multiple correlation signals of hardly predictable strengths occur over several (usually 2–4) bonds,
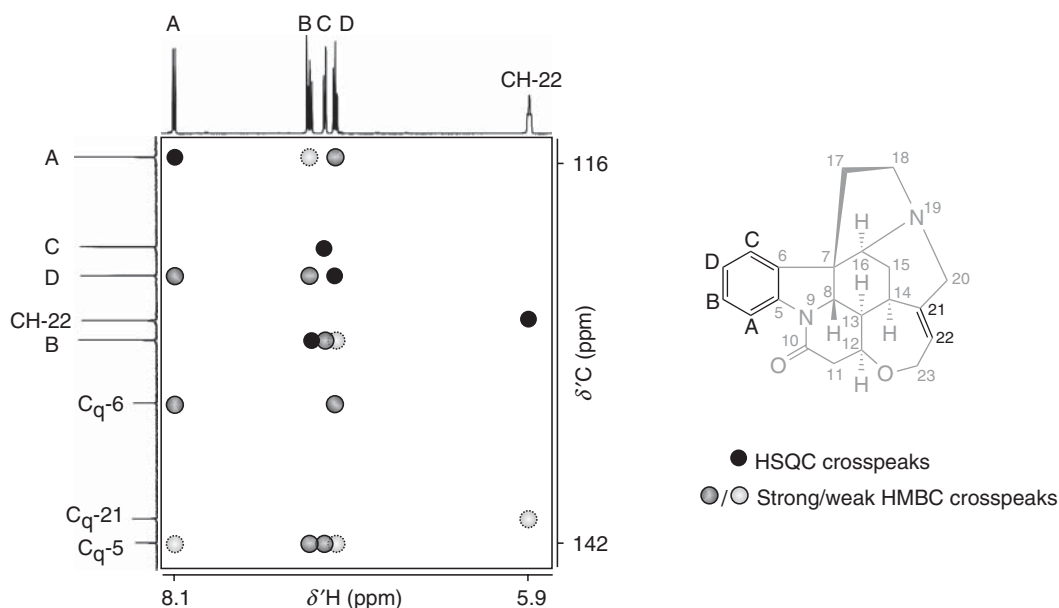
Figure 4 Schematic HSQC/HMBC comparison exemplified on the aromatic substructure and the double bond element of strychnine. In the four-proton spin system of the aromatic ring, the C–H atom pairs and the olefinic C–H element give clear correlation signals (black bullets). The ^{13}C shift of C_A is lower compared to $C_{B,C,D}$ indicating the neighborhood to an electronegative heteroatom. Its presence is confirmed by the HMBC correlation signals between $H_{A–D}$ with a ^{13}C resonance at about $\delta_C = 142$ ppm. The analysis of all HMBC contacts enables to identify quaternary carbon atoms, which are not in the focus of the HSQC experiment. Owing to the occurrence of multiple HMBC correlation signals (i.e., 2–4 for $H_{A–D}$), the interpretation of HMBC spectra can be challenging. In addition, it is frequently observed that owing to the selection mechanism of the HMBC experiment, three-bond correlation (i.e., H_B and H_c to C-5) might give stronger signals than two-bond correlations (i.e., H_A to C-5).

neighbourhood relationships derived from HMBC correlations have to be confirmed by other spectral evidences. Here the heteronuclear two-bond correlation (H2BC) experiment (Nyberg, Duus, and Sorensen, 2005), although less sensitive, is a valuable addition to the HMBC, as strictly two-bond correlations are selected.

The direct detection C–H correlation pendants to HSQC and HMBC are the heteronuclear shift correlation (HETCOR or HETERO-COSY) and the correlation through long-range coupling (COLOC) experiment, respectively. Their advantage over HSQC/HMBC is the high signal resolution in the carbon dimension, which however comes for the prize of low sensitivity, making HETCOR and COLOC usually not utilized with typical natural product samples in the low millimolar concentration range.

A hybrid 2D-NMR experiment extremely useful in natural product chemistry, especially if medium-size analytes with well-separated multi proton spin systems are involved (e.g., triterpenes, saponins, lignans, etc.), is the HSQC-TOCSY, combining the homonuclear ^{1}H/^{1}H TOCSY correlation with the heteronuclear ^{1}H/^{13}C HSQC correlation. The advantage over running TOCSY and HSQC experiments separately is twofold. On the one hand, the ^{1}H/^{1}H spin correlation network identified in the TOCSY can be easily correlated with the ^{13}C resonances of the involved C–H pairs. On the other hand, by changing the scale of the indirect spectral dimension from ^{1}H (as in the TOCSY) to ^{13}C (as in the HSQC), the spectral overlap often hampering the interpretation of TOSCY data is remarkably reduced. Generally, however, the HSCQ-TOCSY is usually not run as replacement but in addition to TOCSY and HSQC spectra for the sake of complete information retrieval. Combining HSQC-TOCSY data interpretation with HMBC derived data is usually sufficient to establish complete C/H coupling networks of natural products.

5 STRUCTURE ELUCIDATION STRATEGIES

Provided that a clean sample with an analyte concentration in the range of 10–50 mM is available, NMR-based structure elucidation of a natural product should always start with a simple ^{1}H NMR experiment. It can be considered a "scout" experiment that will take only some minutes. An experienced operator can deduce from the S/N ratio if the sample concentration is in agreement with the claimed sample amount. In addition, organic impurities such as lipids (e.g., from the plant matrix), solvent residues (e.g., from column chromatography), or plasticizer (i.e., from inappropriate sample storage) are detectable. Inorganic contaminations, for example salts, are often leading to signal broadening, whereas stationary phase material from chromatography shows typical proton resonance frequencies very close to the reference signal of TMS. On the basis of the quality of the "scout" ^{1}H NMR experiment and the analyte concentration found, a set of 2D NMR spectra can be set up (Table 1). If a certain complexity threshold is exceeded, at least a DQF-COSY, an HSQC, and an HMBC spectrum are acquired. TOCSY and HSQC-TOCSY spectra are second-line experiments if complex spin systems are encountered. If region- or stereochemical ambiguities have to be solved, a NOESY or ROESY spectrum is added. Recording of any type of ^{13}C 1D-NMR experiment is only warranted, if HSQC and HMBC spectra show spots of unclear signal assignment or signal overlap, a situation often encountered in triterpenes or polyphenols. Missing coupling networks information as, open valences in the C–H scaffold or unclear presence/absence of a ring closure has to be deduced from other data sources, meaning that in the structure elucidation process, NMR does not stand by its alone (Silverstein, Webster, and Kiemle, 2005). In this context, the value of an exact sum formula derived from high resolution mass spectrometry data must not be underestimated. It provides the number and nature of heteroatoms usually being not NMR active and allows the calculation of double bond equivalents. With this number at hands and the number of double bond elements and ring structures deducible from the given NMR data, the structure elucidator can give an educated guess regarding structural moieties and features not encoded in the information retrieved from the NMR data. This could be the presence of an ether bond instead of two OH functions, a lactone ring closure instead of an acid function and an OH group, an additional ring closure, or a carbonyl function neither detected in the HMBC (e.g., no protons within the four-bond vicinity) nor in the J-MOD spectrum (e.g., analyte concentration too low)).

Modern approaches in computational chemistry including some professional software developers claim that ^{1}H and ^{13}C NMR data can be predicted. Although it is beyond doubt that tremendous progress has been made in recent years, both spectral similarity searches and spectrum prediction have their limitations. While similarity searches strongly depend on the quality of already published data (Robien, 2009), the spectrum prediction has a twofold problem. First of all, all conventional algorithms used are depend on published data – either as learning tool to feed a neuronal network or as basis of structure elements in HOSE code-based prediction approaches. Hence, the data quality of already deposited data strongly influences the quality of the predicted spectra. Generally, if a predicted spectrum of an assumed structure is compared to the experimental data and a good match is found, this finding should be seen as hint toward but not a proof of a structure (Napolitano *et al*., 2013; Stappen *et al*., 2009)

6 CONCLUDING REMARKS

Taken together, NMR spectroscopy is a remarkably versatile tool in the instrumental analysis armamentarium of a natural product laboratory. Interpretation of NMR data accompanies a structure-oriented natural product researcher throughout his or her career. It is not only used to describe novel secondary metabolites on an atomic level but can also be utilized to quantify these analytes in extracts and biofluids. Structure elucidation of organic molecules is an instrumental analysis team effort, with NMR spectroscopy being in the lead. Interpretation of spectral data can be straightforward at first sight; ambiguities in the signal assignment however often spoil first elucidation successes. It is of outmost importance that a proposed structure must be consisted with every bit and piece of experimental data collected. Even small differences in NMR shift values from expected chemical shifts (e.g., as reported in reliable literature) or additional COSY/NOE/HMBC correlation signals which can not be explained within the geometrical limitations of the molecular skeleton might be a clear hint that this structure proposal is not

valid. This has been proven more than once as review literature dramatically tells (Nicolaou and Snyder, 2005, Suyama, Gerwick, and McPhail, 2011). However, once a correct solution is found, the reward is usually high – uncharted biosynthetical terrain is entered, ecosystematic or biogenetic pathways and correlations might be traced, or a bioactive compound of unknown structure matures to a molecular graph stimulating both organic chemists and pharmacologists to further research.

REFERENCES

Abragam, A. (1961) *The Principles of Nuclear Magnetism*, Clarendon Press, Oxford.

Adam, P., Arigoni, D., Bacher, A., *et al.* (2002) *J. Med. Chem.*, **45**, 4786–4793.

Allenmark, S. G. (2000) *Nat. Prod. Rep.*, **17**, 145–155.

Arigoni, D., Sagner, S., Latzel, C., *et al.* (1997) *Proc. Natl. Acad. Sci.*, **94**, 10600–10605.

Berger, S. and Braun, S. (2004) *200 and More NMR Experiments. A Practical Course*, John Wiley & Sons, Inc, Hoboken.

Berger, S. and SickerD. (2009) *Classics in Spectroscopy*, John Wiley & Sons, Inc, Hoboken.

Brenton, R. C. and Reynolds, W. F. (2013) *Nat. Prod. Rep.*, **30**, 501–524.

Bringmann, G., Noll, T. F., Gulder, T., *et al.* (2007) *J. Org. Chem.*, **72**, 3247–3252.

Bringmann, G., Gulder, T. A., Reichert, M., *et al.* (2008) *Chirality*, **20**, 628–642.

Bross-Walch, N., Kühn, T., MoskauD., *et al.* (2005) *Chem. Biodiv.*, **2**, 147–177.

Capdeville, R., Buchdunger, E., Zimmermann, J., *et al.* (2002) *Nat. Rev. Drug Discov.*, **1**, 493–502.

Chmurny, G. N. and Hoult, D. I. (1990) *Concept Magn. Reson.*, **2**, 131–149.

Claridge, T. D. W. (2009) *High-Resolution NMR Techniques in Organic Chemistry*, Elsevier Science, Oxford.

Croasmun, W. R. and Carlson, R. M. K. (1994) *Two-Dimensional NMR Spectroscopy: Applications for Chemists and Biochemists*, John Wiley & Sons, Inc, Hoboken.

Ernst, R. R. (1992) *Angew. Chem. Int. Ed. Engl.*, **31**, 805–823.

Ernst, R. R., Bodenhausen, G. and Wokaun, A. (1990) *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Oxford University Press, Oxford.

Freeman, R. (1998) *Spin Choreography: Basic Steps in High Resolution NMR*, Oxford University Press, Oxford.

Hoye, T. R., Jeffrey, C. S. and Shao, F. (2007) *Nat. Protoc.*, **2**, 2451–2458.

Jonas, J. and Gutowsky, H. S. (1980) *Annu. Rev. Phys. Chem.*, **31**, 1–28.

Keeler, J. (2010) *Understanding NMR Spectroscopy*, John Wiley & Sons, Inc, Hoboken.

Levinson, N. M. and Boxer, S. G. (2012) *PLoS One*, **7**, e29828.

Martin, G. E. and Zektzer, A. S. (1988) *Two-Dimensional NMR Methods for Establishing Molecular Connectivity: A Chemist's Guide to Experiment Selection, Performance, and Interpretation*, John Wiley & Sons, Inc, Hoboken.

Minch, M. J. (1994) *Concept Magn. Reson.*, **6**, 41–56.

Morris, G. A. and Emsley, J. W. (2010) *Multidimensional NMR Methods for the Solution State*, John Wiley & Sons, Inc, Hoboken.

Mulis, K. B. (1994) *Angew. Chem. Int. Ed. Engl.*, **33**, 1209–1213.

Napolitano, J. G., Lankin, D. C., Graf, T. N., *et al.* (2013) *J. Org. Chem.*, **78**, 2827–2839.

Nicolaou, K. C. and Snyder, S. A. (2005) *Angew. Chem. Int. Ed. Engl.*, **44**, 1012–1044.

Nyberg, N. T., Duus, J. O. and Sorensen, O. W. (2005) *J. Am. Chem. Soc.*, **127**, 6154–6155.

Pauli, G. F., Jaki, B. U. and Lankin, D. C. (2007) *J. Nat. Prod.*, **70**, 589–595.

Pescitelli, G., Kurtan, T., Flörke, U., *et al.* (2009) *Chirality*, **21**(Suppl 1), E181–201.

Pieters, L. A. and Vlietinck, A. J. (1989) *J. Pharm. Biomed. Anal.*, **7**, 1405–1471.

Rabenstein, D. L. (1984) *J. Chem. Ed.*, **61**, 909–913.

Ratcliffe, R. G. and Shachar-Hill, Y. (2006) *Plant*, **45**, 490–511.

Reynolds, W. F. and Enriquez, R. G. (2002) *J. Nat. Prod.*, **65**, 221–244.

Robien, W. (2009) *Trends Anal. Chem.*, **28**, 914–922.

Schneider, B. (2007) *Prog. Nucl. Magn. Reson. Spectr.*, **51**, 155–198.

Silverstein, R. M., Webster, F. X. and Kiemle, D. J. (2005) *Spectrometric Identification of Organic Compounds*, John Wiley & Sons, Inc, Hoboken.

Stappen, I., Buchbauer, G., Robien, W., *et al.* (2009) *Magn. Reson. Chem.*, **47**, 720–726.

Suyama, T. L., Gerwick, W. H. and McPhail, K. L. (2011) *Bioorg. Med. Chem.*, **19**, 6675–6701.

Traficante, D. D. and Steward, L. R. (1994) *Concept Magn. Reson.*, **6**, 131–135.

Voelter, W. (1978) *Pure Appl. Chem.*, **48**, 105–126.

Wenzel, T. J. and Chisholm, C. D. (2011) *Chirality*, **23**, 190–214.

Wüthrich, K. (2003) *Angew. Chem. Int. Ed. Engl.*, **42**, 3340–3363.

15

NMR of Large Molecules

**Christoph H. Wunderlich, Sarina Grutsch, Martin Tollinger and
Christoph Kreutz**

*University of Innsbruck and Center for Molecular Biosciences Innsbruck (CMBI), Institute of
Organic Chemistry, Innsbruck, Austria*

1 INTRODUCTION

1.1 A Brief History of NMR Spectroscopy of Large (Bio-)molecules

Over the past three decades, advances in the field of structural biology profoundly changed and still change our view on the structure and function of biologically relevant entities. Proteins and nucleic acids are nowadays no longer "black boxes" as we are familiar with most of the molecular details of their architectures and alike important with their dynamic properties. In this context, solution nuclear magnetic resonance (NMR) spectroscopy is well recognized as an indispensable tool for the elucidation of structure and dynamics of large, that is, high molecular weight, molecules (Bertini, McGreevy, and Parigi, 2012). The time line shown in Figure 1 summarizes milestones in the field of NMR (Figure 1a). It further shows the major hardware, isotope labeling, and methodological developments for the solution structure determination process by NMR (Figure 1b).

With the introduction of high field NMR spectrometers, larger biomolecules and their complexes could be addressed by NMR spectroscopic methods (Frydman, 2011). In the mid-1980s, the main focus lied on the structure elucidation of peptides and small proteins (Kumar, Ernst, and Wuethrich, 1980; Wuethrich *et al*., 1982). One of the first major achievements was the solution structure determination of a 57-amino acid protein – proteinase inhibitor IIA – structure in 1985 using protocols mainly developed by Kurt Wüthrich and coworkers (Williamson, Havel, and Wuethrich, 1985). At that time, the methodology was restricted to proteins with less than 70 amino acids and favorable properties such as good solubility in water and well-dispersed proton resonances.

With the introduction of ^{13}C- and ^{15}N-isotope labeling protocols by the expression of the target protein in minimal medium, the way was paved for sophisticated high dimensional triple resonance NMR experiments that significantly alleviated problems concerning spectral resolution and thus allowed unambiguous signal assignments of larger biomolecules (Hurd, 2011; Kay *et al*., 2011; Muchmore *et al*., 1989; Ruiz-Cabello *et al*., 2011). Using pulsed field gradients, experimental artifacts were suppressed and a sensitivity gain was accomplished by improved coherence selection methods (Hurd, 2011). The developments along with significant advances of hardware components (Figure 1b) led to a boom of NMR structural biology and gave interesting insights into the structure and function of biomacromolecules. In the late-1990s, an additional long-range structural parameter, the residual dipolar coupling, was described and experiments to measure it were added to the toolbox of biomolecular NMR spectroscopy (Lipsitz and Tjandra, 2004; Tjandra and Bax, 1997; Tolman *et al*., 1997). Pervushin *et al*. (1997) published a novel approach, namely
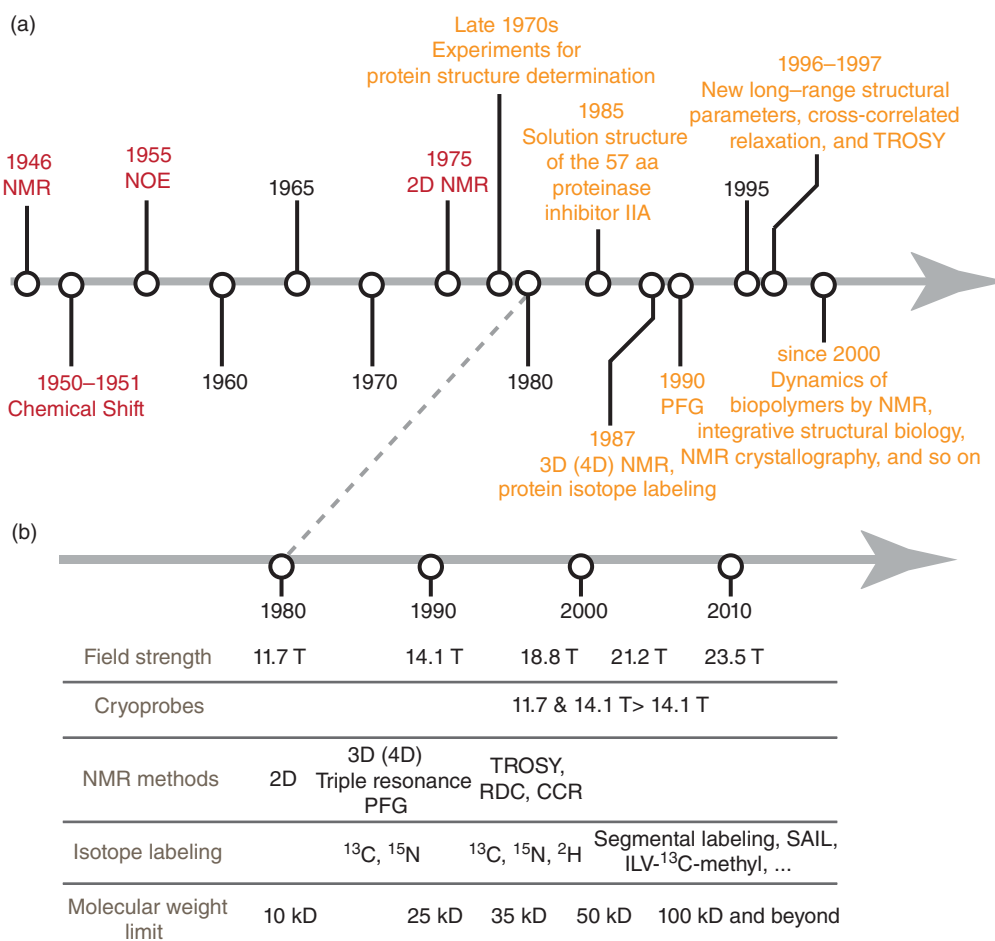
Figure 1 (a) Timeline highlighting milestones of solution NMR spectroscopy. Events in red relate to the basics of NMR physics, whereas major contributions to biomolecular NMR are shown in orange. (b) Condensed history of biomolecular NMR with major hardware, methodological, and isotope labeling developments. Abbreviations: NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; PFG, pulsed field gradients; TROSY, transverse-relaxation-optimized spectroscopy; T, Tesla; RDC, residual dipolar coupling; CCR, cross-correlated relaxation; SAIL, stereo-array isotope labeling; and kD, kilodalton; ILV, Isoleucine, Leucine, Valine.

the transverse relaxation-optimized spectroscopy (TROSY) pulse sequence, which is designed to select the component for which different relaxation mechanisms have almost cancelled, leading to a single, sharp peak in the spectrum. Thereby, a significant increase of both spectral resolution and sensitivity is achieved making NMR studies of large and complex biomolecules with molecular weights higher than 100 kD amenable. The contribution of biomolecular NMR spectroscopy to the field of structural biology is well documented by about 10,000 searchable entries in the Protein Data Bank

(PDB) for structures of proteins, nucleic acids, and their complexes obtained by NMR spectroscopy.

1.2 The Last Decade: Dynamics of Biomacromolecules Studied by Solution NMR Spectroscopy

One of the most recent developments in biomolecular NMR spectroscopy is the description of dynamic phenomena of biopolymers across various timescales ranging from pico- and nanoseconds to micro- and
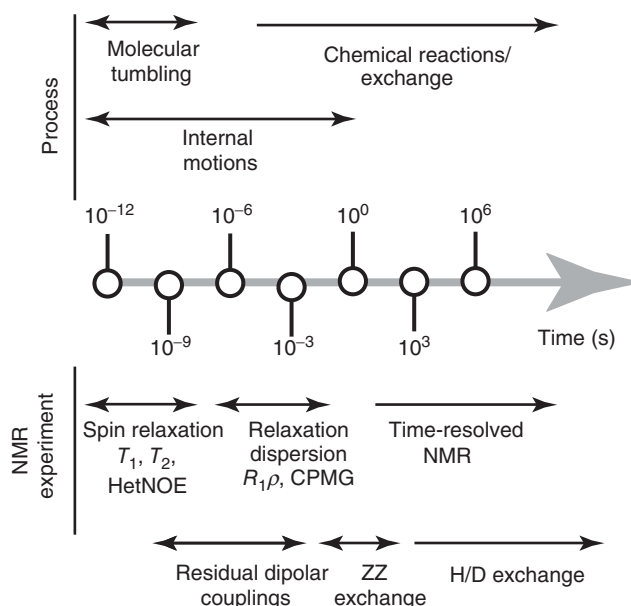
Figure 2 Accessible motions/dynamics by solution NMR spectroscopy. The process and the respective experimental technique are mapped on a timescale ranging from picoseconds to days.

milliseconds and beyond (Figure 2) (Bothe *et al.*, 2011; Fürtig *et al.*, 2007; Kay, 2011; Mittermaier and Kay, 2006; Mittermaier and Kay, 2009). In recent decade, significant efforts were put into the method development to study functional dynamics of biomolecules by observing relaxation properties of ^{15}N- and/or ^{13}C-spins (Akke *et al.*, 1997; Mittermaier and Kay, 2009).

Fast fluctuations in the nano- and picosecond time regimes can be characterized by the analysis of longitudinal (R_1) and transverse (R_2) relaxation rates and the heteronuclear Overhauser effect (hetNOE) of assigned ^{13}C and ^{15}N-resonances in the biomolecule under investigation. In this time regime, residual dipolar couplings (RDCs) can alternatively be used to extract dynamic information. The interpretation of the relaxation and RDC data to delineate quantitative information on the structural flexibility of a biomolecule can be a challenging task, for example, if a motional coupling between local fluctuations and the overall tumbling rate of the molecule occurs (Zhang *et al.*, 2006). Data analysis normally proceeds in the framework of the model-free approach. Order parameters ($0 < S^2 < 1$) are extracted, giving information on the flexibility on a residue-resolved basis (Lipari and Szabo, 1982a; Lipari and Szabo, 1982b).

Slower motions occurring in the micro- to millisecond time regimes can be addressed using relaxation dispersion (RD) techniques. Structural adaptations of biomolecules in this time regime are often involved in functional important processes, such as ligand binding or enzymatic reactions. Two experiments, the Carr–Purcell–Meiboom–Gill (CPMG) RD method and the determination of the power dependence of R1ρ, are suitable to extract the exchange lifetimes between multiple states, the populations of the states, and the chemical shift (CS) difference between the states (Palmer, Kroenke, and Loria, 2001). In proteins, both methods are routinely used to obtain information on backbone dynamics by applying the techniques to the isolated amide ^{15}N–^{1}H spin systems. For nucleic acids, recent findings suggest that nonexchangeable ^{13}C–^{1}H-spin systems will be the reporter spin of choice for the characterization of micro- to millisecond dynamics (Wunderlich *et al.*, 2012).

The description of dynamics in combination with structure determination via X-ray crystallography (and NMR spectroscopy) had an enormous impact on the view of biological function. The functionally important state was often found to be a transiently and weakly populated species, which is, for example, interacting with a ligand or carrying out catalysis.

2 NMR AS A TOOL IN DRUG DISCOVERY

As mentioned earlier, solution NMR spectroscopy is perfectly suited not only to elucidate structures of small molecules but also to reveal molecular details of the structure of biopolymers. In addition, NMR spectroscopy was soon attributed a great potential in the drug discovery process (Fernandez and Jahnke, 2004; Pellecchia *et al.*, 2008; Pellecchia, Sem, and Wuthrich, 2002). Currently, disseminated NMR methods in drug discovery are discussed in the following paragraphs.

2.1 Ligand Screening and Hit Validation by NMR

Two complementary approaches, (i) target-detected and (ii) ligand-detected methods, for ligand screening and hit validation by solution NMR spectroscopy are routinely applied. In case of target-detected techniques, the most widespread technique relies on monitoring changes of the probably most elementary parameter in NMR, the CS. As this observable is highly sensitive to its microenvironment, information about a ligand–host interaction and what parts of the ligand and the host are involved in this cross talk on a molecular level can be extracted. For example, observing the ^{15}N-CS perturbation induced by structural changes linked to ligand binding is the most widespread method to investigate protein–ligand interactions (Carlomagno, 2005; Roberts, 1999). It gives information on the affinity of a ligand toward the protein and it further allows the localization of the interaction site. By the acquisition of a series of two-dimensional ^{15}N/^{1}H correlation spectra (^{15}N/^{1}H-HSQC, heteronuclear single quantum correlation) as a function of increasing amounts of ligand, CS changes can be measured and, if the resonance assignment is known, mapped onto the protein backbone amides (Figure 3a). Additional information on the target protein, for example, a previously determined solution structure, can help in rapidly gain insights into the ligand-binding mode by intermolecular distance restraints obtained from nuclear Overhauser effects (NOEs).

Alternatively, ligand-detected approaches rely on perturbations of ligand's spin parameters caused by substoichiometric additions of the host molecule. These NMR experiments, such as the saturation transfer difference (STD) experiment, and techniques monitoring relaxation parameters are sensitive to the overall molecular motion of the tested compound, which strongly differs in the free or bound state. The methods are fast and can be applied in an isotope label-free way. Moreover, the ligand-detected approaches can further be used to identify a binder compound within a mixture of several compounds. The methods are also very attractive for larger proteins and complexes as only little amounts of the host molecule are needed. However, one has to keep in mind that the experiments are designed to monitor transient ligand binding, that is, the methods are suitable for systems with dissociation constants in the high micromolar or millimolar range.

In a typical STD experiment, simple 1D-proton-detected NMR experiments are acquired at ligand/host ratios of about 100/1 with (by irradiation of protein methyl groups around 0.5–1.0 ppm using a selective 180° shaped pulse) and without (by irradiation with a selective 180° shaped pulse placed at a frequency, for instance, 100 ppm outside the proton CS range) selective irradiation of host resonances (Mayer and Meyer, 1999). By spin diffusion, the saturation effect at on-resonance saturation quickly propagates across the entire receptor nuclei. If the ligand binds to the receptor, the resonance saturation will also spread onto the ligand resonances leading to an attenuation of the signal intensity. For analysis purposes, the spectra with and without selective irradiation are subtracted by appropriate phase cycling yielding the STD spectrum containing only signals of the ligands that bind. Protein signals can effectively be removed by including a filter sensitive to high molecular weight compounds based on their longitudinal relaxation time in the rotating frame ($T_1\rho$) (Figure 3b).

The ligand-detected experiments have a lower informational content as the CS mapping, for instance, the binding site of the ligand on the host cannot be located. Thus, the methods are used primarily in screening and hit validation. Table 1 gives an overview on currently used NMR experiments for screening and hit validation of compounds identified from high throughput screening (HTS) methods (Dalvit *et al.*, 2000; Fejzo *et al.*, 1999; Hajduk, Olejniczak, and Fesik, 1997; Hajduk *et al.*, 1999; Jahnke, Ruedisser, and Zurini, 2001; Yan *et al.*, 2003).
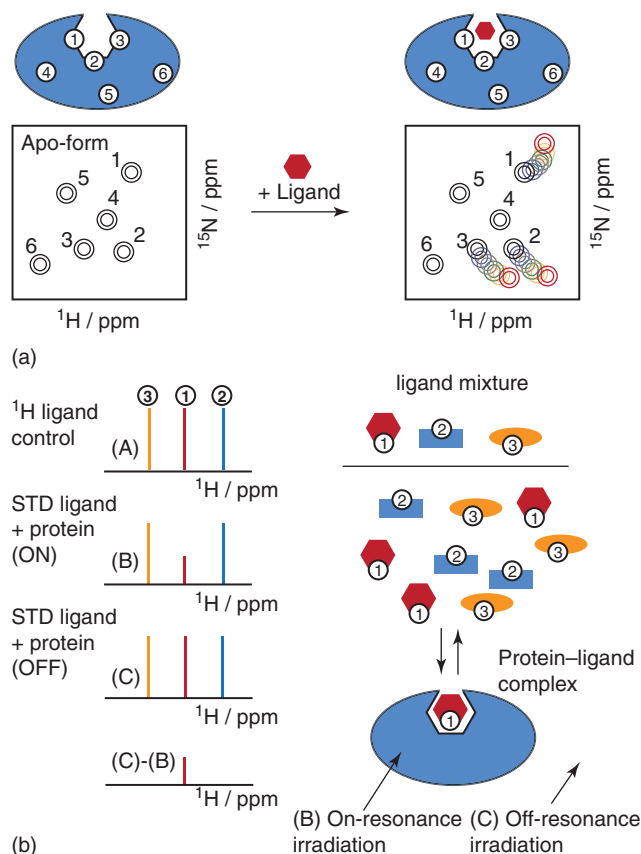
Figure 3 (a) Mapping of a ligand binding site by chemical shift (CS) perturbation. A ^{1}H–^{15}N-HSQC spectrum of the host protein with uniform ^{15}N-labeling pattern is obtained in the apo-form. Then, increasing amounts of ligand are added in a stepwise manner. The CS perturbation is monitored by acquiring HSQC spectra after each ligand addition. Only ^{15}N–^{1}H-spin system nearby the binding site (1, 2, 3) will show pronounced CS changes, whereas distant amide resonances will be unaffected by the ligand binding event (4, 5, 6). (b) Ligand binding detected by the saturation transfer difference (STD) experiment. For details of the experiment, please see text.

2.2 Hit and Lead Optimization by NMR

One of the most recent developments concerning NMR in drug discovery is its implementation into an approach called *fragment-based drug design* (*FBDD*). FBDD is based on screening a small number of molecules with low affinity toward the drug target. The so-called fragments are structurally and chemically diverse; thus, a very efficient sampling of the chemical space in the screening process is possible. Once some low affinity fragments are identified, they serve as starting point for fragment evolution, linking or optimization, or fragment self-assembly. NMR is particularly useful in the characterization of

such low affinity binders with dissociation constants in the micro- to millimolar range. The "SAR by NMR" (structure–activity relationship by NMR) approach uses CS mapping as a tool for hit and lead optimization. In this approach, a mixture of potential binders is screened in the presence of another weak binding ligand. On the basis of the CS perturbation analysis of the host resonances, compounds residing close to the first ligand's binding site are filtered out. Further characterization, for instance analysis of the interligand NOEs, is used to establish a rationally designed linker module between the two low affinity ligands resulting in a bidentate compound with significantly improved binding properties

Table 1 NMR methods useful in drug screening and hit validation

Experiment	Detection	Application	Description of method
Chemical shift perturbation	Host	Primary screen, validation, and localization of binding site	Chemical shift changes of host resonances are monitored as a function of increasing ligand concentration
STD NMR[a]	Ligand	Primary screen and validation	Compound identification and identification of functional groups interacting with host
WaterLOGSY[b]	Ligand	Primary screen	Compound identification by water-mediated NOE
$T_1\rho$[c] and T_2[d] relaxation	Ligand	Primary screen and hit validation	Longer T_1 due to complex formation, line broadening (faster T_2) due to complex formation, and build up curves identify functional groups interacting with host
FABS[e]	Substrate or cofactor	Primary screen and hit validation	Uses ^{19}F NMR to detect reference substrates or cofactors to monitor enzymatic reactions
FAXS[f]	Known ligand (spy)	Primary screen and hit validation	Uses ^{19}F NMR to detect displacement of a fluorinated spy molecule
Measurement of diffusion rate	Ligand	Primary screen and hit validation	Observing difference of diffusion rate for ligands in the bound versus free form

[a] Saturation transfer difference.
[b] Water–ligand observed via gradient spectroscopy.
[c] Longitudinal relaxation time in rotating frame.
[d] Transverse relaxation time.
[e] Fluorine atoms for biochemical screening.
[f] Fluorine chemical shift anisotropy and exchange for screening.

because of a larger number of interactions (enthalpic contribution) and a reduced loss in rotational and translational entropy on binding. Optionally, the protein or the first (known) ligand can be modified using a paramagnetic tag aiding in the identification of compounds, which bind in proximity to the first compound (SLAPSTIC). Table 2 gives a selection of currently disseminated NMR experiments for hit and lead optimization (Becattini and Pellecchia, 2006; Carulla *et al.*, 2005; Chen *et al.*, 2007; Jahnke *et al.*, 2000; Shuker *et al.*, 1996).

The so far discussed methods have proved to be well suited to screen for ligands and subsequently improve their design to yield high affinity binders for challenging targets, for which other approaches failed to come up with viable lead structures. In today's view, NMR spectroscopy has a higher potential for drug discovery purposes, whereas X-ray crystallography is regarded as the method of choice for structure determination.

In the next two sections, we outline recent advances in the field of protein and nucleic acid drug discovery by NMR spectroscopy, which illustrate that NMR spectroscopy is and will be an indispensable tool for designing compounds targeting the two major classes of biomolecules.

2.3 Ribonucleic Acids as Drug Targets

2.3.1 RNA – Small Molecule Interactions

In retrospect, the biological importance of ribonucleic acid (RNA) was underestimated for a long time and proteins and DNAs were regarded as the most promising targets for drug development (Huettenhofer, Schattner, and Polacek, 2005). Historically, the role of RNA in a cellular context was mainly restricted to its function as translator molecule (transfer, messenger, and ribosomal RNAs) to transform the information of DNA into a functional amino acid sequence. Thus, RNA was believed to generally adapt an A-form double helix without special structural features. At that time, RNA was not considered to be an important target for drug development as many interactions with small molecules were thought to be unspecific – either by electrostatic interactions between the positively charged ligands and the negatively charged phosphate backbone or by random stacking interactions. These results were in clear contrast with the complex molecular architecture of the ribosome, for which a high resolution 3D-structure was solved by X-ray crystallography about 15 years ago (Moore and Steitz, 2003; Steitz

Table 2 NMR methods useful in hit validation and lead optimization

Experiment	Detection	Application	Description of method
SAR[a] by NMR	Host and ligand	Structural information, FBDD[e] screen, and lead optimization	Construction of bi- or higher-dentate ligands
SAR by ILOES[b]	Ligand to ligand	FBDD[e] screen and lead optimization	Monitoring of protein-mediated ligand–ligand interactions of binders located at adjacent sites
SLAPSTIC[c]	Ligand	FBDD[e] screen and lead optimization	Monitoring of ligand–ligand interactions of binders located at adjacent sites exploiting the influence of a protein side chain carrying a spin label
INPHARMA[d]	Ligand to ligand	Drug characterization	Monitoring of protein-mediated ligand–ligand interactions competing for a single binding site
Pharmacophore by ILOES	Ligand to ligand	FBDD[e] screen and lead optimization	Monitoring of protein-mediated ligand–ligand interactions, obtained data is used for pharmacophore-based search of bidentate binders

[a] Structure–activity relationship.
[b] Interligand nuclear Overhauser effect (NOE).
[c] Spin labels attached to protein side chains as a tool to identify interacting compounds.
[d] Interligand NOEs for pharmacophore mapping.
[e] Fragment-based drug design

and Moore, 2003). In addition, the ribosome was one of the first examples that RNA is a potential target for small molecule ligands as antibiotics were shown to bind not only to the protein particles but also to the RNA component (Fourmy *et al*., 1996; Ramakrishnan, 2002). These findings and the subsequent reports on novel RNA species with important biological functions led to a radical change in the perception that RNAs are restricted in function and comprise a uniform architecture. Soon, numerous functionally important RNAs with intricate structural features were identified, such as riboswitch aptamer domains, naturally occurring ribozymes, regulatory sequence elements of retroviruses, and micro and small nucleolar RNAs or RNA thermometers (Cochrane and Strobel, 2008; Lu, Heng, and Summers, 2011; Narberhaus, Waldminghaus, and Chowdhury, 2006; Serganov and Patel, 2012; Turner and Summers, 1999; Wu and Feigon, 2007). The versatile nature of RNA was further promoted by applying the SELEX (systematic evolution of ligands by exponential enrichment) approach, which led to the discovery of RNA sequences (i.e., RNA aptamers and enzymes) with the capability to interact specifically with druglike small molecules or to carry out catalysis (Mayer, 2009; Stoltenburg, Reinemann, and Strehlitz, 2007). The 3D structures of naturally occurring and artificially generated aptamers and larger biologically functional RNAs (e.g., group I and II introns and ribosomal RNAs) revealed complex fold architectures comparable to those of proteins (Golden *et al*., 1998; Woodson, 2005). The molecular

details of the interactions between small molecules and RNA residues at the binding site also shed light on the ligand selectivity in RNA–ligand complexes. Intricate hydrogen bonding networks or interactions by appropriately positioned metal ions lead to a high specificity in the small molecule recognition. A selection of such naturally occurring riboswitches and artificially selected binary complexes is shown (Figure 4).

As a first example, we picked a riboswitch aptamer domain. Riboswitches belong to the class of noncoding RNAs and are able to specifically recognize metabolites (Breaker, 2012; Tucker and Breaker, 2005). They are predominantly found in the untranslated leader regions of bacterial messenger RNA (mRNA). After the ligand binding event occurred a structural change in the adjoining downstream expression platform signals either "on" or "off" for gene expression. A riboswitch aptamer domain responsive to the small-molecule cyclic diguanosine monophosphate **1** (c-di-GMP) was identified recently (Sudarsan *et al*., 2008). The highly conserved RNA domain was found in a large number of pathogenic bacterial organisms and resides upstream of the open reading frame (ORF) for proteins involved in the metabolism of c-di-GMP (for example diguanylate cyclase and phosphodiesterase proteins). The recently solved X-ray crystal structure of the class I c-di-GMP riboswitch aptamer domain in complex with its cognate ligand c-di-GMP **1** revealed the molecular details of the ligand recognition (Kulshina, Baird, and Ferre-D'Amare, 2009;
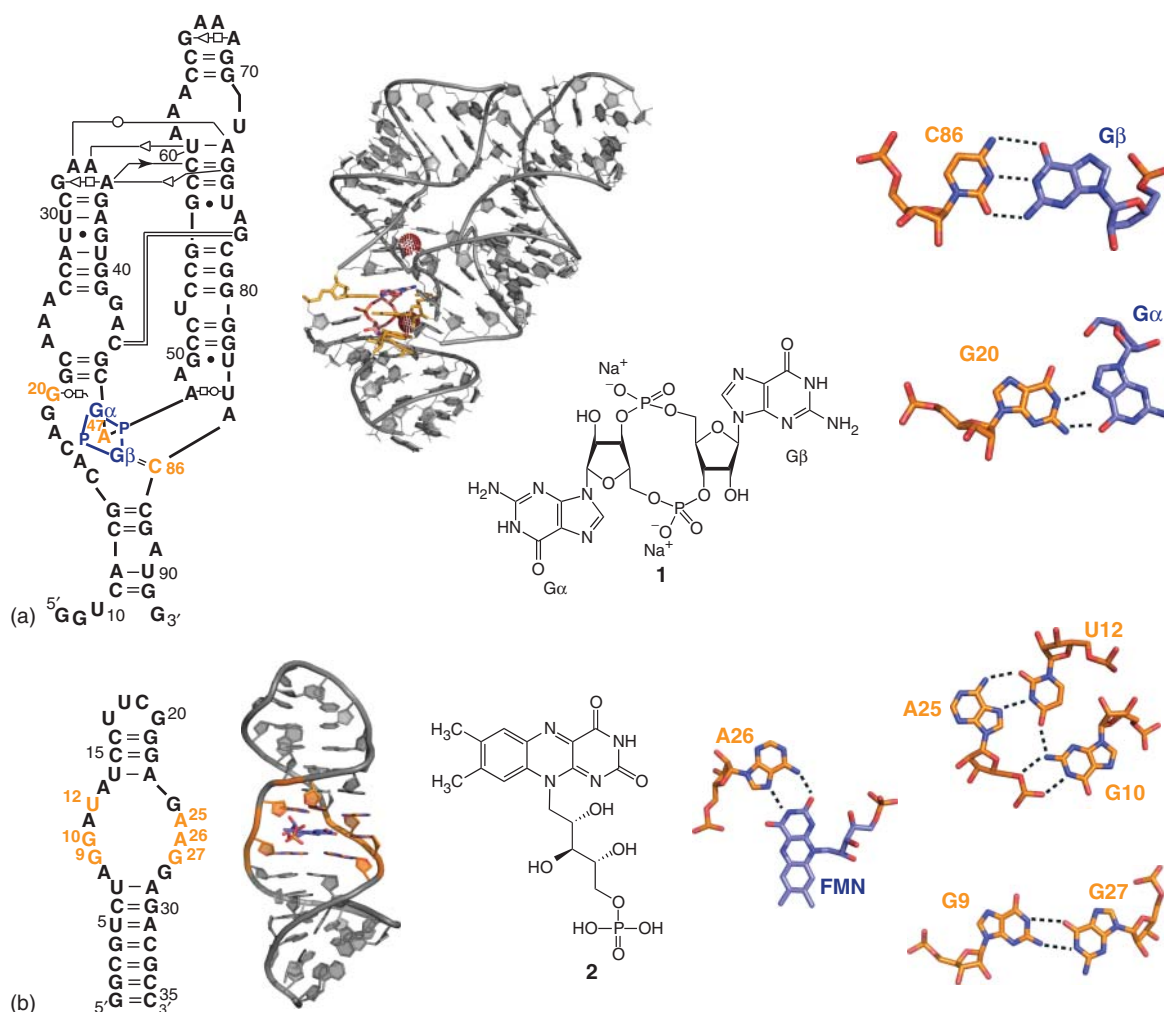
Figure 4 (a) The cyclic-di-guanosine monophosphate (c-di-GMP) riboswitch aptamer domain class I. The secondary and tertiary structures are shown in Leontis–Westhof annotation along with the crystal structure of the binary complex (red dotted spheres magnesium ions, RNA residues in contact with ligand, and c-di-GMP are shown in stick representation and a molecular formula of c-di-GMP **1**). Close-up on C86 and G20 and their interactions with the ligand moiety. (b) The flavin mononucleotide (FMN) aptamer. The secondary structure along with the solution structure of the complex (RNA residues in contact with FMN are shown in stick representation and a molecular formula of FMN **2**). Close-up on A26 interacting via H-bonds with FMN and on noncanonical interactions making up the binding pocket.

Smith *et al.*, 2009). The symmetric ligand c-di-GMP **1** is recognized asymmetrically – by guanosine 20 (G20) via a reverse Hoogsteen base pair with guanosine alpha (Gα) and by cytidine 86 (C86) via a standard Watson–Crick base pair with guanosine beta (Gβ) (Figure 4a).

The second example for complex RNA–ligand interactions is an artificially derived RNA aptamer responsive to flavin mononucleotide (FMN) **2** (Burgstaller and Famulok, 1994; Fan *et al.*, 1996).

Again, a ligand-specific hydrogen-bonding network is formed between **2** and an RNA nucleotide (A26). Furthermore, a ligand-binding pocket comprising noncanonical structural elements, such as a base triple G10-U12-A26 and a non-Watson–Crick G10-G27 base pair, is established resulting in favorable stacking interactions (Figure 4b).

These examples illustrate that RNA–small molecule interactions can be highly discriminative making RNA an interesting drug target. As, for

example, riboswitches are mainly found in pathogenic bacteria – for example, in the bacterium *Vibrio cholerae* causing cholera – strong binders toward this type of RNAs represent a highly desirable goal as this would open the avenue for the development of new antibiotics replacing established compounds (e.g., aminoglycosides), for which more and more bacterial strains show resistance (Blount and Breaker, 2006).

In the next section, we outline NMR spectroscopic methods suitable to monitor RNA–ligand interactions and some aspects and pitfalls that have to be considered when targeting drugs versus RNA. The methods have to consider the intrinsic affinity of RNA to interact in a nonspecific manner (e.g. electrostatic interactions between positively charged ligands or nonspecific stacking interactions with extended π-systems) and that specific RNA–ligand interactions often require extensive structural adaptations either induced by the ligand ('induced fit') or via a conformational selection pathway (Haller, Souliere, and Micura, 2011; Liberman and Wedekind, 2012; Nussinov and Wright, 2009).

2.3.2 NMR Spectroscopic Methods to Monitor RNA–Ligand Interactions

2.3.2.1 Target Observing Methods
As previously outlined, target observed methods mainly rely on monitoring CS perturbations of host resonances induced by binding of the ligand. For that purpose, either homonuclear [1]H- or heteronuclear correlation spectra (for the latter type of experiments [13]C- and/or ^{15}N-labeling is needed) are consecutively recorded at increasing ligand concentrations.

The most straightforward NMR-based method to study the interaction of a small molecule with a nucleic acid target exploits the phenomenon that the imino protons (N^1-H of guanosine and N^3-H of uridine) have a very distinct CS signature (Mayer *et al.*, 2006; Mayer and James, 2005). These protons are involved not only in the hydrogen bonding between nucleobases, like in the Watson–Crick G–C and A–U base pairs, but also in noncanonical interactions, like in G–U wobble or sheared G–A base pairs. The hydrogen bonding interaction leads to a strong deshielding effect on the proton nucleus resulting in a downfield CS for imino protons between 10 and 15 ppm. Hence, the imino proton resonances

are clearly separated from the other ^{1}H resonances making the assignment applicable for medium sized RNAs up to 30 nt by a two-dimensional [1]H-NOE experiment. The CS perturbation of the imino proton resonances is then monitored as a function of increasing ligand amounts. Thereby, the binding site can be mapped onto the RNA by identifying the imino protons with the most rigorous CS change. The affinity of the ligand–RNA interaction can be determined by fitting the following equation (1):

$$\Delta\delta = \frac{(\delta_{\text{complex}} - \delta_{\text{RNA}})\{(R_0 + L_0 + K_D) - \sqrt{(R_0 + L_0 + K_D)^2 - 4R_0L_0}\}}{2R_0} \quad (1)$$

with $\Delta\delta$ CS perturbation, δ_{complex} CS of RNA resonance in bound state, δ_{RNA} CS of RNA resonance in free state, R_0 total RNA concentration, L_0 total ligand concentration, and K_D dissociation constant. The values of K_D and δ_{complex} can be determined from the titration of the RNA aptamer with the respective ligand by a nonlinear least squares fitting of the CS changes as a function of total ligand concentration [L_0] to the Equation (1). The approach is exemplified on the HIV-1 TAR RNA construct (Aboul-ela, Karn, and Varani, 1995). For that RNA, the imino proton assignments are known. Two ligands, argininamide **3** and neomycin B **4**, were used in titration experiments and the imino proton region between 10 and 15 ppm was monitored by acquiring 1D proton NMR spectra at increasing ligand concentrations (Figure 5a). The most pronounced CS changes were observed for imino protons (G5, G10, and U20) located in vicinity of the bulge region, where both ligands are known to bind. The N^1-H-resonance of guanosine residue 5 was used to estimate the dissociation constant (K_D) of the binary HIV-1 TAR RNA–argininamide complex (Brodsky and Williamson, 1997). Fitting the CS perturbation data to the aforementioned equation yielded a K_D value of approximately 500 μM (Figure 5b).

Albeit the imino proton resonances of an RNA represent valuable reporter spins, the major drawback using this approach is their exchangeable nature. That is, these protons are in constant chemical exchange with [1]H-nuclei of the bulk water. In contrast to protein amides, the exchange occurs at a timescale leading to line broadening, as it is found for weak base pair, or the imino proton signals are rendered unobservable by collapsing entirely in the dominating water peak. Accordingly, the exchange with water

(a)

(b)
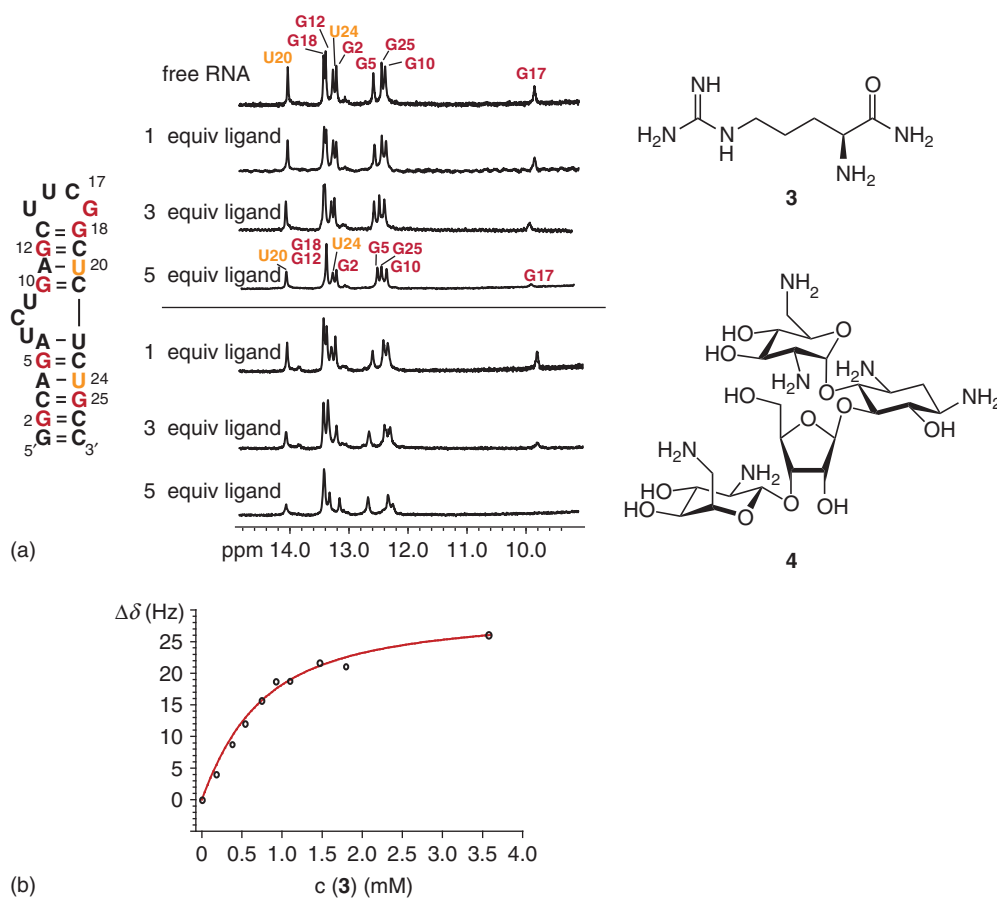
Figure 5 (a) Chemical shift perturbation (CSP) binding assay for the HIV-1 TAR RNA and the ligands argininamide **3** and neomycin B **4**. The imino proton region between 10 and 15 ppm from ^{1}H proton spectra is shown at increasing ligand concentrations. (b) Dissociation constant (K_D) determination of the HIV-1 TAR RNA argininamide complex. The CSP analysis was carried out for the imino N^1-H resonance of G5 and yielded a K_D value of approximately 500 μM.

leads to indiscernible NMR peaks of unprotected imino protons residing in single-stranded regions. Thus, only residues involved in base pairing can be used to monitor the binding of a small molecule.

To address binding studies to single-stranded sequence regions, one has to resort to NMR experiments eligible to monitor nonexchangeable proton resonances. For that purpose, the probably most useful homonuclear correlation experiment is the total correlation spectroscopy (TOCSY) experiment, which correlates the H5 and H6 aromatic protons of the pyrimidine uridine and cytidine residues (Wüthrich, 1986). In well-structured RNAs, these protons normally are well dispersed and can be assigned. An advantage of this approach is that no special stable isotope (^{13}C, ^{15}N) labeling pattern is

needed and that nonexchangeable proton resonances are observed, but at the expense of significantly longer measurement times as compared to the one-dimensional NMR spectroscopy focusing on the imino proton region.

Alternatively, in ^{15}N-labeled RNA, the nitrogen-bound imino protons (H^3-N in uridines and H^1-N in guanosine) can be used for ligand-binding assays. With the ^{15}N- and ^{15}N/^{13}C-labeled ribonucleotide triphosphates being commercially available, T7 RNA polymerase-assisted *in vitro* transcription is today the most common way to obtain isotope-modified RNAs (Lu, Miyazaki, and Summers, 2010; Milligan *et al.*, 1987; Milligan and Uhlenbeck, 1989). The major drawback of the imino ^{15}N–^{1}H spin systems is that as aforementioned these

protons show enhanced water exchange rates as compared to protein amides. However, consistent with the prevalent adaptive recognition mechanism of RNA/ligand interactions, the binding of a ligand to an RNA often leads to a change in the base pairing pattern or the formation of new base pairs resulting in a clear new spectral imino proton resonance signature for the ligand-bound state.

The most useful spin systems for monitoring ligand binding to single-stranded regions of RNA are ^{13}C–^1H spin couples of the sugar or the nucleobase moieties. For the analysis of CS perturbations, the aromatic C6–H6 of pyrimidines, the C8–H8 and C2–H2 of purines, and the anomeric sugar C1′–H1′ spin pair are particularly useful as they offer a sufficiently large CS dispersion. The ^{13}C labeling pattern can introduced either by chemical or by enzymatic means using ^{13}C-modified phosphoramidites or ^{13}C-, ^{15}N-labeled ribonucleotide triphosphates, respectively. Using the former chemical approach, the labeling pattern can be introduced on a site-specific single nucleotide level or using the enzymatic RNA synthesis protocol on a nucleotide level (Kloiber et al., 2011; Wunderlich et al., 2012). As an example, we again picked the HIV-1 TAR RNA–argininamide complex. The ^{13}C-labeling pattern for this target RNA was introduced by substituting all uridine residues by the 6-^{13}C-modified analogs using appropriate phosphoramidite precursors in a chemical solid-phase synthesis approach (Figure 6a). To this ^{13}C-enriched RNA, the ligand argininamide was added successively (0.5 and 5 equivalents ligand). By acquiring ^1H–^{13}C-HSQC (heteronuclear single quantum correlation) spectra, the uridines with the most significantly shifted ^1H–^{13}C correlation peaks were identified. The residues U23, U25, U38, U40, and U42 all clustered at the bulge region show the most drastic CS perturbations, and this bulge is the ligand-binding site of argininamide to the HIV-1 TAR RNA (Figure 6b).

So far, either unlabeled or RNAs with native isotope labeling schemes were discussed in context with ligand-binding assays. In recent few years, an alternative reporter spin nucleus was exploited for drug screening purposes by NMR spectroscopy. Fluorine has very attractive NMR spectroscopic parameters (Kreutz and Micura, 2008). It has a 100% natural abundance, possesses an intrinsic NMR sensitivity almost as high as protons (83%), and offers a CS dispersion that is about 100 times that of protons. These properties make fluorine an ideal candidate

(a)

(b)
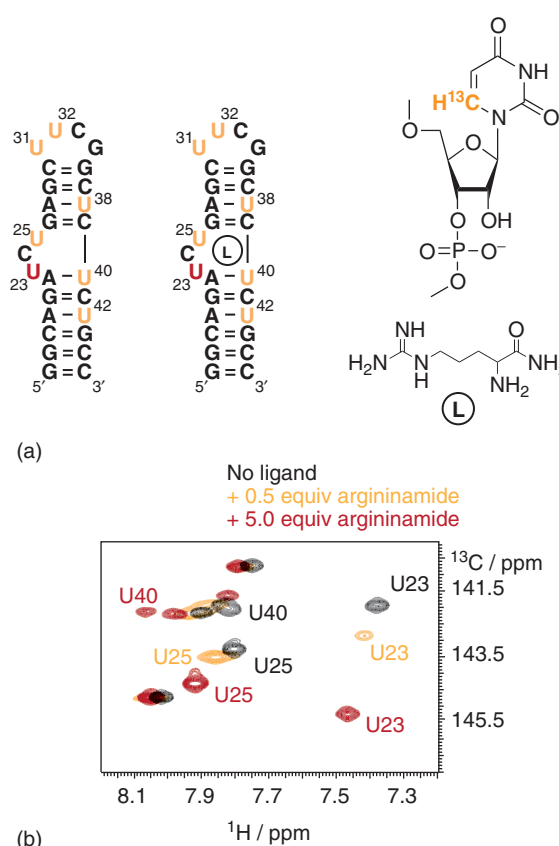
Figure 6 (a) Secondary structure representation of the HIV-1 TAR RNA. The ligand argininamide (L) is shown along with the 6-^{13}C-uridine reporter group. (b) Chemical shift perturbation binding assay by monitoring the ^1H–^{13}C-spin pairs via ^1H–^{13}C-HSQC spectra. The binding site can unambiguously identified at the bulge region, where the largest proton and carbon chemical shift changes are evoked on the addition of argininamide. The uridine residues with the strongest chemical shift changes are highlighted (U23, U25, and U40).

to be used as an alternative spin label complementary to ^{13}C- and ^{15}N-labeling, however, at the cost of introducing a non-native modification (Kreutz et al., 2005). Nonetheless, the introduction of fluorine represents an orthogonal isotope labeling pattern yielding a strong NMR signal that appears against a background devoid of signals from other nuclei.

Hennig et al., (2007) recently elaborated an enzymatic method to incorporate fluorinated nucleobases into a target RNA (Scott et al., 2004). They used 2-fluoroadenosine, 5-fluorocytidine, and 5-fluorouridine triphosphates in an in vitro transcription reaction using T7 RNA polymerase to obtain

uniformly ^{19}F-labeled targets on a nucleotide level. Complementary to the enzymatic labeling protocol, chemical RNA solid-phase synthesis is a very efficient and elegant way to introduce an arbitrary number of fluorine labels at any positions within the target nucleic acid (Kreutz *et al.*, 2006; Puffer *et al.*, 2009). The applicability of this approach is further promoted by the commercial availability of all four 2′-F-modified and the 5-F-uridine phosphoramidites. In addition, a novel fluorine tag, a 2′-SCF$_3$-group, with improved spectral properties was introduced recently (Fauster, Kreutz, and Micura, 2012).

The ^{19}F-nucleus is perfectly suited to conduct ligand binding assays as no other interfering signals are encountered and fluorine is intrinsically very sensitive to changes in its microenvironment because of the high sensitivity of the fluorine CS shielding parameter (Gerig, 1994). However, as the fluorine modification is non-native, the question remains, whether the introduction of fluorine atoms results in a structural perturbation of the target RNA or not. In the case of 5-fluorouridine substitutions, only minor changes on the RNA structure and stability were found indicating that a single 5-fluorouridine replacement represents a structural minimally invasive spin label option. In contrast, the 2′-F-modified nucleotides have to be placed with more care as they exert a more pronounced influence on the RNA structure. The C3′-endo sugar conformation is strongly favored leading to more stability in a double helix containing these fluorinated nucleotide analogs. Further, a functional hydroxyl group is removed.

As an example, we picked a ligand screening study between various antibiotics and an RNA oligonucleotide originally derived from a SELEX selection (Jiang and Patel, 1998). The *in vitro*-selected aptamer was designed to be responsive to the aminoglycoside tobramycin. To test the applicability of ^{19}F NMR as a tool for screening aminoglycoside binders toward the tobramycin aptamer, a difluorinated RNA, bearing a 2′-F-cytidine residue aside the binding pocket and a 2′-F-adenosine moiety placed in the 14-nt bulge region, was synthesized by chemical means. Addition of the cognate ligand, tobramycin **5**, was monitored by ^{19}F NMR spectroscopy. As expected for the high affinity ligand (K_D approximately 12 nM), the binary RNA–small molecule complex and the free RNA are in slow exchange on the chemical shift timescale leading to two observable resonances of the loop fluorine resonance reflecting the free and the bound states. Saturation is reached after adding one

equivalent of ligand. The second fluorine-modified residue (2′-F-U) serves as an internal reference confirming the specific interaction between the RNA and the ligand because only the loop fluorine resonance exhibits CS changes as it is placed within the binding pocket (Figure 7).

In a screening procedure, the affinities of streptomycin **6** and spectinomycin **7** were tested by stepwise addition of these small molecules up to 2.5 equivalents of antibiotic (Figure 7). In case of streptomycin **6**, line broadening and a defined downfield shift of the loop fluorine signal reflected a weak, fast exchange but site-specific loop–aminoglycoside interaction. Finally, spectinomycin **7**, which is not structurally related to tobramycin **5** did not show any detectable affinity toward RNA aptamer. This clearly shows that NMR spectroscopy of fluorine-modified nucleic acids is well suited for ligand binding assays and can be used to discriminate high and weak affinity binders and can be exploited to identify the potential binding pocket.

2.3.2.2 Ligand Observing Methods

The major drawback of host observing methods is that a relatively large amount of the biomolecule is needed and no drug mixtures, that would allow a faster screening procedure, can be used in the assay. In cases, such as only a limited amount of host RNA is available, it is advisable to resort to ligand-based NMR methods. To this end, methods [e.g., STD or water–ligand observed via gradient spectroscopy (WaterLOGSY) experiments], which are established for protein–drug interactions studies, are adapted for nucleic acid systems. All the aforementioned advantages (Section 2) are valid. For the STD experiment, the imino protons with their distinct CS signatures between 10 and 15 ppm are very well suited for the selective irradiation procedure as these signals are normally devoid any ligand resonances. For example, the ligand binding properties of phenothiazines toward biologically relevant RNAs (HIV-1 TAR RNA and ribosomal A-site RNA) were investigated by differential STD NMR spectroscopy (Mayer and James, 2002, 2004; Mayer *et al.*, 2006). In this study, a structure-based *in silico* screen against the solution structure of the ribosomal A-site RNA was carried out yielding phenothazine as a promising scaffold. The experimental verification by NMR spectroscopy identified acetopromazine as a weak affinity, specific binder with a dissociation constant of approximately 350 μM. Using STD amplification
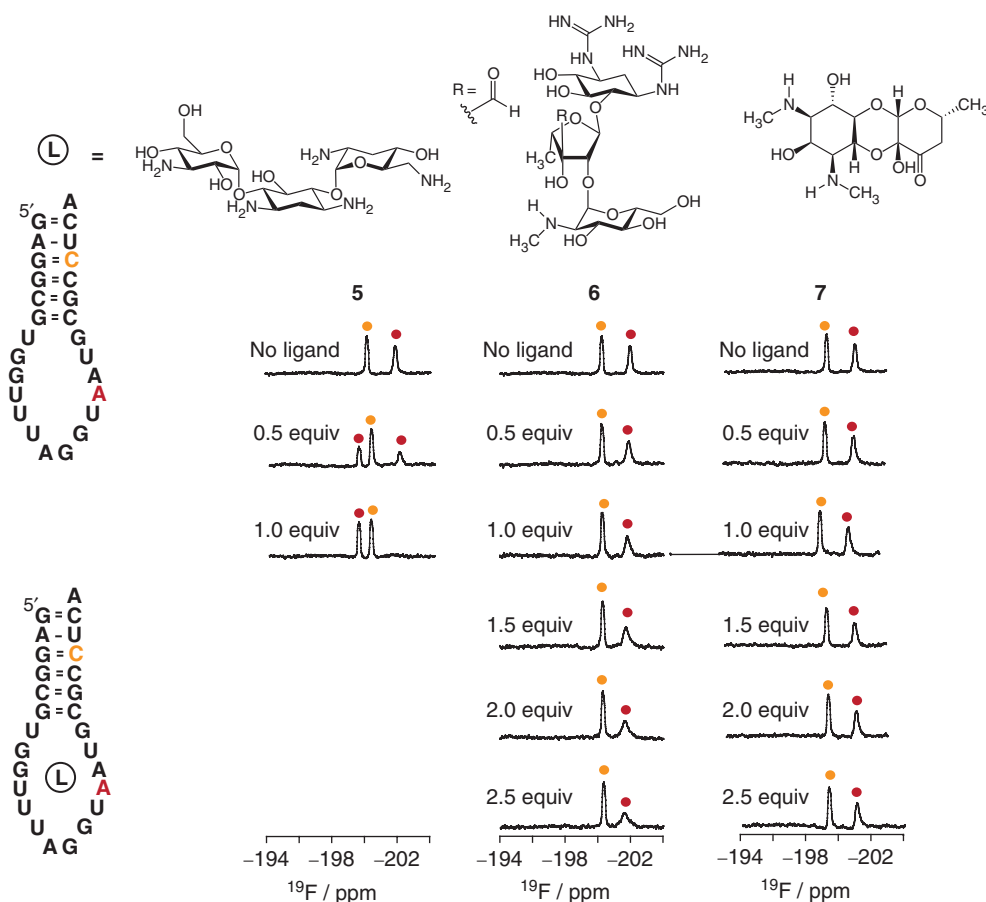
Figure 7 ^{19}F NMR spectroscopic ligand binding assay on a tobramycin-responsive *in vitro*-selected RNA aptamer. The two 2′-F-modified RNA residues are highlighted (2′-F-uridine in orange and 2′-F-adenosine in red). The RNA was screened for its affinity toward various antibiotics. The cognate ligand tobramycin **5** shows a well-behaved two-state equilibrium between bound and free states in slow exchange. Another weak affinity but specific binder, streptomycin **6**, was identified by observing the chemical shift changes of the ^{19}F-loop resonance. The antibiotic spectinomycin **7**, which is not structurally related to tobramycin **5**, was used as a negative control compound. Both fluorine resonances of the RNA remained unaffected by the addition of this ligand indicating no interaction.

factors, a binding epitope mapping was achieved, confirming that the aromatic ring of acetopromazine interacts with the target RNA.

A very elegant approach to probe RNA–ligand interactions was recently introduced (Lombès *et al.*, 2012; Moumne *et al.*, 2010). Building upon previous work, Tisné and coworkers use a fluorinated ligand in binding competition experiments to characterize small molecule–RNA interactions. In a proof of principle study, a ribosomal A-site mimic was used in a binding competition assay using a fluorinated spy molecule and known A-site binders, such as neamine or neomycin B. Using

^{19}F NMR spectroscopy, the progressive release of the fluorine-modified probe could be observed on the addition of competitor yielding an estimation of the dissociation constant (K_D) of the competing binder, as the K_D of the ^{19}F-probe is known. The competition assay can be adapted to the respective RNA–ligand interaction by modulating the affinity of the spy molecule toward the RNA. The approach was further extended to probe an artificial neomycin-B-dependent riboswitch. It proved to be suitable to monitor a conformation selection pathway applied by this RNA to bind its cognate ligand.

2.4 Proteins as Drug Targets

The most disseminated and well-established NMR spectroscopic assays used in protein–drug screening studies were already introduced (Section 2; Tables 1 and 2). In the following few paragraphs, we (exemplarily) picked two more recently developed NMR spectroscopic tools for drug screening.

2.4.1 ^{19}F NMR Spectroscopy as a Tool in Biochemical Screening and Drug Discovery

In recent few years, ^{19}F NMR-spectroscopy-based screening became an established methodology not only to analyze ligand binding but also to run functional assays. Two ^{19}F NMR-based assays were introduced recently: (i) FAXS (*f*luorine chemical shift *a*nisotropy and e*x*change for *s*creening) and (ii) 3-FABS (three-*f*luorine *a*toms for *b*iochemical *s*creening) (Dalvit *et al.*, 2003a, b; Vulpetti and Dalvit, 2012).

The former methodology (FAXS) is a ligand-based binding competition experiment. For this assay, a weak affinity spy molecule bearing a CF or CF_3 group is used. The approach can be extended by the use of a fluorinated control molecule, which exhibits no affinity against the protein target. These two molecules are selected from existing libraries, which is feasible as commercially available libraries contain up to 17% of fluorinated compounds. During the actual screening step, the fluorinated spy molecule is replaced by a competitor from the protein target. Thereby, elementary NMR spectroscopic parameters, such as the CS or the transverse relaxation time (i.e., line width), of the ^{19}F nucleus of the spy molecule are significantly altered. The large chemical shift anisotropy (CSA) of fluorine makes the difference in line width for the spy molecule in the free and bound state very large especially at the high magnetic fields currently used. The FAXS approach performed with a weak affinity-fluorinated ligand (spy molecule) and a ^{19}F-labeled control molecule with no affinity toward the target has proved to be very powerful and sensitive for primary screening of ligands to a protein target of interest. Furthermore, current technological advances, such as cryogenic probes tunable to the ^{19}F resonance frequency, strengthen this NMR-based screening approach (Figure 8a).

The *n*-FABS (*n*-*f*luorine *a*toms for biochemical screening) approach is a functional NMR-based assay to study enzymatic reactions and obtain IC_{50} (concentration of inhibitor at which a 50% inhibition of the enzymatic reaction is reached) values. The approach uses $(CF_n)_m$-tags ($n = 3$, $m = 1, 2,$ …) on the enzyme substrates and ^{19}F NMR spectroscopy. As the fluorinated substrate is converted by the enzyme into another species, the change of the chemical microenvironment of the F-nucleus leads to a CS change. The enzymatic reaction is quenched after a defined delay by the addition of a denaturant, a chelating agent, or a strong inhibitor. For screening purposes, a reference sample without any test molecules is run representing 0% inhibition. Even multiple enzymes can be screened by the presented approach. This is of special interest if the selectivity of an inhibitor for a target enzyme is tested in the presence of another enzyme of the same family (Figure 8b).

2.4.2 The INPHARMA Method

The acronym INPHARMA stands for interligand nuclear Overhauser effects for pharmacophore mapping (Carlomagno, 2012; Orts, Griesinger, and Carlomagno, 2009; Sánchez-Pedregal *et al.*, 2005; Stauch, Orts, and Carlomagno, 2012). This is a rather novel approach that can be used to map the relative orientation of ligands competitively binding to a macromolecule. The approach relies on observing interligand, spin diffusion mediated, and transferred NOE data between two ligands L_A and L_B, which compete for the same binding pocket of a macromolecular receptor R (Figure 9). The two ligands are prepared in a buffered solution in a 10- to 50-fold excess compared to the receptor. Then, NOESY experiments are acquired with varying mixing times. During the mixing, the ligand L_A binds first to the binding pocket and its protons (H_{LA}) transfer their magnetization to the receptor protons (R). Assuming that the off-rate (k_{off}) of the RL_A complex is sufficiently large (k_{off} > 100–1000 s^{-1}), the second ligand L_B can bind to the same site within the same mixing time period. Then, a back transfer of the magnetization originally coming from ligand L_A can occur via the receptor protons on to ligand B's protons (H_{LB}) resulting in observable intermolecular peaks between the protons of ligands A and B, although ligands A and B were
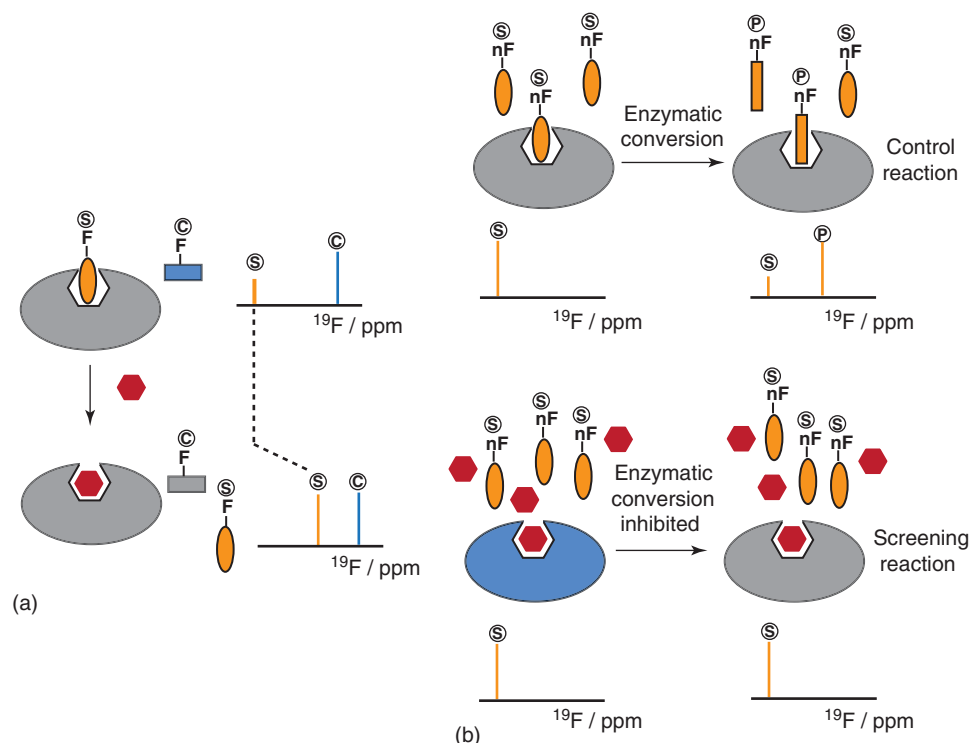
Figure 8 Principle of the FAXS and the *n*-FABS approach. (a) A weak affinity ^{19}F-modified spy molecule is used in a competition assay to identify high affinity protein ligands. (b) Functional screening of enzyme reactions. A fluorinated enzyme substrate is converted into a new species leading to a different ^{19}F-chemical shift signature. Enzyme inhibitors can then be identified by monitoring the ^{19}F-substrate resonance in the presence of the potential inhibitor small molecule.

never close in space during the NOE experiment. An important feature of this approach is the possibility to derive structural information on the binding modes of both ligands, as the relative orientation of the two ligands A and B in the receptor binding pocket is deducible by a thorough analysis of the interligand NOE pattern. The INPHARMA NOEs should not be mistaken for the interligand nuclear Overhauser effects (ILOES), which can occur when two ligands are simultaneously bound at the same binding site or at two nearby binding sites on the host molecule.

In the early stage, the INPHARMA method was used to determine the binding mode of ligand B, when the binding mode of ligand A was known. However, it soon became apparent that the approach is even more powerful and, in well-behaved ligand–host systems, a *de novo* description of the binding mode of both ligands A and B is amenable. This can be achieved using the INPHARMA NOEs

as an additional input data for a molecular docking procedure. To successfully apply the INPHARMA methodology, the most crucial experimental parameter, the length of the mixing step, has to be optimized. The value is predetermined by the molecular weight of the host molecule (i.e., its correlation time τ_c). As a rule of thumb, for a receptor molecule with a molecular weight larger than 400 kDa, the mixing time should be smaller than 100 ms, whereas for macromolecules comprising approximately 20 kDa, the mixing period should be higher than 500 ms. Another important parameter are the off-rates (k_{off}^A and k_{off}^B) of both ligands A and B. As soon as the ratio k_{off}^A/k_{off}^B gets larger than 10, the INPHARMA NOEs become very small. In an optimal case, both ligands roughly spend an equal period in the binding pocket (i.e., $k_{off}^A = k_{off}^B$).

A very impressive example using the INPHARMA approach to derive ligand binding modes is a
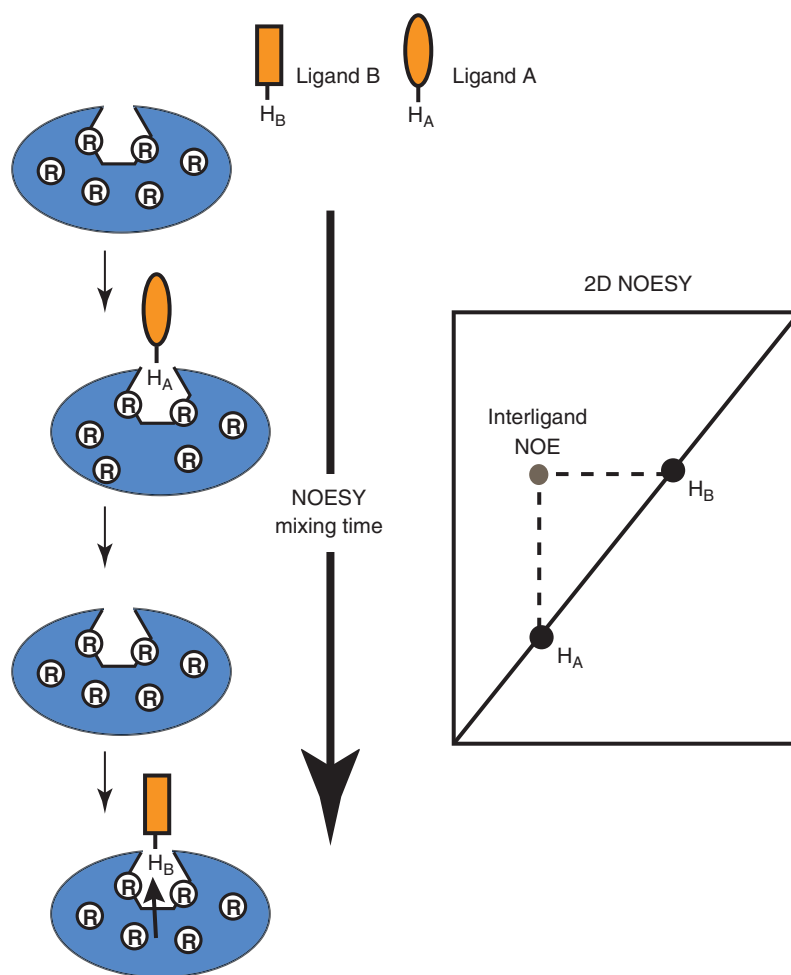
Figure 9 The INPHARMA method. At the beginning of the NOESY mixing time, ligand A binds to the receptor and transfers magnetization to the receptor protons. As ligand A has a low affinity to the host, a displacement of ligand A by B can occur and the "receptor-stored" magnetization of A can be back transferred to ligand B's protons. This results in a spin-diffusion-mediated interligand NOE in a 2D NOESY experiment.

study on tubulin and several natural products (Sánchez-Pedregal *et al.*, 2005). Tubulin is an important target in cancer research as this protein is involved in cell death. Natural products binding to tubulin often possess cytotoxic activity. Among these paclitaxel (PTX), a drug of the taxane family was found to be an effective cancer cell growth inhibitor and is currently widely used for the treatment of solid human cancers. It was shown that epothilones with more favorable pharmacological properties bind to the same binding pocket on tubulin as PTX. The INPHARMA approach was used to characterize the binding mode of epothilone A in this binding pocket. The NMR experiment was performed using a mixture of tubulin, baccatin III (BacIII, a compound structurally related to PTX), and epothilone A (EpoA) in a ratio $1:50:50$. Owing to the high molecular weight intense, receptor-mediated interligand NOEs were observed at a NOESY mixing times smaller than 100 ms. Subsequently, structural models for both complexes were generated and rated on the basis of the observed INPHARMA NOEs. The best-fit model found is in good agreement with SAR data available for EpoA analogs and tubulin mutants. The study further suggests that EpoA and PTX share a common binding mode to tubulin.

3 CONCLUSIONS

In this chapter, we discussed various aspects of biomolecular NMR spectroscopy and its application in the drug discovery process. Starting with a brief overview on the milestones in biomolecular NMR spectroscopy, we then focus on the use of solution NMR spectroscopy for drug screening purposes. The methods are well established for the investigation of proteins interacting with small molecules from various sources, for example, from plant extracts. Notwithstanding, RNAs were recently recognized as a "drugable" target as it gets more and more obvious that RNA is a key component in cellular processes involving among others gene regulation. We discuss similarities between NMR methods suitable to observe ligand–host interactions in both proteins and RNAs and point out important disparities that need to be considered to successfully apply protein NMR experiments for nucleic acids. Finally, we focus on recent developments in the field of NMR spectroscopy of protein–small molecule NMR spectroscopy, which allow functional screening assays based on ^{19}F magnetic resonance spectroscopy and make it possible to elucidate ligand binding modes in challenging protein–ligand systems.

In a nutshell, NMR spectroscopy of large molecules and the characterization of their interactions with small molecules have proved to be indispensable tools in drug discovery studies. Advances with respect to NMR hardware and method development will assure an important role of magnetic resonance in this scientific field.

REFERENCES

Aboul-ela, F., Karn, J. and Varani, G. (1995) *J. Mol. Biol.*, **253**(2), 313–332.

Akke, M., Fiala, R., Jiang, F., *et al.* (1997) *RNA*, **3**(7), 702–709.

Becattini, B. and Pellecchia, M. (2006) *Chem. – A Eur. J.*, **12**(10), 2658–2662.

Bertini, I, McGreevy, K. S. and Parigi, G. (2012) NMR and its place in mechanistic systems biology, in *NMR of Biomolecules*, eds. Bertini, I., McGreevy, K. S. and Parigi, G., Wiley-Blackwell, Weinheim, Germany.

Blount, K. F. and Breaker, R. R. (2006) *Nat. Biotechnol.*, **24**(12), 1558–1564.

Bothe, J. R., Nikolova, E. N., Eichhorn, C. D., *et al.* (2011) *Nat. Methods*, **8**(11), 919–931.

Breaker, R. R. (2012) *Cold Spring Harb. Perspect. Biol.*, **4**(2).

Brodsky, A. S. and Williamson, J. R. (1997) *J. Mol. Biol.*, **267**(3), 624–639.

Burgstaller, P. and Famulok, M. (1994) *Angew. Chem. Int. Ed.*, **33**(10), 1084–1087.

Carlomagno, T. (2005) *Annu. Rev. Biophys. Biomol. Struct.*, **34**(1), 245–266.

Carlomagno, T. (2012) *Nat. Prod. Rep.*, **29**(5), 536–554.

Carulla, N., Caddy, G. L., Hall, D. R., *et al.* (2005) *Nature*, **436**(7050), 554–558.

Chen, J., Zhang, Z., Stebbins, J. L., *et al.* (2007) *ACS Chem. Biol.*, **2**(5), 329–336.

Cochrane, J. C. and Strobel, S. A. (2008) *Acc. Chem. Res.*, **41**(8), 1027–1035.

Dalvit, C., Pevarello, P., Tato, M., *et al.* (2000) *J. Biomol. NMR*, **18**(1), 65–68.

Dalvit, C., Ardini, E., Flocco, M., *et al.* (2003a) *J. Am. Chem. Soc.*, **125**(47), 14620–14625.

Dalvit, C., Fagerness, P. E., Hadden, D. T. A., *et al.* (2003b) *J. Am. Chem. Soc.*, **125**(25), 7696–7703.

Fan, P., Suri, A. K., Fiala, R., *et al.* (1996) *J. Mol. Biol.*, **258**(3), 480–500.

Fauster, K., Kreutz, C. and Micura, R. (2012) *Angew. Chem. Int. Ed.*, **51**(52), 13080–13084.

Fejzo, J., Lepre, C. A., Peng, J. W., *et al.* (1999) *Chem. Biol.*, **6**(10), 755–769.

Fernandez, C. and Jahnke, W. (2004) *Drug Discovery Today: Technol.*, **1**(3), 277–283.

Fourmy, D., Recht, M. I., Blanchard, S. C., *et al.* (1996) *Science*, **274**(5291), 1367–1371.

Frydman, L. (2011) *J. Magn. Reson.*, **213**(2), 213.

Fürtig, B., Buck, J., Manoharan, V., *et al.* (2007) *Biopolymers*, **86**(5–6), 360–383.

Gerig, J. T. (1994) *Prog. Nucl. Magn. Reson. Spectrosc.*, **26**(Part 4), 293–370.

Golden, B. L., Gooding, A. R., Podell, E. R., *et al.* (1998) *Science*, **282**(5387), 259–264.

Hajduk, P. J., Olejniczak, E. T. and Fesik, S. W. (1997) *J. Am. Chem. Soc.*, **119**(50), 12257–12261.

Hajduk, P. J., Gerfin, T., Boehlen, J.-M., *et al.* (1999) *J. Med. Chem.*, **42**(13), 2315–2317.

Haller, A., Souliere, M. F. and Micura, R. (2011) *Acc. Chem. Res.*, **44**(12), 1339–1348.

Hennig, M., Scott, L. G., Sperling, E., *et al.* (2007) *J. Am. Chem. Soc.*, **129**(48), 14911–14921.

Huettenhofer, A., Schattner, P. and Polacek, N. (2005) *Trends Genet.*, **21**(5), 289–297.

Hurd, R. E. (2011) *J. Magn. Reson.*, **213**(2), 467–473.

Jahnke, W., Perez, L. B., Paris, C. G., *et al.* (2000) *J. Am. Chem. Soc.*, **122**(30), 7394–7395.

Jahnke, W., Ruedisser, S. and Zurini, M. (2001) *J. Am. Chem. Soc.*, **123**(13), 3149–3150.

Jiang, L. and Patel, D. J. (1998) *Nat. Struct. Mol. Biol.*, **5**(9), 769–774.

Js, R.-C., Vuister, G. W., Moonen, C. T. W., *et al.* (2011) *J. Magn. Reson.*, **213**(2), 446–466.

Kay, L. E. (2011) *J. Magn. Reson.*, **213**(2), 477–491.

Kay, L. E., Ikura, M., Tschudin, R., *et al.* (2011) *J. Magn. Reson.*, **213**(2), 423–441.

Kloiber, K., Spitzer, R., Tollinger, M., *et al*. (2011) *Nucleic Acids Res.*, **39**(10), 4340–4351.

Kreutz, C. and Micura, R. (2008) Investigations on fluorine-labeled ribonucleic acids by ¹⁹F NMR spectroscopy, in *Modified Nucleosides*, ed. Herdewijn, P., Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, pp. 1–27.

Kreutz, C., Kaehlig, H., Konrat, R., *et al*. (2005) *J. Am. Chem. Soc.*, **127**(33), 11558–11559.

Kreutz, C., Kählig, H., Konrat, R., *et al*. (2006) *Angew. Chem. Int. Ed.*, **45**(21), 3450–3453.

Kulshina, N., Baird, N. J. and Ferre-D'Amare, A. R. (2009) *Nat. Struct. Mol. Biol.*, **16**(12), 1212–1217.

Kumar, A., Ernst, R. R. and Wuethrich, K. (1980) *Biochem. Biophys. Res. Commun.*, **95**(1), 1–6.

Liberman, J. A. and Wedekind, J. E. (2012) *Wiley Interdiscip. Rev.: RNA*, **3**(3), 369–384.

Lipari, G. and Szabo, A. (1982a) *J. Am. Chem. Soc.*, **104**(17), 4559–4570.

Lipari, G. and Szabo, A. (1982b) *J. Am. Chem. Soc.*, **104**(17), 4546–4559.

Lipsitz, R. S. and Tjandra, N. (2004) *Annu. Rev. Biophys. Biomol. Struct.*, **33**(1), 387–413.

Lombès, T., Moumné, R., Larue, V., *et al*. (2012) *Angew. Chem. Int. Ed.*, **51**(38), 9530–9534.

Lu, K., Miyazaki, Y. and Summers, M. (2010) *J. Biomol. NMR*, **46**(1), 113–125.

Lu, K., Heng, X. and Summers, M. F. (2011) *J. Mol. Biol.*, **410**(4), 609–633.

Mayer, G. (2009) *Angew. Chem. Int. Ed.*, **48**(15), 2672–2689.

Mayer, M. and James, T. L. (2002) *J. Am. Chem. Soc.*, **124**(45), 13376–13377.

Mayer, M. and James, T. L. (2004) *J. Am. Chem. Soc.*, **126**(13), 4453–4460.

Mayer, M. and James, T. L. (2005) Discovery of ligands by a combination of computational and NMR-based screening: RNA as an example target, in *Methods in Enzymology*, ed. Thomas, L. J., Academic Press, San Diego, CA, pp. 571–587.

Mayer, M. and Meyer, B. (1999) *Angew. Chem. Int. Ed.*, **38**(12), 1784–1788.

Mayer, M., Lang, P. T., Gerber, S., *et al*. (2006) *Chem. Biol.*, **13**(9), 993–1000.

Milligan, J. F. and Uhlenbeck, O. C. (1989) *Methods Enzymol.*, **180**, 51–62.

Milligan, J. F., Groebe, D. R., Witherell, G. W., *et al*. (1987) *Nucleic Acids Res.*, **15**(21), 8783–98.

Mittermaier, A. and Kay, L. E. (2006) *Science*, **312**(5771), 224–228.

Mittermaier, A. K. and Kay, L. E. (2009) *Trends Biochem. Sci.*, **34**(12), 601–611.

Moore, P. B. and Steitz, T. A. (2003) *Annu. Rev. Biochem.*, **72**(1), 813–850.

Moumne, R., Pasco, M., Prost, E., *et al*. (2010) *J. Am. Chem. Soc.*, **132**(38), 13111–13113.

Muchmore, D. C., McIntosh, L. P., Russell, C. B., *et al*. (1989) Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance, in *Methods in Enzymology*, eds. Norman, J. O. and Thomas, L. J., Academic Press, San Diego, CA, pp. 44–73.

Narberhaus, F., Waldminghaus, T. and Chowdhury, S. (2006) *FEMS Microbiol. Rev.*, **30**(1), 3–16.

Nussinov, R. and Wright, P. E. (2009) *Nat. Chem. Biol.*, **5**(11), 789–796.

Orts, J., Griesinger, C. and Carlomagno, T. (2009) *J. Magn. Reson.*, **200**(1), 64–73.

Palmer, A. G. 3rd, Kroenke, C. D. and Loria, J. P. (2001) *Methods Enzymol.*, **339**, 204–38.

Pellecchia, M., Sem, D. S. and Wuthrich, K. (2002) *Nat. Rev. Drug Discov.*, **1**(3), 211–219.

Pellecchia, M., Bertini, I., Cowburn, D., *et al*. (2008) *Nat. Rev. Drug Discov.*, **7**(9), 738–745.

Pervushin, K., Riek, R., Wider, G., *et al*. (1997) *Proc. Natl. Acad. Sci.*, **94**(23), 12366–12371.

Puffer, B., Kreutz, C., Rieder, U., *et al*. (2009) *Nucleic Acids Res.*, **37**(22), 7728–7740.

Ramakrishnan, V. (2002) *Cell*, **108**(4), 557–572.

Roberts, G. C. K. (1999) *Curr. Opin. Biotechnol.*, **10**(1), 42–47.

Sánchez-Pedregal, V. M., Reese, M., Meiler, J., *et al*. (2005) *Angew. Chem. Int. Ed.*, **44**(27), 4172–4175.

Scott, L. G., Geierstanger, B. H., Williamson, J. R., *et al*. (2004) *J. Am. Chem. Soc.*, **126**(38), 11776–11777.

Serganov, A. and Patel, D. J. (2012) *Curr. Opin. Struct. Biol.*, **22**(3), 279–286.

Shuker, S. B., Hajduk, P. J., Meadows, R. P., *et al*. (1996) *Science*, **274**(5292), 1531–1534.

Smith, K. D., Lipchock, S. V., Ames, T. D., *et al*. (2009) *Nat. Struct. Mol. Biol.*, **16**(12), 1218–1223.

Stauch, B., Orts, J. and Carlomagno, T. (2012) *J. Biomol. NMR*, **54**(3), 245–256.

Steitz, T. A. and Moore, P. B. (2003) *Trends Biochem. Sci.*, **28**(8), 411–418.

Stoltenburg, R., Reinemann, C. and Strehlitz, B. (2007) *Biomol. Eng.*, **24**(4), 381–403.

Sudarsan, N., Lee, E. R., Weinberg, Z., *et al*. (2008) *Science*, **321**(5887), 411–413.

Tjandra, N. and Bax, A. (1997) *Science*, **278**(5340), 1111–1114.

Tolman, J. R., Flanagan, J. M., Kennedy, M. A., *et al*. (1997) *Nat. Struct. Mol. Biol.*, **4**(4), 292–297.

Tucker, B. J. and Breaker, R. R. (2005) *Curr. Opin. Struct. Biol.*, **15**(3), 342–348.

Turner, B. G. and Summers, M. F. (1999) *J. Mol. Biol.*, **285**(1), 1–32.

Vulpetti, A. and Dalvit, C. (2012) *Drug Discov. Today*, **17**(15,Äì16), 890–897.

Williamson, M. P., Havel, T. F. and Wuethrich, K. (1985) *J. Mol. Biol.*, **182**(2), 295–315.

Woodson, S. A. (2005) *Curr. Opin. Struct. Biol.*, **15**(3), 324–330.

Wu, H. and Feigon, J. (2007) *Proc. Natl. Acad. Sci.*, **104**(16), 6655–6660.

Wuethrich, K., Wider, G., Wagner, G., *et al*. (1982) *J. Mol. Biol.*, **155**(3), 311–319.

Wunderlich, C. H., Spitzer, R., Santner, T., *et al*. (2012) *J. Am. Chem. Soc.*, **134**(17), 7558–7569.

Wüthrich, K. (1986) *NMR of proteins and nucleic acids*, John Wiley and Sons.

Yan, J., Kline, A. D., Mo, H., *et al*. (2003) *J. Magn. Reson.*, **163**(2), 270–276.

Zhang, Q., Sun, X., Watt, E. D., *et al*. (2006) *Science*, **311**(5761), 653–656.

On-Line and At-Line LC-NMR and Related Micro-NMR Methods

Nadine Bohni, Emerson F. Queiroz and Jean-Luc Wolfender

School of Pharmaceutical Sciences, University of Geneva, University of Lausanne, Geneva, Switzerland

1 INTRODUCTION

Metabolite profiling by different high performance liquid chromatography (HPLC) techniques (see **HPLC and Ultra HPLC: Basic Concepts**), most notably HPLC-mass spectrometry (HPLC-MS, see **LC and LC-MS: Techniques and Applications**) and HPLC-photo diode array (HPLC-PDA) detection or a combination of these techniques in HPLC-PDA-MS, are extremely useful tools for the early metabolite identification of natural products (NPs) on-line in very complex natural matrices. Recently, this process has been significantly improved by the increase in the use of high resolution (HR) mass spectrometers, such as time-of-flight instruments (time-of-flight mass spectrometry, TOFMS) in the hyphenation with LC (Eugster and Wolfender, 2012). The high mass and spectral accuracies provided by LC-HRMS instruments offer now unambiguous molecular formula assignment when used with an adapted heuristic filtering procedure (Kind and Fiehn, 2007). This represents key steps for dereplication because NP research molecular formulae can be used in combination with chemotaxonomic information to generate putative structure attribution to a given LC peak (Funari *et al.*, 2012).

Because of the lack of generic databases [especially LC–MS and LC–tandem mass spectrometry (LC-MS/MS) databases] that could provide efficient early metabolite identification of previously reported NPs, the identification of NPs in metabolite profiling studies based only on HRMS (high resolution mass spectrometry) and MS/MS in combination with PDA spectra is not sufficient to unambiguously ascertain the structure of the analytes of interest. Recently, efforts regarding the prediction of retention time as support for dereplication (Creek *et al.*, 2011) have also been made but none of these LC approaches can claim a complete on-line structural identification without access to standards or specific databases for confirmation of the metabolite identification.

In this respect, metabolite identification must rely on nuclear magnetic resonance (NMR, see **NMR of Small Molecules**), which is known to provide key structural information on atom connectivities that are complementary to those provided by MS, such as the molecular formula. NMR is thus known to be an essential tool for the *de novo* identification of NPs, whereas structure identification relies on the use of exhaustive one-dimensional (1D) and two-dimensional (2D) homo- and heteronuclear NMR measurements (Sturm and Seger, 2012). Over the past three decades, NMR hardware and pulse sequences (Reynolds and Enriquez, 2002) have evolved considerably, especially the development of 2D NMR experiments, which has provided an efficient way to fully characterized NPs according

to established schemes of spectral interpretation (Breton and Reynolds, 2013). One-dimensional proton NMR (^{1}H NMR) spectra are characteristic for a given compound, and the comparison of such data with those of previously reported compounds is key for the rapid dereplication of NPs.

For all these reasons and also because it is a highly nonselective detection technique, (^{1}H NMR spectroscopy will detect any hydrogen-containing compound present in the HPLC eluent in a sufficient amount, regardless of its structure), NMR can be considered the ideal detector for HPLC hyphenation (Jaroszewski, 2005b; Wolfender, 2010). Coupling HPLC with NMR spectroscopy (LC-NMR) represents one of the most powerful methods for the separation and structural elucidation of NPs in mixtures (Albert, 2002). Since the mid-1990s, LC-NMR has been established as a very efficient method for the on-line identification of organic molecules, and it was readily applied to the structural identification of NPs in crude plant extracts. The on-line coupling of HPLC with NMR, however, suffers from the very low intrinsic sensitivity of NMR. In addition, as discussed in this chapter, various strategies to provide high quality NMR spectra in hyphenation with HPLC, either on-line or at-line, are required for both efficient LC peak identification and absolute quantification purposes. In light of the needs of phytochemical analysis and plant metabolomics (see **Metabolomics**) for both on-line identification and dereplication purposes, the possibilities and limitations of these approaches are discussed in this chapter.

1.1 Sensitivity and NMR Probe Design

Among all detectors for HPLC, NMR is by far the least sensitive. The improvements in sensitivity and resolution that have been achieved in NMR methods in recent years have strongly accelerated the pace at which NPs are identified. NMR magnets have steadily evolved in field strength and consoles in electronics, and over the past few years, in the development of more sensitive probe heads, such as microcoils or cryogenically cooled probe heads. Furthermore, the application of pulsed field gradients has resulted in significant changes (Bross-Walch *et al*., 2005).

1.1.1 Probe and Coil Design of NMR Flow Cells

From a historical viewpoint, the interest in combining separation methods with ^{1}H NMR spectroscopy arose in the late 1970s (Bayer *et al*., 1979). Because of the inherent lack of sensitivity of NMR instruments at that time, the use of LC-NMR to solve analytical problems was delayed by almost two decades (Jaroszewski, 2005b).

The introduction of LC-NMR as a detector for HPLC has to firstly rely on the development of flow probes through which the eluent from the HPLC has to enter, which enables the introduction of the analytes of interest as they are separated by HPLC. These on-flow NMR probes (or continuous-flow probes) consisted of a U-type glass tube fixed in the dewar of the NMR probe body. This design ("saddle"-shaped geometry) was breaking the central symmetry of the conventional tubes setup. Even if no rotation could be applied in such a flow cell, the first application showed excellent resolution, approaching those registered with rotation of the NMR tube (Albert, 1995).

For on-line applications of LC-NMR, the volume of the described flow cell represented a compromise between the needs of chromatography and those of NMR. Indeed, this type of cell employs detection volumes between 40 and 120 μL, which are much larger than conventional UV detection cells (8 μL). LC-NMR is a volume-sensitive detection technique that requires the maximization of NMR-active nuclei by extension of the detection volume that has to match, as closely as possible, that of the elution volume of the HPLC peaks. Furthermore, the line shape in on-flow mode is directly related to the residence time of the analytes in the flow cell, and the measured signal half width increases with increasing flow rate (Albert, 2002).

As will be shown, the development of such a flow probe has permitted the direct practical detection of HPLC peaks by NMR and many applications to NP analysis have demonstrated the usefulness of this type of hyphenation. However, the efficiency of detection of such on-flow probes is optimized when the sample can be preconcentrated using solid-phase extraction (SPE) and eluted in deuterated solvents in the LC-SPE-NMR mode with an optimized elution volume to match that of the flow probe.

The necessity for compromise between the volume of the flow cell and the volume of the LC peak and

the necessity for solvent suppression and other issues described in the following sections have accelerated the development of other probes for at-line NMR measurement of microquantities of analytes collected from HPLC.

In this respect, reduced-diameter solenoid NMR coils represent a particularly attractive approach to enhancing NMR sensitivity for small-volume, mass-limited sample uses (Figure 1b) (Olson *et al.*, 2004). The sensitivity of an NMR experiment increases in inverse proportion to the diameter of the coil (Hoult and Richards, 1976). Thus, for mass-limited samples, such as HPLC peaks, the best sensitivity arises from samples that can be concentrated into a small volume and measured with a small-diameter probe. For such small-diameter probes, solenoidal coils exhibit a several-fold enhancement in NMR sensitivity, when compared to that of the Helmholtz coils (Figure 1b) as discussed in Section 3.
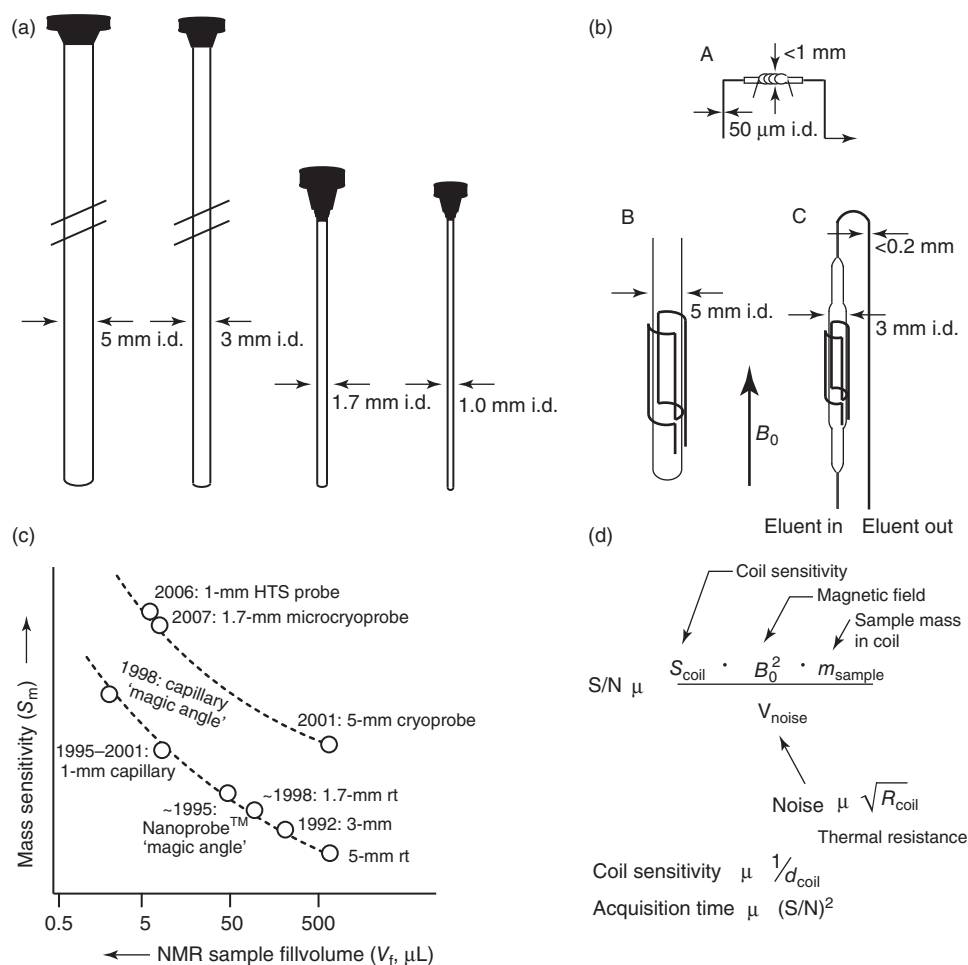
Figure 1   Comparison of different NMR tubes and NMR probe designs for the understanding of NMR sensitivity. (a) On-scale illustration of the inner diameter of different NMR tubes ranging from the conventional 5-mm tube to the smallest available 1-mm microtube. (b) Illustration of a microflow NMR probe with a solenoid coil that has a sample volume of 1.5 μL (A), a conventional 5-mm saddle-type coil NMR probe (B), and a 3-mm NMR flow probe with a saddle-type coil (C). (c) Approximate time line and qualitative comparison of recent milestone NMR probe innovations. Mass sensitivity (S_m, linear "*y*"-scale) for ^{1}H NMR of a hypothetical fixed-mass sample as a function of probe fill volume (V_f, note the logarithmic "*x*"-scale) for room-temperature probes (lower line) and cryoprobes (upper line) at a fixed field, B_0. (*Source*: Adapted from Molinski TF. 2010. NMR of natural products at the 'nanomole-scale'. Nat Prod Rep 27: 321–329, with permission of The Royal Society of Chemistry (http://dx.doi.org/10.1039/B920545B).) (d) Mathematic relations of probe design parameters and NMR sensitivity expressed as S/N.

Solenoid probes are naturally suited for LC-NMR because samples need to be added by flow, which can be manual or interfaced to an LC system. This approach has resulted in a commercially available solenoid probe called the *CapNMR™ probe* (Norcross *et al.*, 2010), which can be simultaneously tuned to multiple frequencies, allowing for near-optimal performance on different channels (Li *et al.*, 2003). These small solenoid coils can also be added as multiple coils in the same probe head, allowing parallel detection and thus a higher analyte throughput (Gökay and Albert, 2012).

1.1.2 Detection Issues

Other than the design of the flow probes, another issue with HPLC-NMR hyphenation is the compatibility of the solvent with NMR detection. In NMR, deuterated solvents are used to record the signal of the analytes present in much smaller amounts than those of the solvent. In HPLC, the analytes are typically separated in an acetonitrile and water mixture because most separations are carried out by reversed-phase (RP) chromatography (Wolfender, 2009). The use of nondeuterated solvents represents a challenging issue for NMR detection because the signals related to these solvents will be several orders of magnitude more intense than those of the analytes of interest. This problem has been overcome by the development of fast, reliable, and powerful solvent suppression techniques such as presaturation [nuclear Overhauser enhancement spectroscopy (NOESY) presaturation], soft pulse multiple irradiation, or WET (water suppression enhanced through T_1 effects). The WET sequence, in particular, consists of a combination of pulsed field gradients, shaped radio-frequency (rf) pulses, and shifted laminar pulses that can be used with carbon-13 (^{13}C) decoupling for efficient removal of the ^{13}C satellites within the solvent (Smallcombe, Patt, and Keiffer, 1995). This type of solvent suppression provides good cancellation of the solvent signal and multiple lines can be suppressed as shown for the analysis of ion-pair chromatography, where many lines must be removed (Ramm *et al.*, 2004). This solvent suppression sequence provides a very selective and efficient suppression; therefore, the maximum receiver gain can be used to maximize the detection of the analyte signals. The analyte signals would otherwise not be detectable in the spectra without suppression because of the restricted dynamic range of the NMR.

However, one main issue is that NMR signals of analytes of interest resonate at the same frequency as acetonitrile (MeCN) or methanol (MeOH) will also be suppressed (Wolfender, 2010). To minimize these effects and to obtain better quality spectra, the water is typically replaced by deuterated water in LC-NMR because it is not too costly. Some authors have also performed the entire LC-NMR separation using fully deuterated solvents. This can notably increase the cost of analysis when standard HPLC flow rates (1 mL min^{-1}) are used. However, the use of fully deuterated solvent in capillary HPLC application is perfectly compatible because of the low solvent consumption of such systems (Xiao *et al.*, 2005). Furthermore, the continuous flow of sample in the detector coil complicates solvent suppression. To ensure satisfactory suppression, these pulse sequences require knowledge of the frequency of the solvent signals during HPLC gradient elution. Moreover, this requires that an initial single-transient NMR spectrum is acquired automatically before solvent suppression and that the frequency for solvent suppression is automatically determined for each time point in the HPLC run. An example of the efficiency of solvent suppression is shown in Figure 3a and b.

As discussed later, other ways of coupling NMR to HPLC, such as SPE-NMR or at-line micro-NMR methods, exist that circumvent such problems and permit recording of spectra using fully deuterated solvents.

1.2 On-Line Versus At-Line Approaches for Obtaining NMR Spectra from an Analyte of Interest within a Mixture

Because NMR sensitivity represents a key issue, different chromatographic strategies and/or modes of operation have been used to preconcentrate the maximum amount possible of an analyte for NMR detection (Figure 2). The goal in such approaches is to place the highest possible number of NMR-active nuclei in the active volume of the detection probe. In direct hyphenation of LC-NMR, this can be performed by increasing the sample loading in HPLC; however, this might results in a loss of chromatographic resolution. This is acceptable only to a certain extent, depending on the chromatographic resolution required for the given analyses.
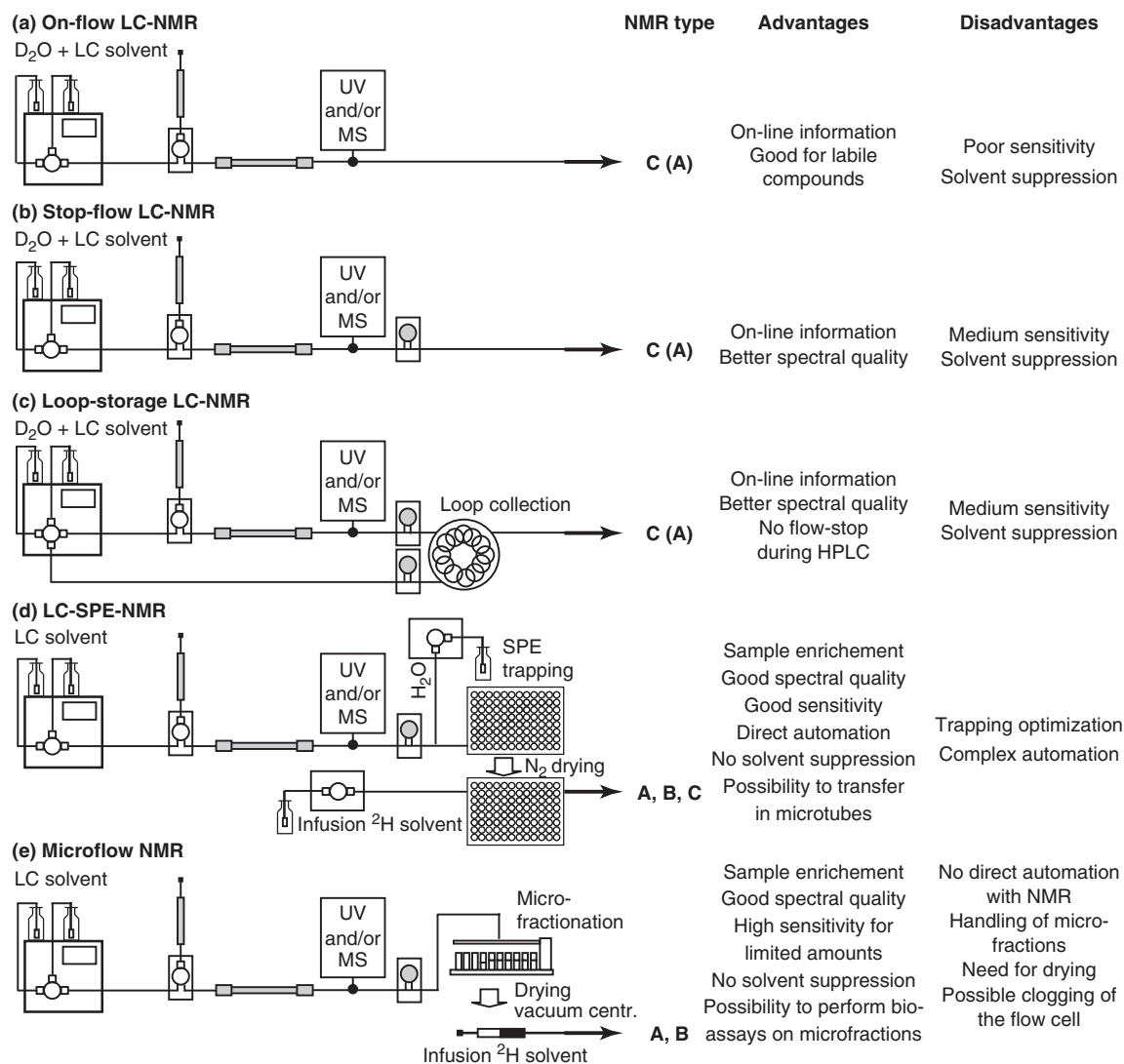
Figure 2 Modes of operation and advantages and disadvantages of the on-line and at-line LC-NMR approaches. Various modes of operation of liquid chromatography-nuclear magnetic resonance (LC-NMR) using direct (a–c) and indirect (d and e) hyphenation techniques. Indication of NMR probe type to be used with the different hyphenation approaches (A: microflow probe with solenoid coil; B: tube probe with saddle-type coil; and C: flow probe with saddle-type coil, see also Figure 1). Advantages and disadvantages of the different LC-NMR techniques. (*Source*: Adapted with permission of Taylor and Francis Group LLC Books, from LC-NMR and related techniques for the rapid identification of plant metabolites, in High Performance Liquid Chromatography in Phytochemical Analysis, eds. M. Waksmundzka-Hajnos and J. Sherma, 2010; permission conveyed through Copyright Clearance Center, Inc.)

Therefore, in relation to the complexity of the mixture to be analyzed and the concentration of the LC peak required to be identified, LC-NMR can be operated in direct hyphenation or analytes may be preconcentrated before NMR detection with approaches such as SPE-NMR, microfractionation, drying and postchromatographic analysis in microflow NMR probes, or through the use of sensitive NMR probes with microtubes (Figure 2). The postchromatographic analyses of microfractions after HPLC profiling require more automation or handling than direct hyphenation; however, the quality and the sensitivity of the NMR spectra can be better optimized.

In this chapter, LC-NMR methods will be divided into two categories. The first type will group all strategies that provide spectra from NMR using direct hyphenation with HPLC with or without preconcentration. In the second category, applications of micro-NMR methods to microfractions obtained after HPLC separation either by flow probes or by microtubes are discussed.

2 DIRECT HYPHENATION OF HPLC WITH NMR

Direct hyphenation of HPLC and NMR allows the acquisition of NMR spectra either directly from the LC peaks as they elute from the HPLC column (on-flow mode) or by stopping the flow (stop-flow mode) using techniques such as collecting the peaks in loops (loop-storage mode) or trapping the peaks on SPE cartridges followed by further elution (SPE-NMR) (Figure 2).

All of these modes of operation will be considered direct hyphenation of HPLC to NMR because an NMR instrument is physically connected to the HPLC even if the LC peaks are brought to the flow probe using automated enrichment procedures.

For all of these modes of operation, a typical LC-NMR setup consists of an HPLC system equipped with a sensitive UV or MS detector coupled to an NMR instrument (medium to high field strength: >400 MHz), which is equipped with a flow probe (Wolfender *et al*., 2013). In this case, the LC-NMR flow probe of the saddle type employs detection volumes between 40 and 120 μL (Figure 1). With such volumes, cryogenized flow probes are also available (Sturm *et al*., 2007) that can provide better signal-to-noise (S/N) ratios. These large detection volumes are required for satisfactory line shape quality and sensitivity but may be detrimental to the quality of the chromatographic separation (Albert, 2002).

2.1 LC-NMR and LC-SPE-NMR

2.1.1 On-Flow and Stop-Flow LC-NMR

Acquiring NMR spectra from eluting LC peaks in the on-flow mode is rather simple because no specific automation between the HPLC and the NMR detector is required (Figure 2a).

The separations are typically run using an RP gradient, and the water is replaced by deuterated water (D_2O) to enhance spectral quality. Because the organic modifier is not deuterated, solvent suppression is performed *in situ*, and the frequencies are calculated based on a single scan (scout scan), which precedes the set of scans that will be acquired to create an increment. Thus, ^1H NMR spectra are continuously recorded as a set of scans (transients) in discrete increments and generate the on-flow LC-NMR plot. Typically, and depending on the flow rate used [1 mL min^{-1} down to 0.1 mL min^{-1} (Queiroz *et al*., 2002)], scans will be accumulated as sets of 16–128 transients for most applications. Because a single scan will take roughly 1 s, the frequency of acquisition in on-flow operation will be between 20 s and 2–3 min.

In on-flow LC-NMR, these slow acquisition frequencies on the detector side will typically yield between one and four NMR spectra over a given LC-NMR peak. This will lead to very low chromatographic resolution compared to LC-MS detection, where frequencies can reach 100 ms or less. The on-flow LC-NMR is not sensitive because of compromises related to the flow cell volume and the line shape, and the general spectral qualities are not optimum because of the short residence time in the probe cell and because of continuous changes in the solvent composition related to the use of chromatographic gradients.

This acquisition mode is, however, very useful for obtaining a rapid overview of the main constituents in an extract and for gaining a general idea of their quantities because NMR is an absolute quantification method.

The results of these on-flow runs are displayed in the form of a 2D plot with NMR frequencies (in parts per million) on one axis and the time corresponding to the chromatographic separation on the other. A typical LC-NMR 2D plot of a crude plant extract is displayed in Figure 3a, where the dichloromethane extract of *Potamogeton lucens* L. (Potamogetonaceae) was analyzed. In this case, the LC-NMR analysis was performed at a rather low flow rate of 0.15 mL min^{-1}, optimizing the number of scans per increment to increase sensitivity (the number of transients is 256). As shown on the 2D plot, several NMR regions of the main peaks could be highlighted, such as the signals of the furano moiety and the exocyclic methylene characteristics of the detected furano-*ent*-labdane (Figure 3a). The corresponding LC ^1H NMR spectrum of one of
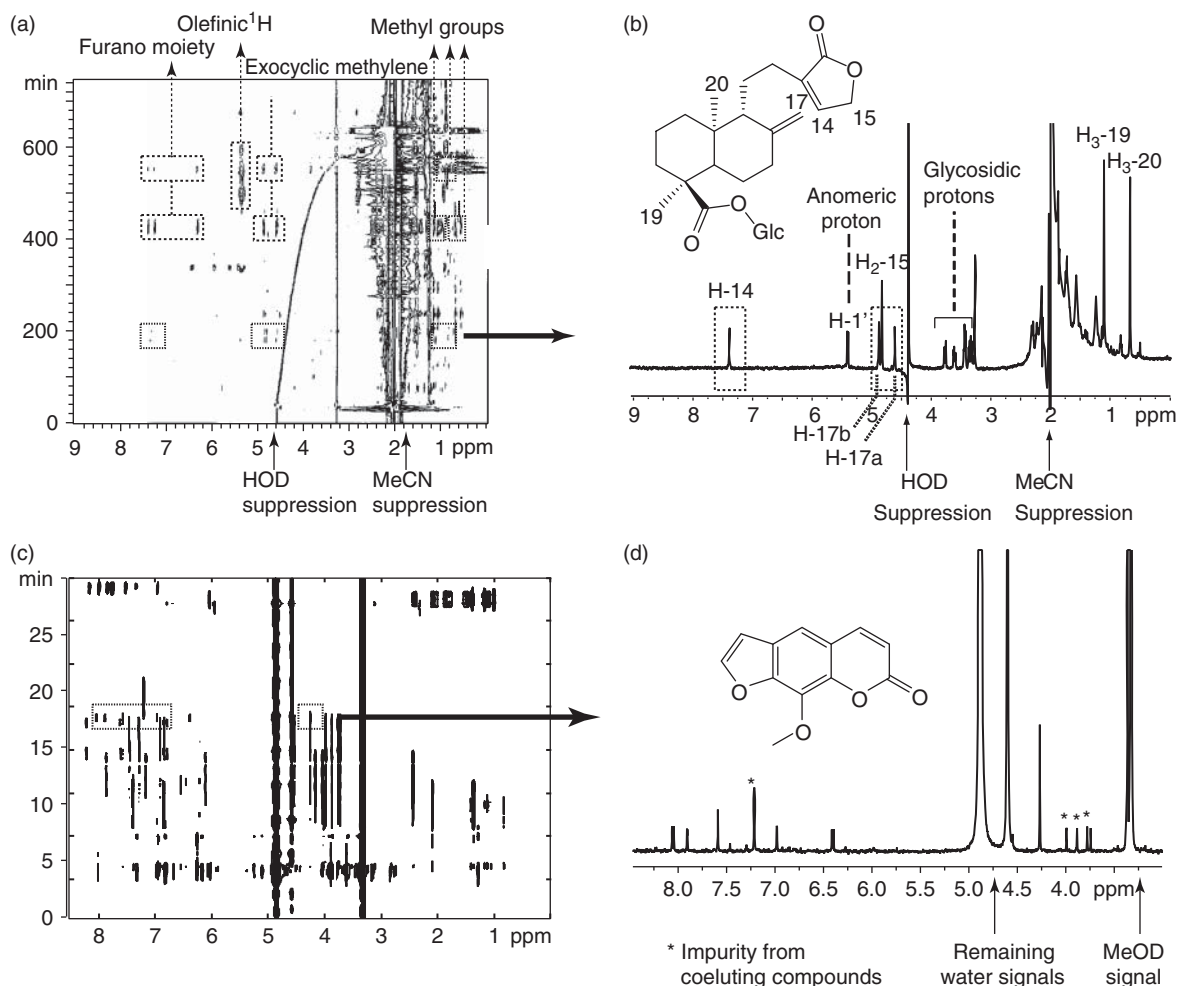
Figure 3   On-line LC-NMR versus at-line LC-NMR. (a) On-flow LC-UV-^{1}H NMR analysis of the dichloromethane extract of *Potamogeton lucens* (10 mg injected). Flow rate: 0.15 mL min^{-1}, 256 scans increment^{-1}. (b) The ^{1}H NMR spectrum of **1** was recorded in stop-flow mode with 512 scans (acquisition time 180 min). *Source*: P. Waridel, et al (2004). Adapted from Elsevier. (c) At-line LC-UV-NMR analysis of a mixture of 19 natural products of diverse polarity (50 mg injected). Flow rate: 4.7 mL min^{-1}. Fractions were collected every 20 s yielding 80 fractions in 30 min. (d) The ^{1}H NMR spectrum of every fraction was recorded with 128 scans using automated injection from a 96-well plate with the OMNMR™ system. The spectrum of xanthotoxin is given as an example, the peak at 7.2 ppm belongs to the partially coeluting naphthazarin, and the methoxy signals are from the partially coeluting papaverine. (*Source*: Data and figures were kindly provided by Samuel Bertrand (Azzollini et al., 2012). Reproduced with permission.)

the main LC peaks is shown in Figure 3b. This spectrum was acquired using the stop-flow mode to improve sensitivity and solvent suppression. On the 2D on-flow plot, at approximately 2 ppm, the residual signals of MeCN were still visible after solvent suppression, and the line corresponding to the HOD shifted with the gradient because of the change in solvent composition (Figure 3a).

In this mode, the quality of the spectra obtained is not optimum, and only 1D ^{1}H NMR spectra can be obtained using this technique; however, other nuclei, such as ^{19}F, have been efficiently monitored by this technique (Shockcor *et al.*, 2000). This acquisition mode is the only one that provides true on-line detection, which has several advantages if unstable products have to be studied (Cogne *et al.*, 2005).

In this case, long acquisitions can be performed, enhancing the S/N ratio and 2D NMR spectra, which allow mainly ^1H–^1H correlations [correlation spectroscopy (COSY) or total correlation spectroscopy (TOCSY)] to be recorded. To optimize the efficiency of the stop-flow measurement, a key aspect is to have a sensitive detector (UV or MS) before the NMR, which allows triggering of a valve (the stop valve) that will halt the HPLC eluent exactly when the analyte is passing through the NMR flow cell (Figure 2b). This requires that a well-calibrated delay between the first detector and the NMR is determined. In the stop-flow mode, the quality of solvent suppression is generally superior to the on-flow mode because the eluent peak is measured in a fix composition. A typical stop-flow spectrum is shown in Figure 3b. However, one drawback is that if many peaks in an extract have to be analyzed during long 2D NMR experiments, the different stops will affect the quality of the LC separation especially if isocratic separations are performed.

To avoid this, the flow rate is stopped for the analysis of each LC peak during the stop-flow mode. It is also possible to collect the peak of interest in loops that have a volume that matches that of the LC-NMR flow cell (loop-storage mode) (Figure 2c). In this case, a valve will trigger the collection of the LC peak in separate loops during the separation of the mixture. Each peak will then be analyzed postchromatographically (Corcoran and Spraul, 2003). This mode is more practical than the stop-flow mode but requires more automation, whereas the sensitivities of the two techniques remain comparable (Wolfender, 2010).

For all of these on-line approaches, analytical HPLC columns providing relatively high loading capacity are commonly used. These C_{18} columns have a large diameter, up to 10 mm, a long length (e.g., 250 mm), or both (Challal *et al.*, 2012). In the RP mode, H_2O is replaced by D_2O to facilitate solvent suppression. The crude extracts or fractions are dissolved in an appropriate deuterated solvent (DMSO-d_6 or CD_3OD) and relatively large volumes are injected. When complex mixtures must be analyzed, 20–50 mg may be injected onto the column. Such high loadings may deteriorate the resolution of the LC separation when compared to standard analytical HPLC conditions; however, these separation conditions provide reasonable NMR detection of the most abundant constituents of a mixture (Wolfender *et al.*, 2013).

2.1.2 LC-SPE-NMR

One of the major drawbacks of the direct hyphenation of HPLC to NMR is that the sensitivity is not optimum because the LC peaks are not very concentrated and elute in volumes that often exceed the volume of the flow cell. Furthermore, the chromatographic resolution is often compromised because of the need for high loading of the samples to achieve satisfactory sensitivity for the detection of the main compounds in the mixture. Another issue is that the spectra are recorded in a mixture of HPLC grade MeOH or MeCN with D_2O. This requires that efficient solvent suppression sequences are used, which might also suppress analyte signals. The NMR shifts also depend on the nature of the solvent and spectra recorded in the HPLC eluent of various compositions because the use of elution gradients will cause the comparisons of the chemical shifts with those of previously reported compounds in standard deuterated solvents to be difficult.

To overcome these problems, efficient methods that enable preconcentration of the analytes before detection in the LC-NMR flow cell and further elution in a fully deuterated solvent have been developed. To date, the most used technique is LC-SPE-NMR (Jaroszewski, 2005a; Sturm and Seger, 2012; Wilson *et al.*, 2007) and the hyphenation of HPLC to NMR via SPE trapping is now fully automated through state-of-the-art NMR setups (Corcoran and Spraul, 2003; Sturm and Seger, 2012).

The development of LC-SPE-NMR resulted from the observation that an SPE cartridge could be used, postcolumn, as an analyte enrichment device (Griffiths and Horton, 1998). To trap a given analyte separated by RP HPLC, an SPE cartridge installed postcolumn and preconditioned with a solvent of low eluotropic strength, such as H_2O or D_2O, can be efficiently used. A sensitive detector used on-line (UV or MS) triggers the collection of a given analyte automatically on individual SPE cartridges. The cartridges can be used in a 96-well plate format for the collection of multiple peaks. Once the analyte flows through the SPE cartridge, it will be retained, provided enough water was added postcolumn for trapping. To transfer the analytes to the NMR flow cell, the SPE cartridges are first dried with nitrogen and then eluted with a suitable deuterated solvent that has sufficient elution power (Figures 2d and 4b).

Such a process has different advantages. With this approach, the elution volumes of the LC peaks do
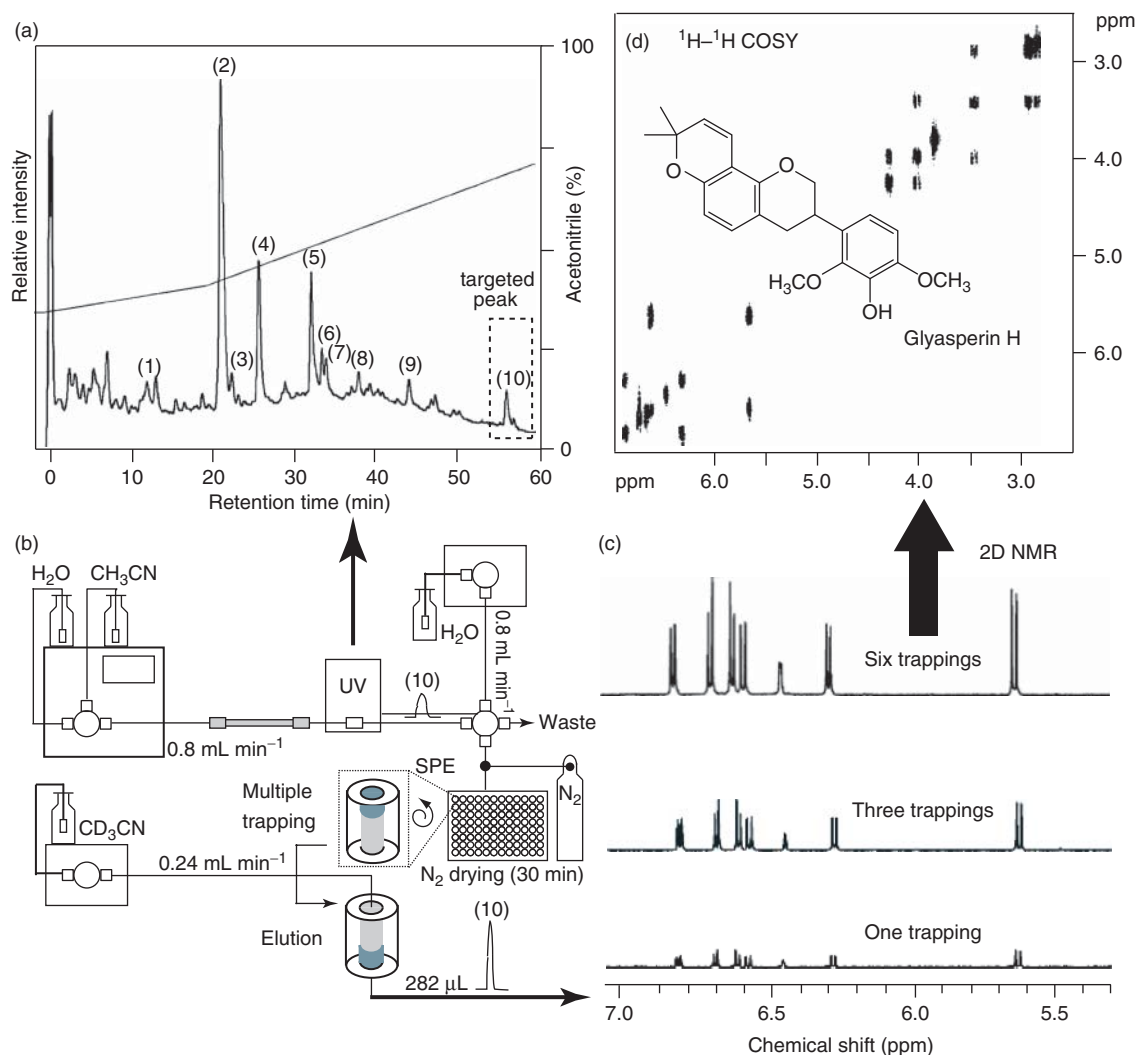
Figure 4 An example of HPLC-SPE-NMR. Workflow of HPLC-SPE-NMR exemplified by the analysis of isoflavonoids from *Smirnowia iranica*. (a) HPLC chromatogram of the ethanolic extract of *S. iranica* roots on a C_{18} column. The acetonitrile gradient profile in water is shown and the chromatogram shows average absorbance at 254 and 300 nm. (b) Principle of operation of the instrumentation used for HPLC-SPE-NMR. The separation is run at a flow rate of 0.8 mL min^{-1} and water is added after UV detection for optimal trapping onto the SPE cartridge (C_{18} HD). The trapping of the individual peaks is triggered by UV detection. The cartridges are dried using a nitrogen gas flow for 30 min and eluted with 282 μL of deuterated acetonitrile for analysis in a 30-μL NMR flow probe. (c) One-dimensional spectra obtained in HPLC-SPE-NMR mode with peak 10 (glyasperin H) after one, three, and six trappings on the same SPE cartridge. (d) COSY spectrum obtained in HPLC-SPE-NMR mode after seven trappings with peak 10 (glyasperin H, total acquisition time 6 h 10 min). (*Source*: Reprinted with permission from Lambert M, Staerk D, Hansen SH, Sairafianpour M, Jaroszewski JW. 2005b. Rapid extract dereplication using HPLC-SPE-NMR: Analysis of isoflavonoids from *Smirnowia iranica*. J Nat Prod 68: 1500–1509. Copyright 2005 American Chemical Society.)

not pose a problem compared to LC-NMR because the trapped peaks elute from the SPE cartridge in a fix volume of deuterated solvent. The elution volume of such SPE cartridges (10 × 2 mm i.d.) is 30–60 μL, which is close to the volume of

an LC-NMR flow probe. This elution volume can even be reduced for mass-limited samples if smaller SPE cartridges are used (10 × 1 mm i.d.) and if detection is achieved using lower volume microflow NMR probes (5 μL flow cell) (Lambert *et al.*,

2007). Owing to adequate matching between the elution and detection volumes, the NMR acquisition is performed on the entire amount of analyte in a given LC peak. Because the elution is performed in fully deuterated solvent, no solvent suppression is required (residual solvent signal resulting from an incomplete drying might require suppression). Contrary to LC-NMR, in LC-SPE-NMR, the entire mixture separation can be performed in standard HPLC grade solvents, and the transfer to the NMR flow cell requires only small volumes of deuterated solvents (ca. 300 µL). Instead of using LC-NMR flow probes, LC-SPE-NMR approaches can also be used with tubes and small-volume top-loading probes (Figure 1a) using a robot to fill racks of tubes and introducing them into the instrument with an automated sample changer (Wolfender *et al.*, 2013).

Typical LC-SPE-NMR analysis mixtures can be separated in optimal LC conditions without overloading, allowing for maximization of the HPLC resolution. This yields the trapping of only a few micrograms of the compound on the SPE cartridges; however, one of the advantages of SPE-NMR is that multiple trapping is possible. In this case, the HPLC separation of the same mixture is repeated several times, and a given analyte is accumulated using multiple subsequent trappings on a given SPE cartridge. The analyte preconcentration provided by SPE-NMR as well as the possibility of multiple trappings substantially increases the NMR sensitivity, especially for minor compounds, and allows for multiple 1D and 2D NMR experiments to be acquired (Wolfender *et al.*, 2013).

An example of the effect of multiple trappings on sensitivity is illustrated by the analysis of isoflavonoids from *Smirnowia iranica* Sabeti (Fabaceae) shown in Figure 4c. The ^{1}H NMR spectra of glyasperin H (peak 10), a minor constituent in the HPLC-UV trace of the extracts, is shown in Figure 4a and shows a significant enhancement in the S/N ratio after six trappings. As a result of this multiple trapping technique, 2D NMR spectra could be successfully recorded as shown by the COSY in Figure 4d (Lambert *et al.*, 2005b).

The main advantage of LC-SPE-NMR resides in the trapping of the analytes. To achieve good LC focusing by this method, the efficiency of SPE trapping has to be optimized so that the analyte elutes as a sharp band in the flow probe (Figure 4b). This can be regarded as a limitation of LC-SPE-NMR

because the trapping can be, to some extent, compound dependent. Thus, the physicochemical properties of the analytes must be accounted for when selecting the best SPE trapping material as well as the nature and amount of the eluent. Furthermore, the transfer of the analyte is also dependent on the nature of the deuterated solvent used (Clarkson *et al.*, 2007). For NP analysis, most of the trapping is performed on cartridges containing a divinylbenzene polymer or C_{18} phase with 1–2 mL of H_2O as the LC-makeup flow. The large majority of neutral compounds can be efficiently trapped by this method; however, charged polar compounds, such as alkaloids or organic acids, might not be retained; therefore, other methods or materials, such as ion exchange or porous carbon, are required (Sturm and Seger, 2012). In multiple trapping experiments, the cartridges normally show a linear improvement in the S/N ratio up to circa 100 µg. However, depending on the nature of the analytes and the trapping material, the saturation of the SPE cartridge might also occur leading to poor S/N ratio (Lambert *et al.*, 2005a).

The entire LC-SPE-NMR operation can be automated using integrated state-of-the-art systems that incorporate MS and UV detections, which are extremely useful, especially if multiple trapping is needed (Wilson *et al.*, 2006).

2.1.3 Multiple Hyphenated Systems Integrating LC-NMR and Automation

NMR represents an ideal detector for *de novo* on-line or at-line structure determination. The complete structure assignment of a given analytes requires gathering spectroscopic information using several different techniques, such as UV and MS spectra, and to some extent, infrared (IR) and circular dichroism (CD) spectra can also be recorded on-line during the LC-NMR experiment (Iwasa *et al.*, 2010; Wilson and Brinkman, 2007). As a result of the introduction of shielded magnets, multiple hyphenation is efficient because the detectors can be physically placed close to the center of the NMR magnet. The multiple hyphenation of various spectroscopic detectors in addition to NMR is possible for online experiments (Wilson and Brinkman, 2003); however, it requires many compromises because of specific needs for detectors. The most efficient combination consists of the LC-NMR-MS platform (Corcoran and Spraul,

2003), where the MS is used to trigger trapping of the LC peak of interest for subsequent NMR measurements and to generate complementary information for metabolite identification. The main issues with such platforms are sample overloading into the MS systems and a resulting shift of the molecular ion species because of proton–deuterium exchange reactions. However, this can be solved using an efficient postcolumn splitter, which diverts a small portion of the flow into the mass spectrometer and enables proton–deuterium back exchange by dilution with an appropriate makeup flow (Wolfender *et al.*, 2013). Efficient LC-SPE-NMR setups have also evolved toward advanced HPLC-PDA-MS-SPE-NMR platforms that can integrate CD measurements when on-line characterization of the chirality of the compounds is required (Sprogoe *et al.*, 2008).

Today, the combination of LC-MS in LC-SPE-NMR is completely integrated into the setup. The LC peaks are analyzed using MS before NMR detection, and the monitoring of the ions in MS during HPLC separation provides a precise MS trigger of the chromatographic peaks of interest (Schlotterbeck and Ceccarelli, 2009). The MS detection can be used to deconvolute coeluting components, thus facilitating the interpretation of the obtained NMR data.

2.2 Applications of LC-NMR in Phytochemical Analysis

2.2.1 LC-NMR Applications

In phytochemical analysis, LC-NMR methods have been predominantly used for dereplication studies or targeted *de novo* identification of given bioactive metabolites in crude extract profiling studies. In Table 1, a summary of the most recent applications (2010–2013) is presented, including the type of experiments and the probe used. For earlier applications, the reader can refer to other reviews (Queiroz, Wolfender, and Hostettmann, 2009; Sturm and Seger, 2012; Wolfender, 2010).

Among the different studies published recently, the on-line identification of the anti-inflammatory constituents of *Angelica dahurica* (Fisch. Ex hoffm.) Benth. Et Hook (Apiaceae) using LC-NMR is representative of the type of data that can be

obtained (Kang *et al.*, 2010). An HPLC-based activity profiling approach was used to investigate the anti-inflammatory properties of the extract using a nitric oxide (NO) inhibitory activity bioassay. In the first step and to localize the compound responsible for the biological activity, the ethanol extract was fractionated by HPLC on an analytical C_{30} RP column (250 × 4.6 mm, i.d.; 3 μm). Five fractions corresponding to the main peaks were collected and submitted to the NO inhibition assay and two LC peaks were found to be responsible for the activity. Then, an additional analysis was performed using the same conditions and replacing H_2O with D_2O for the LC-NMR-MS analysis. On the basis of these data, the active compounds were identified as known furanocoumarins – imperatorin and phellopterin.

Another recently developed method for the rapid identification of NPs having antimicrobial activities is through the direct coupling of a bacterial bioassay with chromatography (Kreiss *et al.*, 2010). This approach was based on a whole-cell bioluminescent reporter gene assay that was coupled with thin layer chromatography (TLC, see **Thin-layer Chromatography, with Chemical and Biological Detection Methods**) for primary hit detection and LC-MS and LC-NMR for dereplication and structure elucidation of the active compounds. Using this strategy, a new gyrA promoter named inthomycin A was identified from a myxobacterial extract without an isolation step. Dereplication of the active constituents was performed directly in the crude myxobacterial extract by LC-HRMS and LC-NMR experiments. LC-NMR coupling experiments were run in the stop-flow mode utilizing similar chromatographic conditions as applied for the HPLC/bioluminescence measurement. The bioactive component was identified by stop-flow ^{1}H–^{1}H TOCSY and COSY experiments. The total time of acquisition was 14.5 h and was performed with 22 μg μL^{-1} of the crude extract. The MS and NMR data obtained were then compared to those in the Chapman & Hall/CRC Chemical Database, which enabled the dereplication of inthomycin A as the active compound.

2.2.2 On-Line Microflow NMR Applications

While microflow NMR (e.g., CapNMR) is mainly used at-line, it can also be used in direct LC-NMR hyphenation in the on-flow mode. In this case, the

Table 1 Recent applications (2010–2013) of LC-NMR and LC-SPE-NMR for on-line metabolite identification.

LC-NMR applications

NMR setup	Sample	Compound class	NMR experiments	Solvent	Sample amounts	References
500-MHz, cryo LC-NMR probe (60 µL flow cell)	*Angelica dahurica* (Fisch. Ex hoffm.) Benth. Et Hook.	Furanocoumarins	^{1}H NMR using on-flow, isocratic conditions with a cryoprobe, LC-NMR-MS, C$_{30}$ column (250 × 4.6 mm, 3 µm), 0.5 mL min^{-1}	D$_2$O/CH$_3$CN	2.5 mg of ethanolic extract (50 mg mL^{-1}, 50 µL injection)	(Kang *et al.*, 2010)
500-MHz, cryo LC-NMR probe (60 µL flow cell)	*Sophora flavescens* Ait	Flavonoids	^{1}H NMR using on-flow LC-NMR with a cryoprobe, presaturation with WET, 0.5 mL min^{-1}	D$_2$O/CH$_3$CN	1 mg of ethanolic extract (50 mg mL^{-1}, 20 µL injection)	(Kim *et al.*, 2010)
500-MHz, LC-NMR probe	*Archangium* cf. *gephyra*	Inthomycins	^{1}H NMR using stop-flow LC-NMR with a cryoprobe, no solvent suppression	D$_2$O/CD$_3$CN (+0.05% CD$_3$OOD)	110 µg of extract (5 µL injection)	(Kreiss *et al.*, 2010)
600-MHz, LC-NMR probe (120 µL flow cell)	Four *Vitis* spp.	Anthocyanins	^{1}H NMR using stop-flow LC-NMR, 0.5 mL min^{-1}, one-dimensional version of the NOESY pulse sequence with presaturation during the relaxation delay, and mixing time at two frequencies	D$_2$O/CH$_3$CN (+0.25% TFA)	Not specified	(Acevedo De la Cruz *et al.*, 2012)
600-MHz, cryo LC-NMR probe (150 µL flow cell)	*Cimicifuga heracleifolia*	Caffeic acid derivatives	^{1}H NMR using stop-flow and on-flow, presaturation with WET	D$_2$O/CH$_3$CN (+0.1% FA)	5 mg of buthanolic extract (100 mg mL^{-1}, 50 µL injection)	(Yim *et al.*, 2012)
500-MHz, cryo-LC-NMR probe (60 µL flow cell)	*Petasites japonicus* (Siebold & Zucc.) Maxim.	Flavonoids and quinic acids derivatives	^{1}H NMR using stop-flow LC-NMR, presaturation with WET	D$_2$O/CH$_3$CN	2 mg of methanolic extract (100 mg mL^{-1}, 20 µL injection)	(Kim *et al.*, 2012)

LC-SPE-NMR applications

NMR setup	Sample	Compound class	NMR experiment	Elution volume and solvent	Sample amounts	Reference
600-MHz, LC-NMR probe (30 µL flow cell)	*Artemisia absinthium* L. and commercial preparation of A. *absinthium*	Sesquiterpene lactone, lignans, and flavonoid	^{1}H and ^{13}C NMR, COSY, HSQC, and HMBC	CD$_3$CN (153 µL)	3 × 0.3 mg of absinthin and degradation products (13 mg mL^{-1}, 20 µL injection)	(Aberham *et al.*, 2010)

Instrument	Species	Compound class	NMR experiments	Solvent	Quantity	Reference
400-MHz, LC-NMR probe (30 µL flow cell)	*Crinum asiaticum* L. var. *sinicum*	Alkaloids	^{1}H and ^{13}C NMR, COSY, NOESY, HSQC, and HMBC	CD$_3$OD (265 µL)	3×0.16 mg enriched fraction (33 mg mL^{-1}, 5 µL injection)	(Chen et al., 2011)
600-MHz, cryo 3-mm LC-NMR probe (60 µL flow probe)	*Boronia megastigma* (Nees)	Norisoprenoids, monoterpenoids, and cucurbates	^{1}H and micro-NMR, COSY, TOCSY, NOESY, HSQC, and HMBC	CD$_3$CN	Not specified	(Cooper et al., 2011)
600-MHz, cryo 5-mm NMR probe (cryofit 30 µL active volume)	*Myrica gale* L.	C-methylated flavanones and dihydrochalcones	^{1}H NMR, COSY, HSQC, and HMBC	CD$_3$CN	6×1.2 mg of extract (200 mg mL^{-1}, 6 µL injection)	(Fang, Paetz, and Schneider, 2011)
500-MHz, cryo 5-mm NMR probe	*Anthriscus sylvestris* L. (Hoffm.)	Lignans	^{1}H NMR, NOE, COSY, NOESY, HSQC, and HMBC	DMSO-d_6	Not specified (5 µL injection)	(Hendrawati et al., 2011)
400-MHz, LC-NMR probe (30 µL flow cell)	*Machilus philippinensis* Merr.	Glycosylated flavonoids	^{1}H and ^{13}C NMR, COSY, NOESY, HSQC, and HMBC	CD$_3$CN	3×1 mg of enriched fraction (50 mg mL^{-1}, 20 µL injection)	(Lin, Tsai, and Lee, 2011)
600-MHz, LC-NMR probe (30 µL flow cell)	*Eriobotrya japonica* (Thunb.) Lindl.	Flavonoids	^{1}H and ^{13}C NMR, COSY, HSQC, and HMBC	CD$_3$CN	Quantity not specified (20 µL injection) and 10 trappings	(Pfisterer et al., 2011)
600-MHz, cryo 5-mm NMR probe (cryofit 30 µL active volume)	*Solanum lycopersicum* L. (tomato mutant)	Flavonoids	^{1}H NMR	CD$_3$OD (227 µL)	Quantity not specified (20 µL injection)	(van der Hooft et al., 2011)
400-MHz, LC-NMR probe (30 µL flow cell); 600-MHz, no specification on NMR probe	*Litsea acuminata* (Blume) Kurata, L. *hypophaea* Hyata, *Neolitsea acuminatissima* (Hyata) Kaneh & Sasaki, and N. *konishii* (Hayata) Kaneh & Sasaki	Flavonoids	^{1}H NMR (SPE-NMR), ^{1}H COSY, HSQC, HMBC, and DEPT-135 (600 MHz)	CD$_3$CN (SPE-NMR); CD$_3$OD	1 mg of enriched fraction (100 mg mL^{-1}, 10 µL injection)	(Tsai and Lee, 2011)

(Continued overleaf)

Table 1 (*Continued*)

Instrument	Species	Compound class	NMR experiments	Solvent	Quantity	Reference
400-MHz, LC-NMR probe (120 μL flow cell) 600-MHz, cryo LC-NMR probe (60 μL flow cell)	*Phyllanthus myrtifolius* Moon and *Syagrus romanzoffiana* (Cham.) Glassman	Lignans and stilbenoids	^{1}H NMR, COSY, and NOESY	CD_3CN and CD_3OD	0.4 mg enriched fraction (40 mg mL^{-1}, 20 μL injection) and 0.4 mg enriched fraction (20.8 mg mL^{-1}, 20 μL injection)	(Wang et al., 2011)
600-MHz, LC-NMR probe (30 μL flow cell)	*Thymus vulgaris* L.	Monoterpines, flavonoids, and phenyl propanoids	^{1}H NMR, COSY, HSQC, and HMBC	CD_3CN	0.1 mg of defatted methanol–water extract (10 mg mL^{-1}, 10 μL injection) and 20 trappings	(Pieri et al., 2012)
800-MHz, cryo 5-mm NMR probe with cryofit 60 μL active volume and 2.5-mm tube	*Origanum vulgare* L.	Polyphenolic compounds	^{1}H- and ^{13}C NMR, COSY, TOCSY, NOESY, HSQC, and HMBC	CD_3CN (200 μL)	0.4 mg of extract (20 mg mL^{-1}, 20 μL injection) and 10 trappings	(Liu et al., 2012)
500-MHz, 3-mm LC-NMR probe (60 μL flow cell)	*Teucrium polium* L.	Flavonoids	^{1}H NMR, COSY, TOCSY, HSQC, and HMBC	CD_3CN	Quantity not specified (30 μL injection)	(Goulas et al., 2012)
400-MHz, LC-NMR probe (30 μL flow cell)	*Alnus formosana* Burk	Diarylheptanoids	^{1}H- and ^{13}C NMR, COSY, NOESY, HSQC, and HMBC	CD_3OD and pyridine-d_5	2 mg and 0.673 mg of enriched fraction (20 μL injection) and three trappings	(Lai et al., 2012)
500-MHz, cryo 5-mm NMR probe (cryofit 30 μL active volume)	*Wachendorfia thyrsiflora* L.	Phenylphenalenones	^{1}H NMR, COSY, HSQC, and HMBC	CD_3CN	2.2 mg injected (110 mg mL^{-1} in DMSO, 20 μL injection) and 4 mg injected (200 mg mL^{-1} in DMSO, 20 μL injection) and six trappings	(Fang, Kai, and Schneider, 2012)
600-MHz, cryo 5-mm NMR probe (cryofit 30 μL active volume)	*Camellia sinensis* (Black, Green, and White Tea Extracts)	Flavonoids	^{1}H NMR and COSY	CD_3OD (227 μL)	Quantity not specified (40 μL injection)	(van der Hooft et al., 2012)

entire separation can be performed using fully deuterated solvents because of the small amount of deuterated solvent required. Such an approach was used, for example, during the separation of essential bioactive isomers of the carotenoid bixin. Here, the mixture profiling was performed on a Pronto-SIL C_{30} (3 μm, 200 Å, 150 mm × 250 μm) in the isocratic mode (acetone-d_6/D_2O = 92 : 8 v/v). The sample was injected in a small volume of 500 nL. This method enables the recording of high quality stop-flow microflow ^1H NMR spectra and offers the possibility to acquire 2D ^1H–^1H COSY-NMR spectra with limited amounts of mixture. Finally, using a combination of LC-MS and LC-microflow NMR, it was possible to elucidate the structures of two stereoisomers of bixin (Rehbein *et al.*, 2007).

2.2.3 LC-SPE-NMR Applications

Because of the limitations of LC-NMR, especially in terms of sensitivity, over the past few years, HPLC-SPE-NMR has been used with growing frequency for the rapid dereplication of complex mixtures in NP research (Seger and Sturm, 2007).

This technique has been used to investigate the phytochemical composition of aerial parts and roots of *Wachendorfia thyrsiflora* L. (Haemodoraceae). The separation of the extracts from aerial and roots plant parts was performed on a Purospher RP18e column (5 μm; 250 × 4.6 mm) with a flow rate of 1 mL min^{-1} and a gradient of MeCN and H_2O. Furthermore, only 2 mg was injected onto the column, and the column was monitored by UV. Six cumulative trappings were performed for each LC peak that was selected for the 1D and 2D NMR analyses, and MeCN-d_3 was used to transfer the metabolite of interest. Identification was performed on 18 compounds including 11 phenylphenalenones in both the root and aerial extracts of this plant (Fang, Kai, and Schneider, 2012).

For *de novo* structure identification of NPs, detection and assignment of ^{13}C resonances are often essential for unequivocal elucidation. However, the ^{13}C acquisition is a challenging task because the gyromagnetic ratio of ^{13}C is only one-fourth that of ^1H and because the natural abundance of ^{13}C is only 1.1% of the total carbon nuclei. In addition, the overall receptivity of ^{13}C is approximately 5700 times lower than that of ^1H at the same magnetic field strength (Wolfender *et al.*, 2013). ^{13}C-NMR data can

be acquired indirectly by ^1H-detected one-bond [heteronuclear single-quantum coherence spectroscopy (HSQC) and heteronuclear multiple-quantum correlation spectroscopy (HMQC)] and multiple-bond (heteronuclear multiple-bond correlation spectroscopy, HMBC) 2D ^1H-^{13}C heteronuclear chemical shift correlation experiments. However, in some cases, carbon atoms and even hydrogen bearing carbon atoms can be difficult to detect because of the dependence on coupling constants, coupling path geometry, and substituent patterns.

Direct ^{13}C measurement was demonstrated to be measurable on sample limited amounts using a combination of a miniaturized cryo-NMR probe (30 μL) with an integrated cooled carbon channel preamplifier and several peak trappings on the HPLC-SPE-NMR system. Such an approach was applied for the identification of triterpenoids from a *Ganoderma lucidum* (Curtis ex Fr.) P. Karst (Ganodermataceae) extract (Wubshet *et al.*, 2012). First, the SPE trapping was optimized using three triterpenoid standards. A defatted methanol extract was fractionated using vacuum liquid chromatography (VLC). Using HRMS and ^1H NMR, a pooled triterpenoid fraction was selected and further fractionated using semipreparative RP HPLC. Then, one of the fractions that contained few chromatographic peaks was submitted to SPE-NMR. To achieve the expected S/N ratios in the ^{13}C NMR experiment, six trappings of 80 μL injections were necessary, which corresponded to approximately 560 μg. Using this procedure, an acquisition time of 13 h resulted in spectra with adequate S/N ratios for the detection of all ^{13}C signals. Structural elucidation was performed based on combinations of 1D and 2D NMR experiments [double quantum filtered correlation spectroscopy (DQF-COSY), HSQC, and HMBC], allowing the identification of two minor ganoderic acids.

Although SPE-NMR is currently applied to a wide variety of organic compounds, certain classes of compounds have limited application. This is the case for the positively charged alkaloids that are often poorly trapped on the commonly used SPE at low pH. To overcome this problem, two new approaches for efficient SPE trapping of such compounds have been developed (Johansen *et al.*, 2012). The authors used a divinylbenzene polymer cartridge (GP resin) and a postcolumn dilution of a 0.1 *M* NaOH solution to enhance the trapping of alkaloids on the commonly used SPE. However, this procedure was incompatible with phenolic alkaloids. The authors also evaluated

the trapping and elution efficacy on an SPE cartridge with a mixed-mode cation exchanger (sulfonyl groups on a polydivinylbenzene backbone). First, this method was evaluated using 24 alkaloid standards that covered a wide range of pK_a values, log P values, aliphatic and aromatic skeletons, and a variety of functionalities, such as alcohol groups and carboxylic that could influence the trapping efficiency. Finally, this method was applied to the analysis of the plant extracts containing alkaloids such as *Huperzia selago* (L.) Bernh. ex Schrank & Mart. (Lycopodiaceae) and *Triclisia patens* Oliv. (Menispermaceae). The methodology proposed was efficient for the trapping and elution of complex aliphatic alkaloids and phenolic alkaloids as evidenced by the acquisition of 2D NMR data for all trapped compounds. In contrast, GP resin proved only viable for the *H. selago* alkaloids, and the trapping and elution of bisbenzylisoquinoline alkaloids was dubious.

3　MICROFLOW NMR AND MICRO-NMR FOR SENSITIVE AT-LINE DETECTION

The use of directly coupled LC-NMR applications has decreased substantially in recent 10 years because of the major drawbacks of this technique, such as the need for solvent suppression or chemical shifts because of solvent composition changes when used in the gradient mode LC. In addition, owing to the intrinsically low sensitivity of the NMR detection, only major mixture constituents are detectable. This has resulted in an increased use of LC-SPE-NMR (Sturm and Seger, 2012), which has overcome most of the issues encountered in on-flow LC-NMR but at the price of extensive automation.

With the emergence of NMR systems with improved mass and/or concentration sensitivity, alternative methods to LC-SPE-NMR based on at-line NMR detection of LC peaks that can be collected by an independent and efficient microfractionation procedure have increased (Jansma *et al.*, 2005; Wolfender, Queiroz, and Hostettmann, 2005).

3.1　Design of Micro-NMR Probes with Enhanced Sensitivity

Three strategies have been employed to enhance NMR sensitivity: the use of a different coil type compared to conventional NMR systems in microflow NMR, the reduction of noise during the acquisition through cryogenic cooling of the probe in micro-cryoprobes, and the development of new coil material in high temperature superconducting (HTS) probes. The influence of the coil diameter and the noise of the receiver on the NMR S/N ratio are illustrated in Figure 1d. With these advances, the amount of sample needed for 1D and 2D NMR acquisitions decreased to the nanomolar range (Molinski, 2010), enabling NMR data acquisition on small molecules (molecular mass <1000 Da) in the microgram range (Dalisay and Molinski, 2009; Gronquist *et al.*, 2005; Williams *et al.*, 2012b) or below for the direct ^{13}C-NMR detection of LC peaks (Wubshet *et al.*, 2012).

3.1.1　At-Line Microflow NMR

As mentioned, the sensitivity of an NMR experiment increases in inverse proportion to the diameter of the coil (Hoult and Richards, 1976) (Figure 1d). Thus, flow probes with small coils (microflow probes) have been placed on the market for better performance (Olson *et al.*, 2004). These solenoid probes are naturally suited for LC-NMR because samples need to be introduced using flow. In this respect, the CapNMR, a commercially available solenoid probe, represents an interesting option for mass-limited samples.

With such probes, the collection of LC peaks is not directly hyphenated to HPLC as is the case with LC-SPE-NMR but the LC peaks are microfractionated from independent HPLC profiling analyses (see the following text) and the microfractions of interest are dried and analyzed postchromatographically in the microflow probe (Bohni *et al.*, 2013). To achieve the best sensitivity, the dried analytes (typically a few micrograms) are dissolved in a volume of deuterated solvent that slightly exceeds the volume of the microflow cell (5 μL). The very concentrated samples are then parked into the center of the probe by injection and further pushed with a calibrated volume of deuterated solvent (push volume). The loading can be performed using manual injection with a syringe of a given analyte, or microfractions can be injected automatically from a 96-well plate (Jansma *et al.*, 2005). These microtiter plates can be used for HPLC collection. The LC microfractions are dried and redissolved in a minimal amount of deuterated solvent.

An LC autosampler system connected to a microliter pump providing the corrected deuterated solvent for pushing each individual sample into the microflow probe is then used to measure the NMR spectra of all of the microfractions automatically. Compared to a conventional NMR setup with 5-mm tubes, a sensitivity gain of up to fivefold can be obtained, and 2D NMR measurement becomes addressable with only a few tenths of micrograms of analyte (Gronquist *et al.*, 2005; Wolfender, Queiroz, and Hostettmann, 2005). An example of the spectra that can be obtained from a mixture of NPs analyzed overnight from consecutive automated injection from a 96-well plate is shown in Figure 3c and d. Figure 3c consists of a pseudo-LC 2D plot that was reconstituted by merging all of the CapNMR spectra recorded for the consecutive microfractions obtained from the LC separation. As shown, when compared to an on-flow LC-NMR 2D plot (Figure 3a), this mode of acquisition does not require any solvent suppression because the spectra are measured in a deuterated solvent. Furthermore, because all of the microfractions are analyzed in the same solvent, such as deuterated methanol, there is no shift as a result of the change of composition because of the gradient elution mode (Figure 3c). The spectrum of a furanocoumarin acquired under these conditions is shown in Figure 3d. Compared with LC-NMR (Figure 3b), a good quality spectrum (Figure 3d) was obtained directly from the analysis of the corresponding microfraction without the requirement of solvent suppression.

As for LC-SPE-NMR, it has been shown that this type of microflow probe can also be integrated in microscale LC-MS-NMR platforms via a thoroughly optimized automated droplet microfluidic NMR loading method (microdroplet NMR). With such an approach, by postchromatographic analysis of the analytes of interest, interpretable 1D NMR spectra were obtained from analytes down to the 200 ng level in 1 h per well using an automated NMR data acquisition. Such platforms were efficient in the dereplication of known cyanobacterial compounds and the prioritization of the isolated unknown using a single 30 μg injection of a crude extract (Lin *et al.*, 2008).

Excellent line shape can be obtained with small solenoid probes and because of the minute amount of solvent used, the residual solvent line in the spectra is much less important than in conventional probes, which also helps to maximize gain values for enhanced sensitivity (Figure 3d). Another advantage of solenoid probes is that the coil can be simultaneously tuned to multiple frequencies, allowing for near-optimal performance on different channels and thus the analysis of multiple samples at the same time (Li *et al.*, 2003). These are now available commercially from Protasis (Norcross *et al.*, 2010).

3.1.2   1-mm and Other Microtubes

In line with the development of microflow probes, microprobes with optimized sensitivity using disposable 1-mm microtubes (Figure 1a) with a sample volume of 5 μL, similar to that of the CapNMR probe, have also been brought to the market (Schuetz *et al.*, 2011). In this case, the automation is less easily achieved than with tubeless approaches, and the samples must be transferred from the dried collected fraction to the microtubes (Figure 2e). One advantage of this approach compared to microflow NMR is that the microtubes do not interact with the probe itself, and problems related to the poor dissolution of samples, such as precipitation and clogging that might arise in microflow approaches, are avoided. Microtubes, such as Shigemi microtubes, can also be used to enhance the sensitivity of mass-limited samples in probes with larger diameters; however, the filling factor will not be optimized to the same extent as for dedicated microprobes, and the gains in sensitivity will be decreased (Marcourt *et al.*, 2011).

3.1.3   Additional Sensitivity with Cryo-NMR Probes

An additional advantage of using microtubes is that the probe can be cryogenically cooled, and these small volume microprobes (e.g., a 1.7-mm cryoprobe) are commercially available. The development of cryo-NMR probes has brought NMR sensitivity to a new level. As the noise of the measurement is reduced by cooling the coil and the associated electronics with helium (V_{noise} in Figure 1d), the S/N ratio is improved by as much as four times compared to room-temperature probes (Styles *et al.*, 1984). Different probe configurations exist from all major NMR vendors, from 5-mm tube probes to 1.7-mm microtube probes (MicroCryoProbe™). The cryogenized option is also available for LC-NMR flow probes

(CryoFlowProbe). Such probes have boosted the sensitivity of NMR measurements of samples with limited amounts (Figure 1c) as well as for compounds with relatively low solubility, such as proteins that require a relatively large volume to be solubilized (Molinski, 2010).

With such a setup, high sensitivities can be attained and 2D NMR measurements of LC peaks obtained from a single conventional analytical HPLC run are feasible (Williams *et al*., 2012a).

An example of the advantages in terms of sensitivity that can be observed using microtubes with a cryoprobe compared to a room-temperature microflow probe is shown by the analysis of the same amount of an indole alkaloid on both probes of the same 600-MHz NMR system (Figure 5). As shown, the ^{1}H spectra could be recorded with only 2 µg of the compound; however, a much better S/N ratio was obtained with the MicroCryoProbe using only 64 scans (Figure 5a), whereas 1024 scans were required for the CapNMR probe (Figure 5b). With 10 µg of the same compound in the MicroCryoProbe, an HSQC spectrum could be recorded in 1 h (Figure 5c).

To the best of our knowledge, probes with extremely small active volumes, such as the one holding 1-mm microtubes or the microflow solenoid probes, are not available as cryogenated probes most likely because of the difficulty in adapting the technology to such small volumes. The main drawback of this technology is the price of the cryogenated probes and that the maintenance costs associated with these probes are significantly higher than that of room-temperature probes.

Another way to increase sensitivity is to improve the quality of the coil material. For most NMR probes, the rf coils are constructed from copper-based conductors (Wolfender *et al*., 2013). The use of HTS materials such as yttrium barium copper oxide (YBCO), to build the rf coil in cryoprobes results in higher sensitivity, and together with the other advantages (reduced noise) of cryoprobes, gives HTS coils the best S/N ratio among commercially available NMR probes (Brey *et al*., 2006; Wolfender *et al*., 2013). Owing to geometric restrictions in the construction of HTS coils, only probes with small-diameter coils and tube probes are currently commercially available. As for the other probes, a higher sensitivity is achieved with HTS coils of small diameters, and 1-mm probes using the materials described have been built (Brey *et al*., 2006).

3.2 Isolation and Purification of NPs for Sensitive NMR Analysis

At-line microisolation procedures before sensitive NMR could be performed either by multiple collection of a given peak of interest using standard HPLC conditions (4 mm i.d. columns) or by semipreparative separations with columns of relatively small diameters (e.g., 10 mm i.d.). These latter types of HPLC columns are usually preferred because they allow relatively high loadings (few milligrams up to tens of milligrams of crude mixture), they can be operated at flow rates compatible with collection in 96-deep well plates (2–3 mL min^{-1}) and the microfraction volume is relatively restricted (1–2 mL) (Bohni *et al*., 2013). Such volumes enable the collection directly in 96-deep well plates that can be evaporated in a single step. According to the complexity of the extracts or fractions studied, this microfractionation procedure typically yields microfractions that contain between a few micrograms to approximately hundreds of micrograms for the most abundant metabolites (Bohni *et al*., 2013; Glauser *et al*., 2008). At-line microfractionation can also be performed on an LC-SPE-NMR platform, and in this case, the peak trapped on SPE can be eluted into microtubes of various volumes (Schmidt *et al*., 2012; Wubshet *et al*., 2012).

Efficient procedures exist for the scaling up of the high resolution metabolite profiling of extracts obtained at the analytical level using LC-MS, HPLC, or UHPLC (ultra high performance liquid chromatography) and the further optimization of separating the analyte of interest at the semipreparative scale (Eugster, Glauser, and Wolfender, 2013). For the optimization of the chromatographic resolution of a given analyte in a complex mixture, optimized elution conditions can be calculated using chromatographic software based on the comparison of the LC peak that is to be isolated in two linear gradient elutions of different slopes (Figure 6a). On the basis of this, a simulated chromatogram can be generated to evaluate the efficiency of the separation obtained (Figure 6b). To keep the same selectivity at the analytical and semipreparative scale, the analytical conditions have to be transferred to the semipreparative conditions using geometric transfer on the correctly characterized LC instruments (Figure 6b) (Guillarme *et al*., 2008). Such rules have been reported and applied successfully for the isolation of minor crude extract constituents for
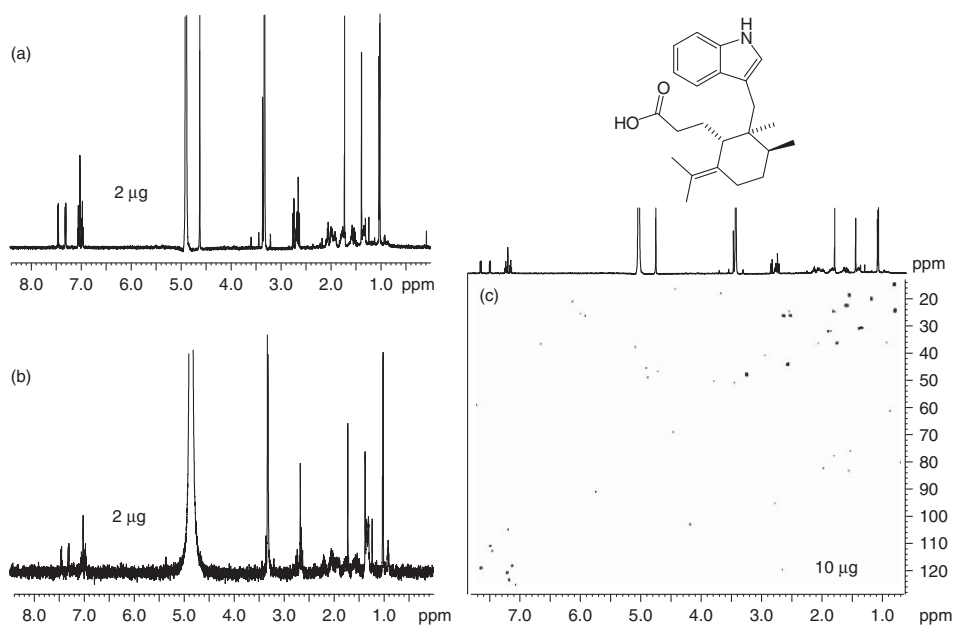
Figure 5 Comparison of ^{1}H NMR spectra of 2 µg of natural product suaveolindole acquired on a microflow and a microcryo NMR probe. The increased NMR sensitivity of the microflow NMR system compared to a conventional NMR system (5-mm tube probe, saddle-type coil) is because of its improved coil design (solenoid type), whereas the improved sensitivity of the microcryo system is both a result of the cryogenically cooled coil system and the smaller dimensions of the coil (saddle type). The NMR mass and concentration sensitivity of the microcryo NMR system is clearly superior to the sensitivity of the microflow NMR system. (a) For NMR analysis on a 1.7-mm microcryo system (MicroCryoProbe), 2 µg of the natural product suaveolindole was dissolved in 13 µL of deuterated methanol, placed in a 1-mm-diameter tube, and then, the ^{1}H NMR spectrum was acquired using 64 scans (approximate analysis time: 5 min). (b) For the analysis on the microflow NMR system (CapNMR™), 2 µg of the natural product suaveolindole was concentrated in 6 µL of deuterated methanol, injected into the probe with the aid of a syringe, and the ^{1}H NMR spectrum was acquired using 1024 scans (approximate analysis time: 1 h 20 min). (c) An HSQC spectrum was acquired on a microcryo system using a 10-µg sample with an acquisition time of 1 h. (*Source*: Spectra were kindly provided by Mark O'Neil-Johnson, Sequoia Sciences Inc. Reproduced by permission.)
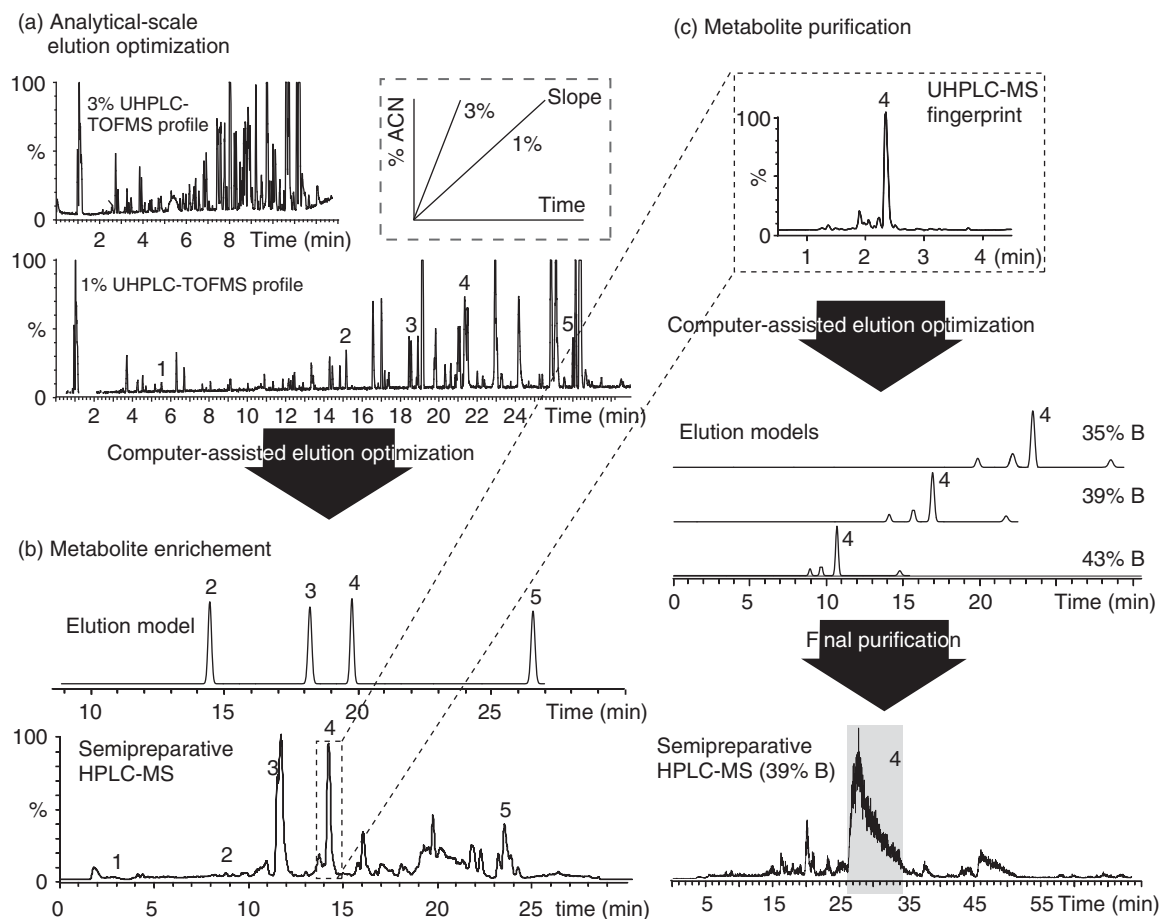
Figure 6   Two-step software-driven purification of natural products using microfractionation. (a) UHPLC-TOFMS analysis of crude extract performed with the two gradient slopes necessary for optimizing the separation of metabolites (**2–5**) through software modeling. (b) Targeted metabolite enrichment with simulated separation of these four metabolites and corresponding semipreparative LC-MS separation. (c) Final targeted purification of metabolite **4** with fingerprint of the enriched fraction containing **4** that was used for modeling, elution models of three different isocratic separation conditions for the final purification step, and corresponding final semipreparative LC-MS purification using the isocratic condition at 39% B. (*Source*: Reprinted with permission from Bertrand et al., 2013a. *De novo* production of metabolites by fungal co-culture of *Trichophyton rubrum* and *Bionectria ochroleuca*. J Nat Prod 76: 1157-1165. Copyright 2013 American Chemical Society.)

further microflow NMR characterizations (Bertrand *et al*., 2013b; Glauser *et al*., 2008).

As an example, the UHPLC-TOFMS profiling of the crude extract of a fungal co-culture extract is shown in Figure 6a. The analysis was performed with the two gradient slopes necessary for optimizing the separation of the targeted metabolites through software modeling. On the basis of this information, a simulated separation could be computed (Figure 6b), which was then used for the first enrichment step by semipreparative separation. A second simulation (Figure 6c) aided in finding the isocratic conditions

for a final isolation step, which yielded a few tens of micrograms of the desired biomarker for further microflow NMR analysis (Bertrand *et al*., 2013b).

At-line microisolation procedures, before sensitive NMR measurement at the microgram scale, require careful optimization because when working with such small amounts of analytes, signals from solvent impurities or column bleeding may considerably deteriorate the quality of the NMR spectra. Indeed, most HPLC-grade solvents are not NMR-grade and protonated impurities might interfere. Moreover, according to the type of stationary phase used, small

amounts of the packing material might bleed and generate signals visible in the ^{1}H NMR spectrum that are further strengthened by the concentration of the fractions. When dealing with microgram amounts of samples, the drying procedures are also important, as residual water (Figure 3d) will deteriorate the spectral quality and thus introduce unwanted solvent lines that might need to be suppressed. Therefore, efficient eluent evaporation procedures, such as vacuum centrifugation, must be used, and hygroscopic deuterated solvents need to be stored under inert gas to avoid the presence of residual water. Contamination can also come from the vials, the microtiter plate used for collecting microfractions or the SPE cartridge used for SPE-NMR. The quality and the background NMR signals generated by these devices must be tested according to the conditions and solvents used for each of the measurements, and control NMR spectra using only solvent blanks are recommended to compare NMR spectra and correctly discriminate impurity signals from those of the analytes (Wolfender *et al.*, 2013).

3.3 Applications of At-Line Micro-NMR in Phytochemical Analysis

At-line NMR applications that are considered herein are those where the NMR flow cells are not used in direct hyphenation with HPLC. This includes at-line microflow measurements after microfractionation and all experiments using microtubes as well as those involving SPE-NMR with transfer to microtubes, which is a method that is frequently used in this field. All of these applications are summarized in Table 2.

3.3.1 At-Line Microflow NMR Applications

A significant number of recent applications (Table 2) were performed using microflow NMR (CapNMR). For example, the identification of the antimicrobial compounds from the rhizome of *Peucedanum ostruthium* W.D.J. Koch (Apiaceae) was performed using this method, following a classical bioguided isolation procedure based on assays with three pathogenic bacteria (*Bacillus cereus*, *Escherichia coli*, and *Staphylococcus aureus*) (Gökay *et al.*, 2010). The activity of the crude ethyl acetate extract

and fractions obtained using different chromatographic methods were investigated by the agar disk diffusion method. All fractions were analyzed using an LC-UV-electrospray ionization-MS (LC-UV-ESI-MS) and a microflow NMR probe on a 600-MHz instrument. The constituents of the fraction with the highest antibacterial activity were rapidly identified as the coumarin derivatives, oxypeucedanin and oxypeucedanin hydrate. The ^{1}H NMR analysis of oxypeucedanin was obtained using a 15-µg (∼52 nmol) sample, whereas the analysis of oxypeucedanin hydrate was achieved using a 22-µg (∼72 nmol) sample. The small amounts of these two compounds were sufficient to obtain unambiguous ^{1}H NMR spectra in approximately 2 min and 20 s.

In the frame of a high throughput phytochemical investigation based on the initial normal-phase flash chromatography of crude extracts by a purification step including a semipreparative RP HPLC, the constituents of the apolar root extract of *Greenwayodendron suaveolens* (Engl. & Diels) Verdc. (Annonaceae) were efficiently studied. This strategy afforded milligram amounts of the major compounds (39.8–194 mg) and micrograms of the minor constituents (710–980 µg) including three new and two known indole sesquiterpene alkaloids that were characterized by HRMS and 1D and 2D microflow NMR. One of the isolated compounds, named pentacyclindole, was determined to possess a new NP framework and showed interesting antibiotic activity against clinical isolates of *S. aureus* (Williams *et al.*, 2010).

3.3.2 At-Line Micro-NMR Applications

An alternative to microflow NMR requires the use of 1-mm microtubes having approximately the same sample volume. Recently, this technique was used for the structure elucidation of five antiplasmodial bisabololoxide sesquiterpene diesters from the aerial parts of *Artemisia persica* Boiss. (Asteraceae) (Moradi-Afrapoli *et al.*, 2013). An HPLC-time-based activity profiling using 350 µg of extract enabled an efficient correlation of biological activity to a specific chromatogram region. LC-MS^n analysis together with at-line micro-NMR analysis revealed the presence of several compounds in the active fraction. Targeted isolation of the active compounds was performed by a combination of normal-phase medium pressure liquid chromatography (MPLC)

Table 2 Recent applications (2010–2013) of microflow NMR (CapNMR™) and micro-NMR (3.0–1.0-mm microtubes) for at-line metabolite identification.

NMR setup	Sample	Compound class	NMR experiments	Solvent	Sample amounts	References
At-line microflow NMR (CapNMR)						
600-MHz, microflow NMR (5 µL flow cell)	*Radix imperatoriae* (*Peucedanum ostruthium* W.D.J. Koch)	Oxypeucedanin and corresponding hydrate (psoralene and coumarin)	^{1}H NMR	$CDCl_3$	Injection of 52 and 72 nmol (150 and 220 µg in 15 µL)	(Gökay *et al.*, 2010)
600-MHz, microflow NMR (5 µL flow cell)	*Oncidium* spp.	Stilbenoids	^{1}H NMR, COSY, NOESY, HSQC, and HMBC	CD_3OD and $CDCl_3$	50–480 µg in 6.5 µL	(Williams *et al.*, 2012a)
500-MHz, microflow NMR (5 µL flow cell)	*Eschscholzia californica*	Alkaloids	^{1}H NMR	$DMSO\text{-}d_6$	200 ng to 8 µg	(Gathungu *et al.*, 2012)
At-line microNMR (1.7- and 1-mm microtubes) after isolation or LC-SPE-NMR trapping						
600-MHz, cryo 1.7-mm NMR probe	*Oncidium* spp.	Stilbenoids	^{1}H NMR, COSY, NOESY, HSQC, and HMBC	CD_3OD and $CDCl_3$	5–33 µg	(Williams *et al.*, 2012a)
600-MHz, cryo 1.7-mm NMR probe	*Ganoderma lucidum* (Curtis ex Fr.) P. Karst	Triterpenoids	^{13}C NMR (13 h acquisition time and detection limit 0.5 µg)	CD_3OD	600 µg of enriched fraction (7 mg mL^{-1} and 80 µL injection), HPLC-SPE-NMR and transfer to NMR tube, and six trappings	(Wubshet *et al.*, 2012)
600-MHz, cryo 1.7-mm NMR probe and 5-mm NMR probe	*Carthamus oxyacantha* M. Bieb.	Glycosylated spiranes, lignin glycoside, and vanilic acid	^{1}H NMR, COSY, NOESY, HSQC, and HMBC	CD_3CN (30 µL) and CD_3OD	500 µg (20 mg mL^{-1}, 25 µL injection), HPLC-SPE-NMR and transfer to NMR tube, and eight trappings	(Johansen *et al.*, 2011)

Method / Instrument	Species / Source	Compound class	NMR experiments	Solvent	Amount / details	Reference
500-MHz, LC-NMR probe (60 μL flow cell); 600-MHz, cryo 1.7-mm NMR probe	24 alkaloid standards and extracts of *Huperzia selago* (L.) Bernh. ex Schrank & Mart. (Lycopodiaceae) and *Triclisia patens* Oliv.	Alkaloids	^{1}H NMR, COSY, NOESY, HSQC, and HMBC	CD_3OD, $CDCl_3$, and CH_3OD + 5% NH_4OH	250 and 200 μg of alkaloid residue (5 and 20 mg mL^{-1}, 50 and 10 μL injection); HPLC-SPE-NMR and transfer to NMR tube, and 4, 6, or 10 trappings	(Johansen et al., 2012)
600-MHz, cryo 5-mm NMR probe with 1.7-mm tube	*Cinnamomum subavenium* Miq.	Dibenzocyclo-heptanoids	^{1}H- and ^{13}C NMR, COSY, NOESY, HSQC, and HMBC	CD_3OD	1 mg of enriched fraction (50 mg mL^{-1}, 20 μL injection), HPLC-SPE-NMR and transfer to NMR tube, and three trappings	(Lin and Lee, 2012)
500-MHz, LC-NMR probe (60 μL flow cell); 600-MHz, cryo 1.7-mm NMR probe	*Carthamus oxyacantha* M. Bieb. and *Penicillium namyslowski* K.M. Zalessky	Glycosylated spiranes, lignin glycosides, and fatty acids	^{1}H NMR, COSY, NOESY, HMBC, and HSQC	CD_3CN (30 μL)	0.6 mg (24 mg mL^{-1}, 25 μL injection) and 0.05 mg (5 mg mL^{-1}, 10 μL injection) and HPLC-SPE-NMR and transfer to NMR tube	(Johansen, Wubshet, and Nyberg, 2013)
500-MHz, 1-mm and 5-mm NMR probes	*Artemisia persica* Boiss.	Bisabololoxide sesquiterpene	^{1}H NMR, COSY, NOESY, HSQC, and HMBC (1-mm probe); ^{13}C NMR (5-mm probe)	C_6D_6	Microgram and milligram scales	(Moradi-Afrapoli et al., 2013)
500-MHz, 1-mm and 5-mm NMR probes	*Abrus precatorius* ssp. *africanus*	Isoflavan quinones	^{1}H NMR, ^{13}C NMR, COSY, NOESY, HSQC, and HMBC	$CDCl_3$ and CD_3OD	Microgram and milligram scales	(Hata et al., 2013)

Table 2 *(Continued)*

At-line microtube NMR (2.5- and 3-mm tubes) after LC-SPE-NMR trapping

NMR setup	Sample (plant)	Compound class	NMR experiments	Solvent	Sample amounts	References
400-MHz, 3-mm and 5-mm NMR probes with 3-mm tube	*Nauclea pobeguinii* (Pobég.) E.M.A. Petit	Alkaloids	^{1}H and ^{13}C NMR, HSQC, and HMBC	CD_3CN	1 mg of enriched fraction (200 mg mL^{-1}, 5 µL injection) and HPLC-SPE-NMR and transfer to NMR tube	(Xu *et al.*, 2012a)
400-MHz, 3-mm and 5-mm NMR probe with 3-mm tube	*Ormocarpum kirkii* S. Moore	Isoflavonoids	^{1}H and ^{13}C NMR, HSQC, and HMBC	CD_3OD	1 mg of enriched fraction (50 mg mL^{-1}, 20 µl injection) and HPLC-SPE-NMR and transfer to NMR tube	(Xu *et al.*, 2012b)
500-MHz, cryo-NMR probe with 2.5-mm tubes	*Strychnos usambarensis* Gilg	Alkaloids	^{1}H and ^{13}C NMR, COSY, HSQC, and HMBC	CD_3CN	Quantity not specified, HPLC-SPE-NMR and transfer to NMR tube, and three trappings	(Cao *et al.*, 2012)
800-MHz, 5-mm cryo-NMR probe with 2.5-mm tube	Apple peel (*Malus* x *domestica* Borkh.)	Flavonoids	^{1}H NMR	CD_3OD	200 µg (10 mg mL^{-1}, 20 µL injection) and HPLC-SPE-NMR and transfer to NMR tube	(Schmidt *et al.*, 2012)

and semipreparative RP HPLC. Structure elucidation was achieved by 1D and 2D NMR experiments using a 1-mm NMR probe. Relative configurations of the isolated compounds were established on the basis of three J ^{1}H–^{1}H coupling constants and nuclear Overhauser enhancement (NOE) difference spectra. Finally, the relative and absolute configurations of the active compounds were determined by comparing the experimental electronic circular dichroism (ECD) spectroscopy with the simulated ECD data to find possible stereoisomers using time-dependent density function theory (TDDFT). Some of the isolated compounds exhibited *in vitro* antimalarial activity against *Plasmodium falciparum* in the low micromolar range.

The same strategy has been used for the identification of the active compounds in *Abrus precatorius* L. ssp. *africanus* (Fabaceae) against the protozoan parasite *Trypanosoma brucei rhodesiense*. An HPLC microfractionation of the extract in a 96-well plate, followed by biological activity, was used to localize the active peaks that were analyzed by 1D and 2D NMR in 1-mm microtubes. Using this approach, two isoflavan hydroquinones and three isoflavan quinones were characterized *de novo* (Hata *et al.*, 2013).

3.3.3 MicroCryo Probe Applications

The use of MicroCryoProbes represents further improvement in NMR sensitivity. Such an approach was found to be key in the investigation of Orchidaceae metabolites that were isolated at the microgram scale. In this respect, the search for new anticancer agents in several species of the Orchidaceae family led to the identification of 15 stilbenoids, including a new phenanthraquinone and two new dihydrostilbenes from *Oncidium microchilum* Bateman ex Lindl., *Oncidium isthmi* Schltr., and *Myrmecophila humboldtii* Rolfe (Orchidaceae). The compounds were isolated using the combination of an initial normal-phase flash chromatography followed by a semipreparative HPLC separation with an RP-C_{18} stationary phase. Despite the small quantity obtained (between 5 and 480 µg per compound), high quality 2D NMR spectra were measured, which allowed structural elucidation of all of the compounds. Several of the compounds were found to inhibit the proliferation of NCI-H460 and M14 cancer cell lines (Williams *et al.*, 2012a).

The combination of SPE-NMR (to trap target compounds) and 1.7-mm microtubes (to enhance the NMR analysis sensitivity) has been used with success by several groups (Table 2). This approach was used for the phytochemical investigation of the *Carthamus oxyacantha* M. Bieb. (Asteraceae) extract and was also compared to results from a traditional approach. The classical approach was time consuming and was used for a nontargeted fractionation of the extract, which led to the isolation of only a few compounds. However, HPLC-PDA-HRMS-SPE-NMR was used for efficient targeted isolation and identification of the selected constituents. This hyphenated approach, involving VLC and semipreparative HPLC, was found to be considerably faster and required less solvent and other consumables than the classical approach. The SPE-NMR system employed 1.7-mm tubes, and the NMR spectra were acquired using a cryogenically cooled probe. The elution of the SPE cartridges into the microtubes was performed with a robotic liquid handler, and the tubes were managed using an automated sample changer. For the preliminary analysis, 30 peaks were selected for adsorption onto SPE cartridges to obtain ^{1}H NMR spectra. On the basis of the UV or MS data, two or three trappings were performed for each peak. Looking at their ^{1}H NMR spectra, most of the peaks were found to correspond to saturated fatty acids and flavonoids. On the basis of these initial ^{1}H NMR data, 11 peaks, presenting unusual NMR signals, were selected for SPE-NMR cumulative trapping for additional studies using 2D NMR. A total of 15 compounds, of which four were new spiro compounds, were finally characterized (Johansen *et al.*, 2011).

Recently, a similar approach involving LC-SPE-*tt*NMR (*tt*NMR, tube transfer nuclear magnetic resonance) was used to quickly screen α-glucosidase inhibitors in complex matrices (Schmidt *et al.*, 2012). Alpha-glucosidase inhibitors are oral antidiabetic drugs that prevent carbohydrate digestion and are used for the treatment of diabetes mellitus type 2 (van de Laar *et al.*, 2005). This approach of using LC-SPE-*tt*NMR was divided into three main steps: first, the active extract was fractionated using a chromatographic scout separation in a 96-well plate for activity evaluation; second, the glucosidase inhibition of the analytes was assessed in the microtiter plates, which permitted the establishment of an HPLC biochromatogram for the localization of the active compounds; and finally, the LC-SPE-NMR analysis was targeted on the active peaks to elucidate

their structures. This methodology was applied to the analysis of the "Pink Lady" apple peel extract. The 180 fractions that were collected during the scout separation were assessed for α-glucosidase inhibitory activity after evaporation of the HPLC solvent. Analytes with α-glucosidase inhibitory activity were correlated with their corresponding peaks in the HPLC chromatogram. The information obtained was then used for the setup of the UV-threshold-based trapping of the active peaks for multiple trapping onto the SPE cartridges. Trapped compounds were automatically dried for 45 min using a nitrogen gas flow and subsequently eluted into 2.5-mm NMR tubes with 140 μL of methanol-d_4. This strategy allowed the identification of epicatechin, reynoutrin, and avicularin as the active compounds.

3.3.4 Quantification Using At-Line LC-NMR Applications

Direct quantification by NMR (qNMR) after on-line or at-line LC-NMR has been attempted for a limited number of applications. For on-line LC-qNMR, quantification can be achieved using an internal standard that is either added to the sample before chromatographic separation for the isocratic mode or added to the LC solvent in gradient mode to compensate for the changing solvent composition (Godejohann and Preiss, 1998). For at-line LC-NMR approaches, isolated compounds are sometimes quantified using an internal standard (Pauli, Jaki, and Lankin, 2007), by comparison with the ^{13}C satellite signals (which is adapted particularly for use with cryoprobes) (Dalisay *et al.*, 2009) or with the residual solvent signal (Pierens *et al.*, 2008). An innovative approach was used in the field of environmental toxicology that involved using NMR to quantify cyanide in water, whereby the ions were derivatized with a fluorinated reagent; then, the reaction mixture was purified by LC-SPE-NMR; and finally, after trapping and using an internal standard that was added to the analyte with the NMR solvent, the compound was quantified by at-line ^{19}F and ^{1}H NMR (Mazumder, Kumar, and Dubey, 2013). Through the combination of different information, such as the retention time from the LC and the chemical shift of NMR spectroscopy, higher specificity for cyanide quantification was achieved compared to that previously observed with common analyses techniques such as gas chromatography.

3.4 Advantages of Microfractionation for Tracking Complementary Activity

As full structural characterization of NPs eluted in microgram amounts from an HPLC column has become feasible thanks to the improvement in NMR sensitivity, the direct analysis of the generated microfractions in a biological assay seems the logical next step in NP drug discovery. Several on-line biochemical assays have already been developed (de Jong *et al.*, 2006; Heus *et al.*, 2010), whereby detection is achieved using either UV fluorescence or MS spectrometry. Drawbacks of these methods (termed *high resolution screening*, *HRS*) are, on the one hand, the limited validity of biochemical assay compared to *in vivo* or high content assays and, on the other hand, the lack of determination of the potency of the detected activity as these detectors do not show a quantitative response to analytes in complex mixtures such as plant extracts. Microfractionation coupled to NMR offers the possibility to quantify the generated microfractions in one step with the analyte characterization and directly use the generated microgram amount samples for dose–response assessment.

This strategy was applied for the detection of anti-angiogenic and anti-inflammatory compounds from the Tanzanian plant *Rhynchosia viscosa* (Roth) DC. (Fabaceae) (Bohni *et al.*, 2013) and is shown in Figure 7. Here, the use of an internal standard for NMR quantification is not desired (so that the samples are not contaminated before biological testing) and an external standardization method [pulse length based concentration determination, PULCON (Wider and Dreier, 2006)] was adopted. At first, the enriched methanolic extract was microfractionated using time-triggered collection (Figure 7a and b). Then, all generated microfractions were screened in an *in vivo* zebrafish assay for their anti-angiogenic activity (Figure 7c). Bioactive microfractions were subsequently analyzed using ^{1}H NMR (Figure 7d) and their content was characterized (Figure 7e). The same analyses permitted the quantification of the isolated compounds (Figure 7f), and thus, the microfractions containing compounds in a pure form and at highest amount were taken to perform dilution series for concentration–response analysis directly on the microgram amount samples.
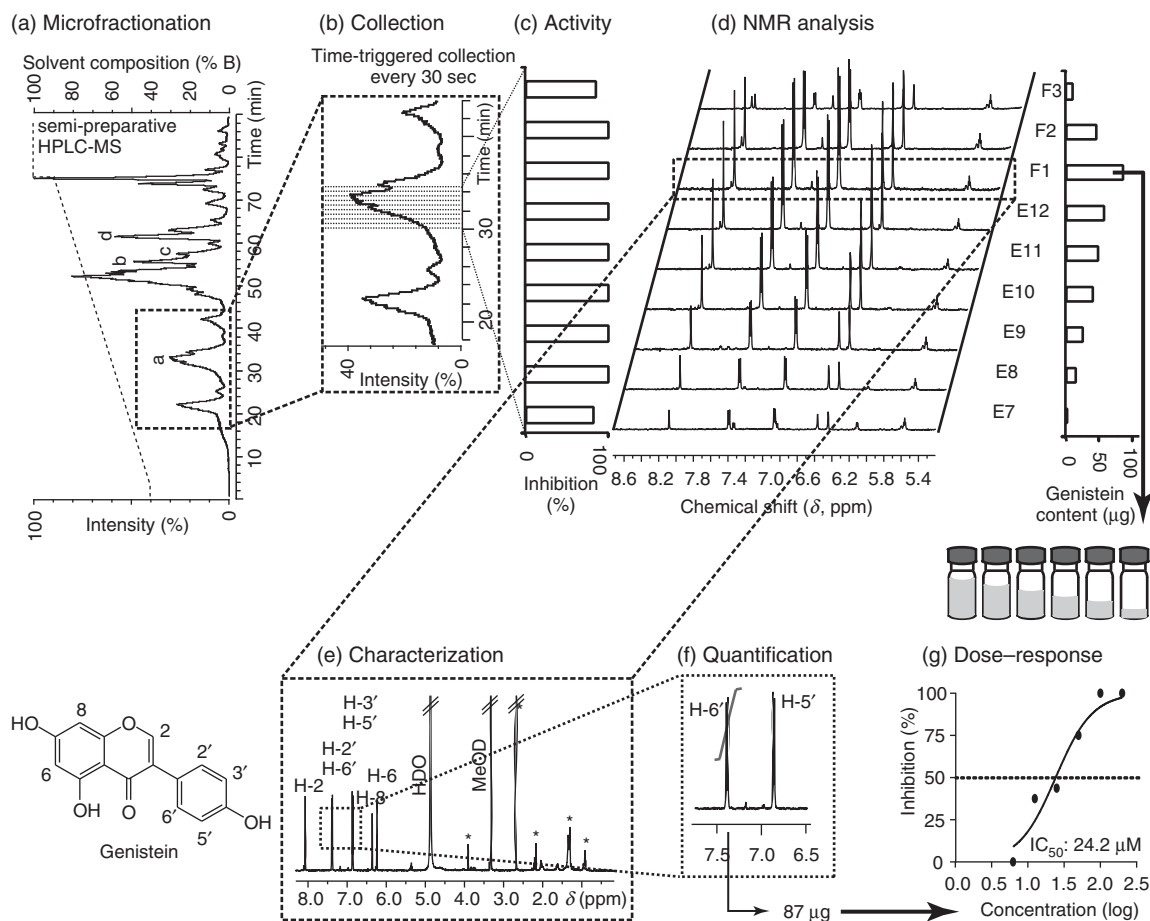
Figure 7 High resolution screening of plant extract with direct estimation of potency of isolated microgram amount samples using quantitative NMR and microfractionation. (a) Semipreparative HPLC chromatogram for the microfractionation of the enriched extract of *Rhynchosia viscosa* (Roth) DC. HPLC conditions: XBridge™ BEH C_{18} column (250 × 10 mm i.d., 5 μm); A: 0.1 vol% formic acid–H_2O, B: 0.1 vol% formic acid–MeOH, 40–90% in 74.9′; 2.3 mL min^{-1}; and ESI-MS detection in negative ionization mode. (b) Microfractions were collected every 30 s into 96-deep well plates. (c) All collected microfractions were tested for anti-angiogenic activity on a zebrafish model. Several microfractions eluting after 30 min were inhibiting angiogenesis to 70–100% and were analyzed by ^1H NMR using a microflow NMR probe [CapNMR™, (d)] and contained the isoflavone genistein (e). *: impurity. (f) The amount of genistein in the individual microfractions was quantified with an external standard using PULCON. (g) The microfraction with the highest amount of genistein (F1) was used to perform a dilution series for a concentration–response analysis of the anti-angiogenic activity. (*Source*: Adapted from Integration of Microfractionation, qNMR and Zebrafish Screening for the *in vivo* Bioassay-Guided Isolation and Quantitative Bioactivity Analysis of Natural Products as published in PLoS One, 8(5): e64006, 2013, doi:10.1371/journal.pone.0064006. © Bohni *et al.*, 2013. distributed under the Creative Commons Attribution License.)

4 CONCLUSION

There is an important need for the rapid and efficient identification of NPs in complex biological mixtures, especially for NP-based drug discovery programs and for many aspects related to food supplements or phytopharmaceuticals. This need has been increased considerably by the rapid advancement of metabolomics where peak annotation represents a major bottleneck in the approach and structures of a large number of biomarkers need to be correctly assigned. Despite the considerable advances in metabolite profiling methods based on advanced LC-MS or LC-MS/MS methods and as a result of the lack of unified MS databases for small molecule, the unambiguous identification of NPs must rely mainly on NMR.

Over the past three decades, the sample amount required for *de novo* structure identification has decreased from the range 20–50 mg to less than 1 mg for conventional 1D and 2D NMR measurements. With access to higher magnetic field strengths and the development of dedicated micro-NMR probes, sensitivity has greatly increased and ^{1}H NMR spectra can be recorded in the submicrogram level, whereas 2D NMR experiments can be obtained with only a few micrograms using state-of-the art NMR technology. These amounts can now be readily obtained from the single HPLC metabolite profiling analysis of a crude natural extract.

The hyphenation of NMR to HPLC has evolved over the years to achieve rapid metabolite identification in crude natural extracts. Direct hyphenation of on-flow LC-NMR has been demonstrated to be practical and feasible at the expense of some compromises. To overcome most of the limitations associated with on-flow LC-NMR and to provide access to key 2D NMR experiments, efficient methods of analyte enrichment via multiple collections of given LC peaks have been integrated and have evolved toward automated and efficient LC-SPE-NMR-MS platforms providing high quality spectra, on-line in a flow cell, or at-line after transfer in microtubes.

To further improve sensitivities for limited amounts of sample, room-temperature microflow probes and small-diameter cryoprobes have been developed. Such probes are no longer used on-line but are used at-line after HPLC separation. Their integration for efficient peak identification requires that metabolite profiling methods be scaled up, enabling the separation of milligram amounts of complex extracts with a good prediction of the accuracy. For efficient microfractionation of the extracts, LC-MS-triggered fractionation strategies can be used, which allow the precise collection of given LC peaks based on their extracted ion traces. This can be performed directly in 96-well plates, and subsequent NMR analysis of the wells can be automated to yield high quality 1D and 2D NMR spectra using only a few micrograms of NPs. However, the spectral quality is still strongly dependent on the quality of the chromatographic separation; therefore, the HPLC high resolution separations at the semipreparative scale need to be further advanced, as this has been the case at the analytical level with the introduction of columns packed with sub-2 μm particles used for UHPLC.

For applications involving unresolved LC peaks or the analysis of fractions still containing a mixture of a few NPs, multivariate data analysis (see **Multivariate data analysis**) methods for the NMR data sets and covariance analysis with orthogonal methods such as LC-MS may be useful to extract NMR information of pure constituents from simple mixtures (Robinette *et al.*, 2011).

The quality of metabolite profiling studies will continue to rely on NMR for unambiguous identification. With the development of HPLC biological profiling assays, NMR will also be used, with increasing frequency, for the absolute quantification of the amounts collected by microfractionation and for the assessment of the bioactive potency of given NPs.

The ^{1}H NMR spectra that are generated by a well-defined deuterated solvent using sensitive at-line approaches have the advantage of being universal and comparable between instruments. In this respect, the creation of tools and NMR databases of NPs that are easily searchable and implementable should improve the efficiency of peak identification in HPLC. The high throughput acquisition of NMR data obtained with these latest technologies and the access to NMR databases should very significantly boost NP research in the years to come.

ACKNOWLEDGMENTS

LIST OF ABBREVIATIONS

[13]C	Carbon-13
CD	Circular dichroism
COSY	Correlation spectroscopy
1D	One-dimensional
2D	Two-dimensional

DEPT	Distortionless enhancement by polarization transfer
DQF	Double quantum filtered
ECD	Electronic circular dichroism
ELSD	Evaporative light scattering detection
ESI	Electrospray ionization
HCS	High content screening
^{1}H NMR	Proton NMR
HMBC	Heteronuclear multiple-bond correlation spectroscopy
HMQC	Heteronuclear multiple-quantum correlation spectroscopy
HPLC	High performance liquid chromatography
HRMS	High resolution mass spectrometry
HRS	High resolution screening
HSQC	Heteronuclear single-quantum coherence spectroscopy
HTS	High temperature superconducting
IR	Infrared
MeOH	Methanol
MeCN	Acetonitrile
MPLC	Medium pressure liquid chromatography
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MS^n	Multiple stage mass spectrometry
MW	Molecular weight
NMR	Nuclear magnetic resonance
*tt*NMR	Tube transfer NMR
NO	Nitric oxide
NOESY	Nuclear Overhauser enhancement spectroscopy
NPs	Natural products
PDA	Photo diode array
PULCON	Pulse length based concentration determination
qNMR	Quantitative NMR
SPE	Solid-phase extraction
TDDFT	Time-dependent density function theory
TLC	Thin layer chromatography
TOCSY	Total correlation spectroscopy
TOFMS	Time-of-flight mass spectrometry
UHPLC	Ultra high performance liquid chromatography
VLC	Vacuum liquid chromatography
WET	Water suppression enhanced through T_1 effects
YBCO	Yttrium barium copper oxide

REFERENCES

Aberham, A., Cicek, S. S., Schneider, P., *et al.* (2010) *J. Agric. Food Chem.*, **58**, 10817–10823.

Acevedo De la Cruz, A., Hilbert, G., Rivière, C., *et al.* (2012) *Anal. Chim. Acta*, **732**, 145–152.

Albert, K. (1995) *J. Chromatogr. A*, **703**, 123–147.

Albert, K. (2002) *On-line LC-NMR and Related Techniques*, John Wiley & Sons, Chichester.

Azzollini, A., Bertrand, S., Nievergelt, A., *et al.* (2012) *Chimia*, **66**, 487.

Bayer, E., Albert, K., Nieder, M., *et al.* (1979) *J. Chromatogr. A*, **186**, 479–507.

Bertrand, S., Schumpp, O., Bohni, N., *et al.* (2013a) *J. Nat. Prod.*, **76**, 1157–1165.

Bertrand, S., Schumpp, O., Bohni, N., *et al.* (2013b) *J. Chromatogr. A*, **1292**, 219–228.

Bohni, N., Maldonado, A. M., Siverio-Mota, D., *et al.* (2013) *PLoS One*, **8**, e64006.

Breton, R. C. and Reynolds, W. F. (2013) *Nat. Prod. Rep.*, **30**, 501–524.

Brey, W. W., Edison, A. S., Nast, R. E., *et al.* (2006) *J. Magn. Reson.*, **179**, 290–293.

Bross-Walch, N., Kuhn, T., Moskau, D., *et al.* (2005) *Chem. Biodiv.*, **2**, 147–177.

Cao, M., Muganga, R., Nistor, I., *et al.* (2012) *Phytochem. Lett.*, **5**, 170–173.

Challal, S., Bohni, N., Buenafe, O. E., *et al.* (2012) *Chimia*, **66**, 229–232.

Chen, C. K., Lin, F. H., Tseng, L. H., *et al.* (2011) *J. Nat. Prod.*, **74**, 411–419.

Clarkson, C., Sibum, M., Mensen, R., *et al.* (2007) *J. Chromatogr. A*, **1165**, 1–9.

Cogne, A. L., Queiroz, E. F., Marston, A., *et al.* (2005) *Phytochem. Anal.*, **16**, 429–439.

Cooper, C. M., Davies, N. W., Motti, C. A., *et al.* (2011) *J .Agric. Food Chem.*, **59**, 2610–2617.

Corcoran, O. and Spraul, M. (2003) *Drug Discov. Today*, **8**, 624–631.

Creek, D. J., Jankevics, A., Breitling, R., *et al.* (2011) *Anal. Chem.*, **83**, 8703–8710.

Dalisay, D. S. and Molinski, T. F. (2009) *J. Nat. Prod.*, **72**, 739–744.

Dalisay, D. S., Morinaka, B. I., Skepper, C. K., *et al.* (2009) *J. Am. Chem. Soc.*, **131**, 7552–7553.

Eugster, P. J. and Wolfender, J.-L. (2012) UHPLC in natural products analysis, in *Basics of UHPLC and Application in Life Science Analysis*, eds. J.-L. Veuthey and D. Guillarme, RSC Publishing, Cambridge, pp. 354–386.

Eugster, P. J., Glauser, G. and Wolfender, J.-L. (2013) Strategies in biomarker discovery. Peak annotation by MS and *de novo* structure identification by micro-NMR, in *Metabolomics Tools for Natural Product Discoveries*, eds. U. Roessner and A. Dias, Humana Press, New York.

Fang, J. J., Paetz, C. and Schneider, B. (2011) *Biochem. Syst. Ecol.*, **39**, 68–70.

Fang, J. J., Kai, M. and Schneider, B. (2012) *Phytochemistry*, **81**, 144–152.

Funari, C. S., Eugster, P. J., Martel, S., *et al.* (2012) *J. Chromatogr. A*, **1259**, 167–178.

Gathungu, R. M., Oldham, J. T., Bird, S. S., *et al.* (2012) *Anal. Methods*, **4**, 1315–1325.

Glauser, G., Guillarme, D., Grata, E., *et al.* (2008) *J. Chromatogr. A*, **1180**, 90–98.

Godejohann, M. and Preiss, A. (1998) *Anal. Chem.*, **70**, 590–595.

Gökay, O. and Albert, K. (2012) *Anal. Bioanal. Chem.*, **402**, 647–669.

Gökay, O., Kuhner, D., Los, M., *et al.* (2010) *Anal. Bioanal. Chem.*, **398**, 2039–2047.

Goulas, V., Gomez-Caravaca, A. M., Exarchou, V., *et al.* (2012) *Sci. Technol.*, **46**, 104–109.

Griffiths, L. and Horton, R. (1998) *Magn. Reson. Chem.*, **36**, 104–109.

Gronquist, M., Meinwald, J., Eisner, T., *et al.* (2005) *J. Am. Chem. Soc.*, **127**, 10810–10811.

Guillarme, D., Nguyen, D. T. T., Rudaz, S., *et al.* (2008) *Eur. J. Pharm. Biopharm.*, **68**, 430–440.

Hata, Y., Raith, M., Ebrahimi, S. N., *et al.* (2013) *Planta Med.*, **79**, 492–498.

Hendrawati, O., Woerdenbag, H. J., Michiels, P. J. A., *et al.* (2011) *Phytochemistry*, **72**, 2172–2179.

Heus, F., Giera, M., Kloe, G., *et al.* (2010) *Anal. Bioanal. Chem.*, **398**, 3023–3032.

van der Hooft, J. J. J., Mihaleva, V., de Vos, R. C. H., *et al.* (2011) *Magn. Reson. Chem.*, **49**, S55–S60.

van der Hooft, J. J. J., Akermi, M., Unlu, F. Y., *et al.* (2012) *J. Agric. Food Chem.*, **60**, 8841–8850.

Hoult, D. I. and Richards, R. E. (1976) *J. Magn. Reson.*, **24**, 71–85.

Iwasa, K., Cui, W., Takahashi, T., *et al.* (2010) *J. Nat. Prod.*, **73**, 115–122.

Jansma, A., Chuan, T., Albrecht, R. W., *et al.* (2005) *Anal. Chem.*, **77**, 6509–6515.

Jaroszewski, J. W. (2005a) *Planta Med.*, **71**, 795–802.

Jaroszewski, J. W. (2005b) *Planta Med.*, **71**, 691–700.

Johansen, K. T., Wubshet, S. G., Nyberg, N. T., *et al.* (2011) *J. Nat. Prod.*, **74**, 2454–2461.

Johansen, K. T., Ebild, S. J., Christensen, S. B., *et al.* (2012) *J. Chromatogr. A*, **1270**, 171–177.

Johansen, K. T., Wubshet, S. G. and Nyberg, N. T. (2013) *Anal. Chem.*, **85**, 3183–3189.

de Jong, C. F., Derks, R. J. E., Bruyneel, B., *et al.* (2006) *J. Chromatogr. A*, **1112**, 303–310.

Kang, S. W., Kim, C. Y., Song, D. G., *et al.* (2010) *Phytochem. Anal.*, **21**, 322–327.

Kim, S. J., Kim, S. M., Kim, M. C., *et al.* (2010) *J. Med. Plants Res.*, **4**, 2452–2459.

Kim, S. M., Kang, S. W., Jeon, J. S., *et al.* (2012) *Biomed. Chromatogr.*, **26**, 199–207.

Kind, T. and Fiehn, O. (2007) *BMC Bioinf.*, **8**, 105.

Kreiss, W., Froede, R., Moehrle, V., *et al.* (2010) *Anal. Bioanal. Chem.*, **398**, 2081–2088.

van de Laar, F. A., Lucassen, P. L., Akkermans, R. P., *et al.* (2005) *Diabetes Care*, **28**, 154–163.

Lai, Y. C., Chen, C. K., Lin, W. W., *et al.* (2012) *Phytochemistry*, **73**, 84–94.

Lambert, M., Staerk, D., Hansen, S. H., *et al.* (2005a) *Magn. Reson. Chem.*, **43**, 771–775.

Lambert, M., Staerk, D., Hansen, S. H., *et al.* (2005b) *J. Nat. Prod.*, **68**, 1500–1509.

Lambert, M., Wolfender, J.-L., Staerk, D., *et al.* (2007) *Anal. Chem.*, **79**, 727–735.

Li, Y., Logan, T. M., Edison, A. S., *et al.* (2003) *J. Magn. Reson.*, **164**, 128–135.

Lin, H. C. and Lee, S. S. (2012) *J. Nat. Prod.*, **75**, 1735–1743.

Lin, Y. Q., Schiavo, S., Orjala, J., *et al.* (2008) *Anal. Chem.*, **80**, 8045–8054.

Lin, H. C., Tsai, S. F. and Lee, S. S. (2011) *J. Chinese Chem. Soc.*, **58**, 555–562.

Liu, H. B., Zheng, A. M., Liu, H. L., *et al.* (2012) *J. Agric. Food Chem.*, **60**, 129–135.

Marcourt, L., Ndjoko-Ioset, K., Pinto, A., *et al.* (2011) *Chimia*, **65**, 488.

Mazumder, A., Kumar, A. and Dubey, D. K. (2013) *J. Chromatogr. A*, **1284**, 88–99.

Molinski, T. F. (2010) *Nat. Prod. Rep.*, **27**, 321–329.

Moradi-Afrapoli, F., Ebrahimi, S. N., Smiesko, M., *et al.* (2013) *Phytochemistry*, **85**, 143–152.

Norcross, J. A., Milling, C. T., Olson, D. L., *et al.* (2010) *Anal. Chem.*, **82**, 7227–7236.

Olson, D. L., Norcross, J. A., O'Neil-Johnson, M., *et al.* (2004) *Anal. Chem.*, **76**, 2966–2974.

Pauli, G. F., Jaki, B. U. and Lankin, D. C. (2007) *J. Nat. Prod.*, **70**, 589–595.

Pfisterer, P. H., Shen, C. X., Nikolovska-Coleska, Z., *et al.* (2011) *Bioorg. Med. Chem.*, **19**, 1002–1009.

Pierens, G. K., Carroll, A. R., Davis, R. A., *et al.* (2008) *J. Nat. Prod.*, **71**, 810–813.

Pieri, V., Sturm, S., Seger, C., *et al.* (2012) *Metabolomics*, **8**, 335–346.

Queiroz, E. F., Wolfender, J.-L., Atindehou, K. K., *et al.* (2002) *J. Chromatogr. A*, **974**, 123–134.

Queiroz, E. F., Wolfender, J.-L. and Hostettmann, K. (2009) *Curr. Drug Targets*, **10**, 202–211.

Ramm, M., Wolfender, J.-L., Queiroz, E. F., *et al.* (2004) *J. Chromatogr. A*, **1034**, 139–148.

Rehbein, J., Dietrich, B., Grynbaum, M. D., *et al.* (2007) *J. Sep. Sci.*, **30**, 2382–2390.

Reynolds, W. F. and Enriquez, R. G. (2002) *J. Nat. Prod.*, **65**, 221–244.

Robinette, S. L., Ajredini, R., Rasheed, H., *et al.* (2011) *Anal. Chem.*, **83**, 1649–1657.

Schlotterbeck, G. and Ceccarelli, S. M. (2009) *Bioanalysis*, **1**, 549–559.

Schmidt, J. S., Lauridsen, M. B., Dragsted, L. O., *et al.* (2012) *Food Chem.*, **135**, 1692–1699.

Schuetz, C., Quitschau, M., Hamburger, M., *et al.* (2011) *Fitoterapia*, **82**, 1021–1026.

Seger, C. and Sturm, S. (2007) *LC GC Eur.*, **20**, 587–597.

Shockcor, J. P., Unger, S. E., Savina, P., *et al.* (2000) *J. Chromatogr. B Biomed. Sci. Appl.*, **748**, 269–279.

Smallcombe, S. H., Patt, S. L. and Keiffer, P. A. (1995) *J. Magn. Reson., Ser. A*, **117**, 295–303.

Sprogoe, K., Staerk, D., Ziegler, H. L., *et al.* (2008) *J. Nat. Prod.*, **71**, 516–519.

Sturm, S. and Seger, C. (2012) *J. Chromatogr. A*, **1259**, 50–61.

Sturm, S., Seger, C., Godejohann, M., *et al*. (2007) *J. Chromatogr. A*, **1163**, 138–144.

Styles, P., Soffe, N. F., Scott, C. A., *et al*. (1984) *J. Magn. Reson.*, **60**, 397–404.

Tsai, S. F. and Lee, S. S. (2011) *J. Chinese Chem. Soc.*, **58**, 376–383.

Wang, C. Y., Lam, S. H., Tseng, L. H., *et al*. (2011) *Phytochem. Anal.*, **22**, 352–360.

Waridel, P., Wolfender, J.-L., Lachavanne, J. B., *et al*. (2004) *Phytochemistry*, **65**, 945–954.

Wider, G. and Dreier, L. (2006) *J. Am. Chem. Soc.*, **128**, 2571–2576.

Williams, R. B., Hu, J. F., Olson, K. M., *et al*. (2010) *J. Nat. Prod.*, **73**, 1008–1011.

Williams, R. B., Martin, S. M., Hu, J.-F., *et al*. (2012a) *Planta Med.*, **78**, 160–165.

Williams, R. B., Martin, S. M., Hu, J.-F., *et al*. (2012b) *J. Nat. Prod.*, **75**, 1319–1325.

Wilson, I. D. and Brinkman, U. A. T. (2003) *J. Chromatogr. A*, **1000**, 325–356.

Wilson, I. D. and Brinkman, U. A. T. (2007) *TrAC, Trends Anal. Chem.*, **26**, 847–854.

Wilson, S. R., Malerod, H., Petersen, D., *et al*. (2006) *J. Sep. Sci.*, **29**, 582–589.

Wilson, S. R., Malerod, H., Petersen, D., *et al*. (2007) *J. Sep. Sci.*, **30**, 322–328.

Wolfender, J.-L. (2009) *Planta Med.*, **75**, 719–734.

Wolfender, J.-L. (2010) LC-NMR and related techniques for the rapid identification of plant metabolites, in *High Performance Liquid Chromatography in Phytochemical Analysis*, eds. M. Waksmundzka-Hajnos and J. Sherma, CRC Press, Boca Raton, pp. 287–329.

Wolfender, J.-L., Queiroz, E. F. and Hostettmann, K. (2005) *Magn. Reson. Chem.*, **43**, 697–709.

Wolfender, J.-L., Bohni, N., Ndjoko-Ioset, K., *et al*. (2013) Advanced spectroscopic detectors for identification and quantification: nuclear magnetic resonance, in *Liquid Chromatography: Fundamentals and Instrumentation*, eds. S. Fanali, P. Haddad, C. F. Poole, P. J. Schoenmakers and D. Lloyd, Elsevier, Oxford, pp. 349–384.

Wubshet, S. G., Johansen, K. T., Nyberg, N. T., *et al*. (2012) *J. Nat. Prod.*, **75**, 876–882.

Xiao, H. B., Krucker, M., Putzbach, K., *et al*. (2005) *J. Chromatogr. A*, **1067**, 135–143.

Xu, Y. J., Foubert, K., Dhooghe, L., *et al*. (2012a) *Phytochem. Lett.*, **5**, 316–319.

Xu, Y. J., Foubert, K., Dhooghe, L., *et al*. (2012b) *Phytochemistry*, **79**, 121–128.

Yim, S.-H., Kim, H., Park, S.-H., *et al*. (2012) *Arch. Pharm. Res.*, **35**, 1559–1565.

New Developments of Laser Desorption Ionization Mass Spectrometry in Plant Analysis

Andreas Schinkovitz, Denis Séraphin and Pascal Richomme

University of Angers, UFR des Sciences Pharmaceutiques, Angers, France

1 INTRODUCTION: HISTORICAL BACKGROUND

The development of matrix-assisted laser desorption ionization (MALDI) and related methods dates back to the late 1970s when it was shown that the irradiation of small organic molecules by a pulsed laser at high intensity leads to the formation of ions that could be analyzed through mass spectrometry (MS) (Posthumus *et al.*, 1978). Most generally using an UV laser as the irradiation source, this observation led to the conception of the so-called *laser desorption/ionization* (LDI) source. Later on, this soft ionization source was coupled with a time-of-flight (TOF) detector in LDI-TOF mass spectrometers. From that time on, it was possible to produce and characterize ions originating from nonvolatile and/or thermolabile compounds. However, it also became clear that not every entity could enter the LDI process thus limiting the technique to very few and particular compounds. An important step forward was taken during the 1980s by introducing a so-called "*matrix*," that is, a small organic molecule carrying a strong UV chromophore, to generalize the LDI process (Karas and Hillenkamp, 1988; Tanaka *et al.*, 1988). This concept was called *MALDI* (Figure 1) and extended the range of MS applications to bioanalyticals and synthetic polymers. Ever since MALDI

MS has gained immense popularity and by today represents one of the most commonly applied analytical techniques for the analysis of large molecules.

2 LASER DESORPTION IONIZATION MASS SPECTROMETRY

2.1 Laser Desorption Ionization (LDI)

Apart from MALDI, other LDI applications were contemporaneously developed over time. With substantial progress made in nanotechnologies, matrix-free LDI techniques, such as "surface-assisted laser desorption ionization" (SALDI) MS, were established. In this method, classical matrix molecules are replaced by nanostructured materials such as germanium, carbon, and particularly silicon (Chen and Vertes, 2006; Law, 2010; Law and Larkin, 2011). It has been shown that under optimized conditions, SALDI may facilitate the detection of selected compounds in the low femtomole range (Chen and Vertes, 2006). Further, nanostructured metal or transitional metal derivatives, like, silver, titanium nitride, and gold, could be used in a similar way (McLean, Stumpo, and Russell, 2005; Schuerenberg, Dreisewerd, and Hillenkamp, 1999).
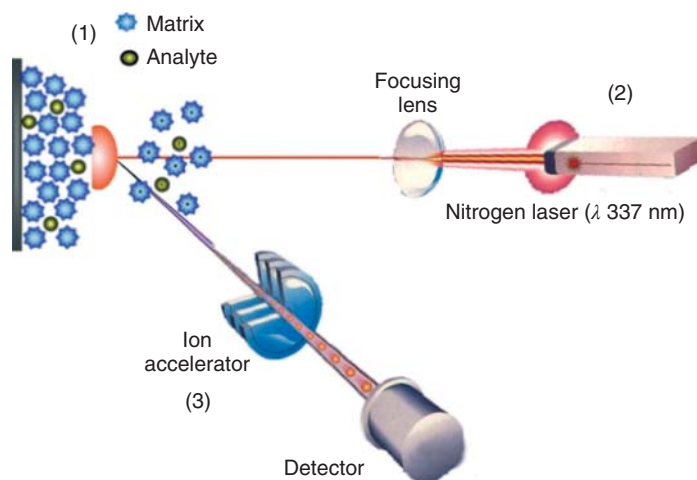
Figure 1   Basic principle of MALDI (TOF) MS. (1) A sample (analyte) is cocrystallized with an excess of matrix on a sample carrier (e.g., a stainless steel plate) (2). A laser pulse (a few nanosecond) ionizes the matrix molecules. Sample molecules are ionized through charge transfer ($[M+H]^+$, $[M+Na]^+$, $[M+K]^+$, etc.) (3) ions are accelerated by an electric field (4) small ions reach the detector earlier than large ones. A time-of-flight (TOF) analyzer measures the time required by the ions to reach the detector: the mass to charge ratio (m/z) of an ion is proportional to the square of its drift time.
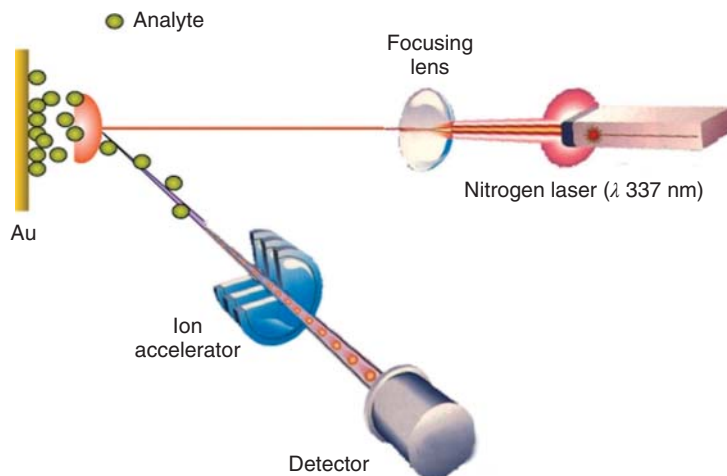
Figure 2   Laser ionization/desorption on a gold surface (GFLDI). The sample (analyte) is directly crystallized on the sample plate covered with gold.

Depending on their specific properties, these materials may be utilized as free nanoparticles, or as surface material attached on a sample carrier (Figure 2). A good example for the practical application of the latter is represented by the spectrum of 3′,4′,5′,5,7-pentamethoxyflavone displayed in Figure 3, which was obtained by gold film-assisted LDI (GFLDI). In that case, a thin gold layer (35 nm) was deposited onto the sample carrier by physical

vapor deposition (PVD), a method that facilitates a particularly uniform surface formation.

As shown in Figure 3, solely analyte signals were detected representing the ideal case of an LDI spectrum.

Nevertheless, like any other technique, both LDI and MALDI have their specific benefits and drawback, and choosing a suitable system might be difficult even for the advanced users. Therefore, some of
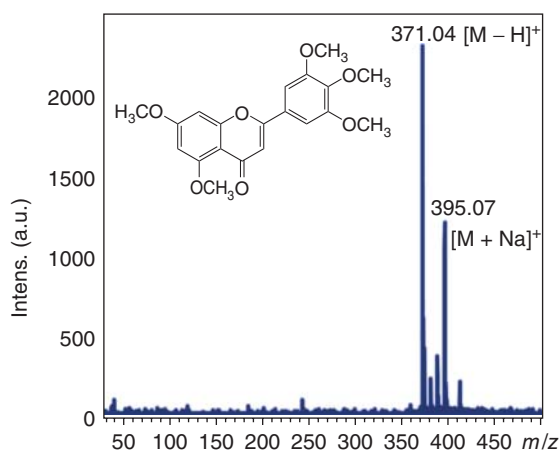
Figure 3 Gold film-assisted LDI mass spectrum (positive mode) of 3',4',5',5,7-pentamethoxyflavone ($C_{20}H_{20}O_7$, monoisotopic mass: 372.1209) acquired at a laser energy of 40% (38.4 µJ). Solely analyte-related signals are detected.

the essential aspects for choosing a working method will be discussed.

2.2 Matrix-Assisted LDI (MALDI)

2.2.1 The Concept

As previously outlined, the concept of MALDI MS (Figure 1) was firstly introduced by Hillenkamp and Karas in the mid-1980s. Since then, MALDI MS has become a powerful analytical tool in MS. Although hyphenations with chromatographic techniques are rare and require sophisticated instrumentation, MALDI MS represents a very useful and most sensitive technique of soft ionization in analytical chemistry. Further, when coupled with a TOF analyzer, the method facilitates accurate mass analysis and consequently access to molecular formulae.

However, while the method has been successfully applied to characterize large organic molecules such as proteins, sugars, and polymers, its utilization for small molecules ($\leq$600 Da) is significantly impaired by the coformation of matrix ions in the same mass region (Figure 4).

The matrix – most generally a small UV-absorbing compound – plays a key role in MALDI by absorbing laser energy and redirecting it to the analyte molecules. This leads to ion formation through complex processes, which have been intensively
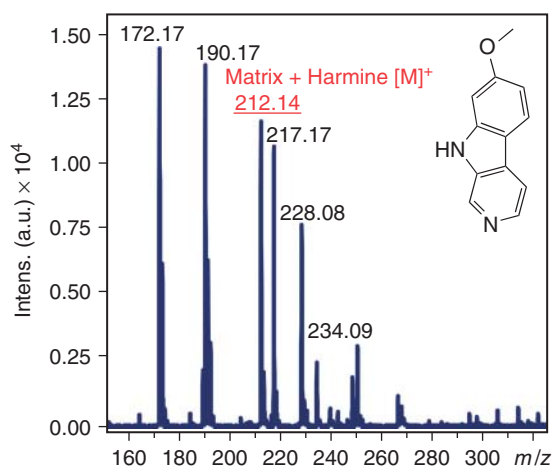
Figure 4 MALDI-TOF mass spectrum (positive mode) of the alkaloid harmine ($C_{13}H_{12}N_2O$, monosiotopic mass: 212.0950). The spectrum was recorded using α-cyano-4-hydroxycinnamic acid (CHCA) [$C_{14}H_{10}O_3$, monosiotopic mass: 179.0426] as the matrix. The low mass region is very crowded and the molecular ion of harmine is buried below a matrix-related signal also showing up at 212 *m/z*. Consequently, it is impossible to distinguish between the two signals.

studied and for which various mechanisms have been proposed (Beavis, Chait, and Fales, 1989; Beavis, Chaudhary, and Chait, 1992). In principle, the latter can be divided into two groups: Processes that are taking place immediately after laser irradiation (5–20 ns, primary ion formation) and those occurring at a later stage (0.2–5 µs) in the MALDI plume (secondary ion formation). In any case, ionization is based on the interaction of matrix and analyte molecules after laser irradiation, inducing the formation of their respective ions. Most commonly utilized matrices such as the popular 2,5-dihydroxybenzoic acid (DHB) or α-cyano-4-hydroxycinnamic acid (CHCA) (Figure 5) and also many others display a molecular mass in the range 0–600. The formation of matrix ions therefore severely impairs the detection of small molecules in that mass region. For this reason, MALDI-TOF MS is predominantly applied for the analysis of macromolecules.

2.2.2 Attempts to Minimize Matrix Peaks – Desorption/Ionization on Self-Assembled Monolayer Surfaces (DIAMS)

Successful attempts to reduce or eliminate matrix noise by lowering the applied laser energy or
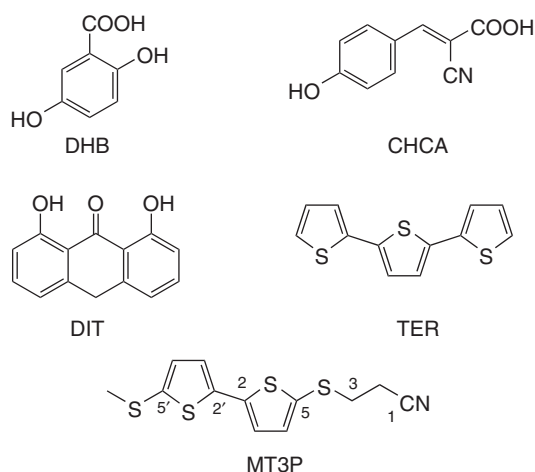
Figure 5 Some matrix molecules used in MALDI-TOF mass spectrometry. DHB, 2,5-dihydroxy benzoic acid; CHCA, α-cyano-4-hydroxycinnamic acid; DIT, dithranol (anthralin); TER, tertiophene; MT3P, 3-[5′-(methylthio)-2, 2′-bithiophen-5-ylthio]propanenitrile.

modifying the analyte to matrix mixing ratio have been reported (Beavis, Chait, and Fales, 1989; Strupat, Karas, and Hillenkamp, 1991; Ayorinde *et al*., 1999). According to these studies, matrix to analyte ratios of ≤ 100 and low laser energy powers may induce a so-called "*matrix suppression effect*" (MSE). Unfortunately, an efficient MSE is frequently accompanied by an "analyte suppression effect," as both mechanisms share a common underlying principle (Knochenmuss *et al*., 1996; McCombie and Knochenmuss, 2004). This is of particular relevance when several analytes are present in the same sample. Consequently, MSE may yield cleaner spectra, but signal intensities of certain analytes might be significantly reduced or even entirely suppressed.

Over recent years, desorption/ionization on self-assembled monolayer surfaces (DIAMS) has been described as an alternative method for eliminating matrix noise from MS spectra (Sanguinet *et al*., 2006; Bounichou *et al*., 2008). This method replaces classical matrix molecules by light absorbing self-assembled monolayers (SAMs), which are covalently bonded to a gold surface (Figure 6).

In order to facilitate optimal energy transfer, these SAMs are equipped with a strong UV chromophore, for example, a bithiophene moiety, exhibiting an absorption maximum in the range of the emission wavelength of a standard nitrogen laser (337 nm), which is used in many MALDI instruments. In detail, SAMs for DIAMS were build up from a 5′-*S*-substituted 5-*S*-methyl-2,2′-bithiophene as the redox chromophore, linked to a spacer terminated by a thiol, the anchoring functional group to the gold surface. The spacer must be carefully designed to provide high stability of the SAMs. Moreover, the synthesis of the organic precursors may be easily adapted for studying various SAMs, with modified properties. Following this approach, the concept of click chemistry recently developed by Sharpless (Rostovtsev *et al*., 2002) and used for facilitating high yield reactions under soft conditions was applied for the synthesis of the latest generation of SAMs for DIAMS. Indeed, interlocking azidoalkylbithiophenes and alkynthiols of various lengths led, by a convergent click reaction, to triazolobithiophene thiols. With a triazol ring in the linker that might exihibit $\pi-\pi$ stacking, the corresponding SAMs showed high surface coverage and stability (Kenfack *et al*., 2011).

As previously mentioned, the bithiophene moiety with its specific absorption properties represents a core feature of DIAMS. In principle, the light-absorbing monolayer is designed to act like a classical matrix by taking up laser energy and subsequently redirecting it to the sample but without getting desorbed from the gold surface. Consequently, no matrix ions should be present in the plume, yielding spectra solely exhibiting analyte signals.

Hitherto DIAMS MS was successfully applied for the detection of fatty acids, glycerides, and salicylic acid (Bounichou *et al*., 2008). One report directly compared signal intensities of triglycerides obtained by DIAMS and GFLDI. Out of eight analytes, five displayed stronger analyte signals in GFLDI, whereas two showed better results in DIAMS. For one sample, no difference was detected. Further the alkaloid totum extracted from *Thalictrum flavum* L. (Ranunculaceae) was successfully detected by DIAMS. The "common meadow rue" or "yellow meadow rue" is an herb of 0.5–1.5 m height growing in Europe in tall grass meadows, ditches, marshes, and other habitats close to water bodies (Ropivia *et al*., 2010). The isoquinolines shown in Figure 7 presents key compounds of the crude extract and LC sub-fractions that could be directly indentified by DIAMS-TOF MS without further processing.

However, it has also been observed that, contrary to classical MALDI-TOF spectrometry, DIAMS-TOF mass spectra frequently suffer from
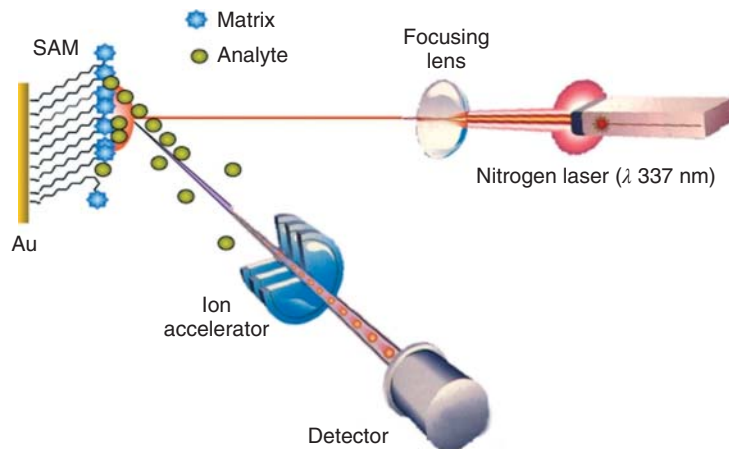
Figure 6 The principle of DIAMS MS. Desorption/ionization on self-assembled monolayer surfaces (DIAMS) mass spectrometry: A self-assembled monolayer (SAM) incorporating the matrix is covalently bound to a gold surface.
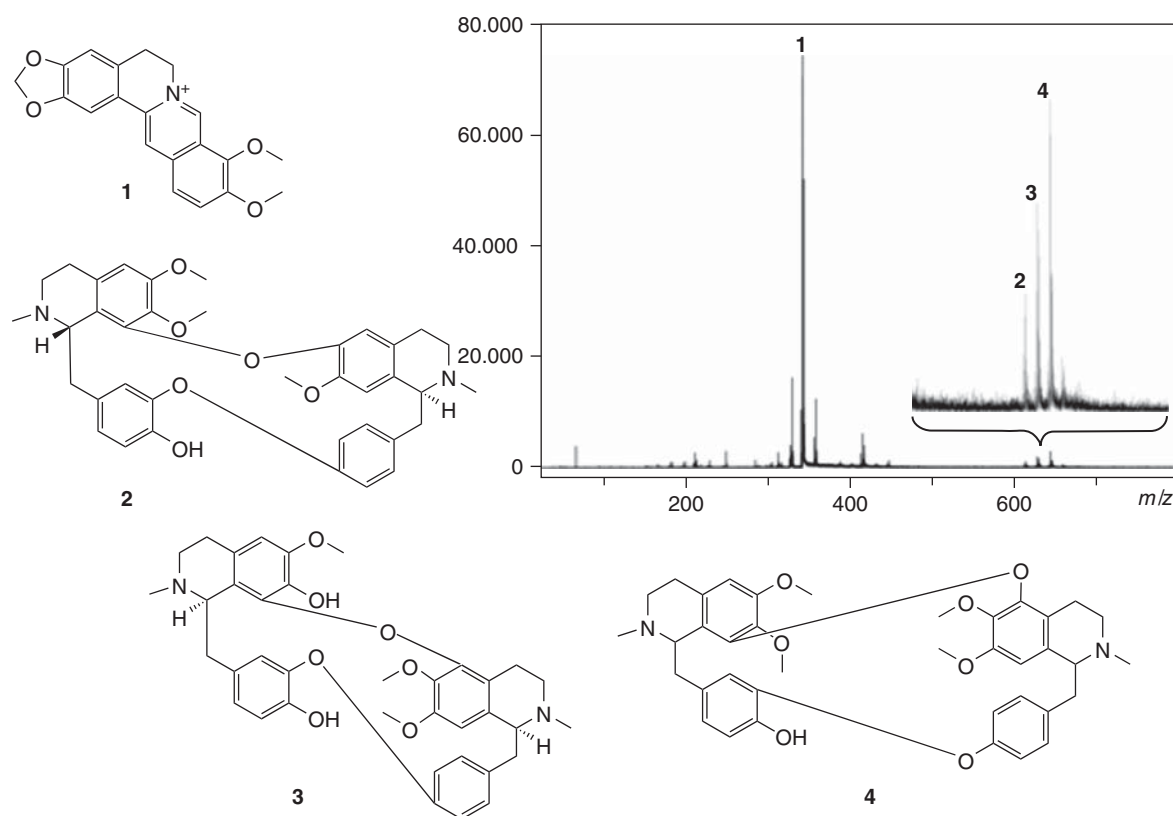
Figure 7 DIAMS mass spectrum (positive mode) of an alkaloid totum extracted from *Thalictrum flavum* L. (Ranunculaceae). Key compounds of the crude extract and LC subfractions: berberine (**1**), thalicberine (**2**), thaligoside (**3**), and thalfoetidine (**4**).
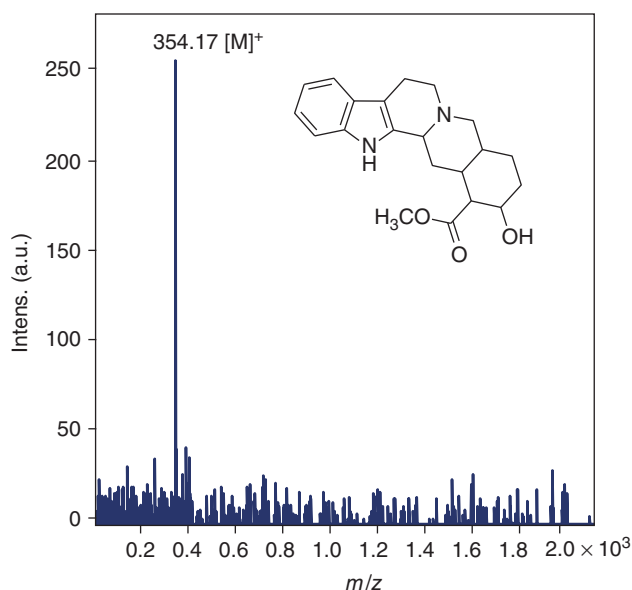
Figure 8 DIAMS-TOF mass spectrum (positive mode) of yohimbine ($C_{21}H_{26}N_2O_3$, calculated monoisotopic mass: 354.1943, observed: 354.17). Laser energy 30% (30.8 μJ).

low signal-to-noise ratios as shown by the spectrum of the indole alkaloid yohimbine (Figure 8).

Therefore, a study on the elucidation of ion formation processes occurring on the surface of light absorbing SAMs was performed. Ion yields obtained by free and immobilized matrix molecules and those generated by GFLDI (*vide supra*) were compared (Schinkovitz *et al*., 2011). In short, this study showed that the formation of strong analyte signals is essentially linked to the presence of "free" matrix molecules. Immobilizing the latter redirected the absorbed laser energy away from the analyte molecules. This effect inversely correlated with the surface coverage of SAMs and could be quantified through cyclic voltammetry (Figure 9).

However, seen from another perspective, the phenomenon may offer potential applications for analyzing particularly sensitive sample material. The impact of laser light, indeed, may be limited to just the extent necessary for inducing ionization while minimizing dimerization or degradation processes. Further, the very organized surface structure of SAMs seems to facilitate a most homogeneous distribution of sample material. Consequently, and unlike in classical MALDI, the observed MS signals acquired from different spots of the sample deposition area tentatively showed little variation in signal intensity.

This aspect might be of high interest for quantitative applications.

2.2.3 Specific Matrices for Small Compounds: A New Concept Starting from MT3P

Following the rapid expansion of proteomic studies, considerable efforts have been made for the development and application of MALDI-TOF MS, whereas fundamental understanding of underlying processes is still limited. This is because MALDI is quite a complex phenomenon, where sample material embedded into a solid, mostly crystalline, matrix is instantly vaporized and consecutively diluted from a condensed state to a very thin aerosol. MALDI ionization itself predominantly takes place in the gaseous phase and is known to be particularly soft, minimizing the formation of fragmentation and dimerization products. Consequently, as far as matrices are concerned, MALDI-TOF was developed based on empirical approaches. Indeed, early efforts were made to identify "good matrices." During these experiments, putative candidates were briefly tested on a few analytes and given qualitative ratings. Nowadays, actually very few of them are in common use. They were identified early on (Knochenmuss,
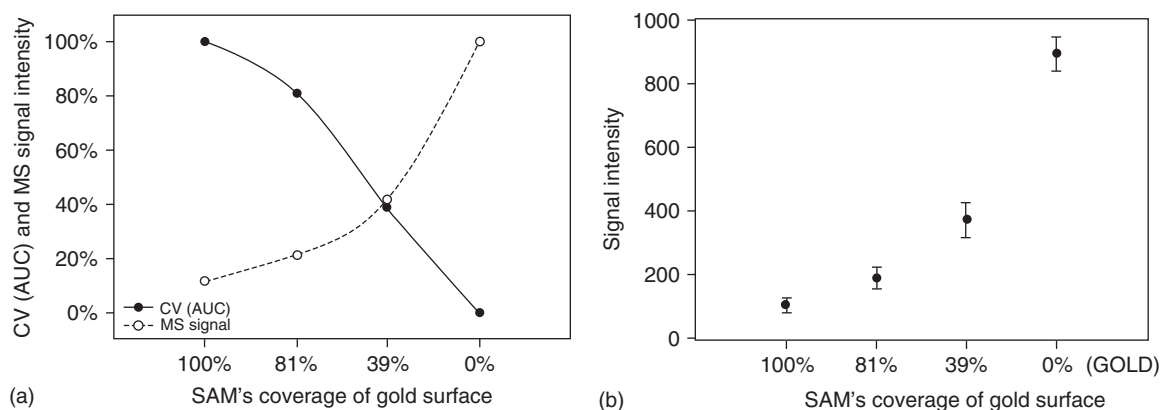
Figure 9 Inverse correlation of SAM "quality" (surface coverage) and MS signal intensity observed for the alkaloid claviculine ($C_{18}H_{19}NO_4$, monoisotopic mass: 313.1314). As the bithiophene moiety exhibits a reversible redox system, SAM degradation may be monitored by cyclic voltammetry (CV, **A**). Charge transfers and consequently ion intensities clearly increase with the decay of the SAM surface. On the other hand, the small standard deviation indicates that the SAM's surface facilitates an excellent sample distribution, limiting "sweet spot" phenomena known from classical MALDI (**B**). This effect might be beneficial for quantitative applications. (*Source*: A. Schinkovitz, *et al*. (2011). Reproduced from John Wiley & Sons, Ltd.)

2006) and are today mainly applied to the analysis of peptides, proteins, polymers, and other macromolecules. Further, certain matrices show some substrate specificity, another favorable feature of MALDI, which may allow the selective detection of target molecules in complex mixtures. Hitherto, the latter has been primarily reported for cinnamic acid-based matrices targeting large molecules such as peptides, proteins, or lipids (Beavis, Chait, and Fales, 1989; Beavis, Chaudhary, and Chait, 1992; Strupat, Karas, and Hillenkamp, 1991). In contrast, very little is known about small molecules, as their analysis is generally impaired by the coformation of matrix ions. Many of most commonly utilized MALDI matrices exhibit their molecular ions in the range 0–600 *m/z*; therefore, this region is highly critical for correct assignment of analyte and matrix signals. Thus, with respect to small molecules, the reduction of matrix noise is almost equally important as the adequate ionization of analytes by matrix molecules. Various strategies have been discussed in order to improve spectrum qualities for low mass molecules in MALDI. Some matrices (Figure 5) such as 1,8-dihydroxy-9,10-dihydroanthracen-9-one (DIT) produce fewer signals than others, for example, CHCA. Further, compounds of higher molecular weights such as *meso*-tetrakis(pentafluorophenyl) porphyrin may be utilized, provided they show sufficient ionization of target molecules (Ayorinde *et al*.,

1999). Alternatively, the so-called "*MSEs*" (*vide supra*) may offer useful strategies to minimize the formation of matrix noise in MALDI (Knochenmuss *et al*., 1996; McCombie and Knochenmuss, 2004). These effects have been intensively discussed in literature, and approaches to limit matrix noise range from specifically adapted matrix molecules or mixtures of matrices to matrix noise-inhibiting additives (Guo *et al*., 2002; Vaidyanathan, Gaskell, and Goodacre, 2006; Komori *et al*., 2009; Fujita *et al*., 2010). Further, a simple decrease of applied laser energy may improve spectrum qualities, and also matrix to analyte ratios seem to be of significant relevance. Following this approach, McCombie and Knochenmuss have introduced the "MSE score", a simple but powerful equation for evaluating spectrum qualities, which provides a valuable tool for spectrum optimization. In this equation, the sum of all analyte-related signals are divided by the sum of all detected signals (analyte + matrix) yielding a factor between 0 and 1. MSE scores close to 1 indicate strong analyte signals with low matrix noise, whereas factors around 0 represent high matrix noise and low analyte signals.

With these elements in mind, we decided to explore a new concept of LDI, that is, matrices that would specifically interact with compounds of a selected chemical family. With focus on secondary metabolites from natural products,

Table 1   Fifty-five analyzed compounds by MALDI-TOF MS using MT3P as the matrix; originally acquired signal intensities for molecular ions were divided by 1250 for illustration purposes.

Compound	MT3P	STD_{MT3P}	DIT	STD_{DIT}	CHCA	STD_{CHCA}	TER	STD_{TER}	DHB	STD_{DHB}	MF	MT3P versus CHCA
Acetylsalicylic acid[LN10]	nd	—	nd	—	nd	—	nd	—	nd	—	$C_9H_8O_4$	—
Aconitine[a]	14.00	4.86	0.22	0.09	3.90	1.17	0.48	0.21	nd	—	$C_{34}H_{47}NO_{11}$	Yes $P = 0.002$
Amentoflavone	17.68	3.76	0.25	0.10	14.24	7.44	0.18	0.03	nd	—	$C_{30}H_{18}O_{10}$	No $P = 0.589$
Angiotensin II	nd	—	nd	—	10.85	4.49	nd	—	nd	—	$C_{50}H_{71}N_{13}O_{12}$	Yes $P = 0.002$
Atropine[a]	18.61	5.24	0.04	0.03	9.67	2.72	0.02	0.01	nd	—	$C_{17}H_{23}NO_3$	Yes $P = 0.009$
Aucuparin	nd	—	nd	—	nd	—	nd	—	nd	—	$C_{14}H_{14}O_3$	—
β-Carotene	0.24	0.16	0.01	0.00	0.15	0.05	0.69	0.27	nd	—	$C_{40}H_{56}$	No $P = 0.198$
Benzocaine	nd	—	nd	—	1.51	0.46	nd	—	nd	—	$C_9H_{11}NO_2$	Yes $P = 0.002$
Berberine[a]	23.56	5.67	0.34	0.16	18.75	4.48	6.16	5.49	0.01	0.01	$C_{20}H_{18}NO_4$	No $P = 0.134$
Bergaptene	0.07	0.02	nd	—	8.72	3.88	nd	—	nd	—	$C_{12}H_8O_4$	Yes $P = 0.002$
Boldine[a]	11.22	4.32	0.20	0.24	7.94	5.25	0.36	0.30	nd	—	$C_{19}H_{21}NO_4$	No $P = 0.265$
Caffeic acid [LN10]	2.48	2.30	nd	—	nd	—	1.30	0.79	nd	—	$C_9H_8O_4$	Yes $P = 0.002$
Caryophyllen[LP15]	nd	—	nd	—	nd	—	nd	—	nd	—	$C_{15}H_{24}$	—
Chlorogenic acid [LN10]	nd	—	nd	—	0.02	0.02	nd	—	nd	—	$C_{16}H_{18}O_9$	Yes $P = 0.015$
Cholesterol	nd	—	nd	—	nd	—	nd	—	nd	—	$C_{27}H_{46}O$	—
Claviculine[a]	20.43	4.75	0.19	0.11	6.51	1.85	1.74	2.74	nd	—	$C_{18}H_9NO_4$	Yes $P = 0.002$
Codeine[a]	12.07	3.59	0.12	0.08	16.58	3.62	0.65	0.40	nd	—	$C_{18}H_{21}NO_3$	No $P = 0.056$
Colchicine[a]	12.11	3.77	0.45	0.21	2.98	1.39	0.12	0.07	nd	—	$C_{22}H_{25}NO_6$	Yes $P < 0.001$
Coumarin	nd	—	nd	—	1.12	0.36	nd	—	nd	—	$C_9H_6O_2$	Yes $P = 0.002$
Curcumin	5.41	0.43	0.29	0.15	4.77	2.65	0.05	0.06	nd	—	$C_{21}H_{20}O_6$	No $P = 0.573$
Digitoxin[LP20]	8.27	4.07	1.22	0.92	3.34	1.83	4.42	1.67	nd	—	$C_{41}H_{64}O_{13}$	Yes $P = 0.022$
1.3-Dipalmitoyl-glycerol	nd	—	nd	—	nd	—	nd	—	nd	—	$C_{35}H_{68}O_5$	—
Emetine[a]	28.02	1.37	0.32	0.45	5.34	2.42	0.18	0.15	nd	—	$C_{29}H_{40}N_2O_4$	Yes $P < 0.001$
E-Notopterol	nd	—	nd	—	nd	—	nd	—	nd	—	$C_{21}H_{22}O_5$	—
Fumaric acid [LN10]	0.31	0.00	nd	—	nd	—	nd	—	nd	—	$C_4H_4O_4$	—
Fumaritine[a]	25.72	2.90	0.23	0.10	2.14	2.15	1.29	1.08	nd	—	$C_{20}H_{21}NO_5$	Yes $P < 0.001$
Geraniol [LP15]	nd	—	nd	—	nd	—	nd	—	nd	—	$C_{10}H_{18}O$	—
Glyceryl-1,3-distearate	nd	—	nd	—	nd	—	nd	—	nd	—	$C_{39}H_{76}O_5$	—
Harmine[a]	24.68	3.42	2.39	2.43	9.34	2.83	2.57	1.16	0.01	0.00	$C_{13}H_{12}N_2O$	Yes $P < 0.001$
Hesperidin [LP15]	0.24	0.30	0.08	0.06	0.82	0.67	0.10	0.10	nd	—	$C_{28}H_{34}O_{15}$	Yes $P < 0.041$
L-Hyoscyamine[a]	11.26	3.02	0.32	0.27	19.38	3.38	0.61	0.30	nd	—	$C_{17}H_{23}NO_3$	Yes $P < 0.001$
Isoimperatorine	nd	—	nd	—	5.47	13.41	nd	—	nd	—	$C_{16}H_{14}O_4$	Yes $P = 0.002$
Khellin	2.32	1.46	nd	—	12.95	6.56	nd	—	nd	—	$C_{14}H_{12}O_5$	Yes $P = 0.003$
Leucine enkelphalin	nd	—	nd	—	9.10	2.82	nd	—	nd	—	$C_{28}H_{37}N_5O_7$	Yes $P = 0.002$
Limogine[a]	26.03	2.46	0.52	0.19	7.92	1.69	0.09	0.07	nd		$C_{20}H_{17}NO_5$	Yes $P < 0.001$
Nicotine[a]	0.43	0.20	0.01	0.01	1.29	0.73	0.01	0.00	nd	—	$C_{10}H_{14}N_2$	Yes $P = 0.018$
Pentamethoxyflavone	26.00	4.27	2.06	0.36	13.19	5.24	1.70	1.77	0.00	0.00	$C_{20}H_{20}O_7$	Yes $P < 0.001$
Pilocarpine[a]	11.03	4.07	0.86	0.80	7.56	2.81	0.03	0.01	nd	—	$C_{11}H_{16}N_2O_2$	No $P = 0.117$
Pregnolon	nd	—	nd	—	nd	—	nd	—	nd	—	$C_{21}H_{32}O_2$	—
Quercetine	0.15	0.08	0.05	0.03	14.47	3.41	0.00	0.01	nd	—	$C_{15}H_{10}O_7$	Yes $P = 0.002$
Quinidine[a]	20.96	6.16	2.05	1.24	3.08	1.99	0.15	0.11	nd	—	$C_{20}H_{24}N_2O_2$	Yes $P < 0.001$
Rutin[LP20]	3.63	2.24	0.65	0.20	3.92	2.44	0.23	0.19	nd	—	$C_{27}H_{30}O_{16}$	No $P = 0.837$
Scopolamine[a]	13.56	2.59	0.10	0.08	3.03	1.92	0.05	0.02	nd	—	$C_{17}H_{21}NO_4$	Yes $P < 0.001$
Senecionine[a]	12.05	3.73	0.14	0.08	3.64	2.55	0.21	0.24	nd	—	$C_{18}H_{25}NO_5$	Yes $P < 0.001$
β-Sitosterol	nd	—	nd	—	nd	—	nd	—	nd	—	$C_{29}H_{50}O$	—

Table 1 (*Continued*)

Compound	MT3P	STD_{MT3P}	DIT	STD_{DIT}	CHCA	STD_{CHCA}	TER	STD_{TER}	DHB	STD_{DHB}	MF	MT3P versus CHCA
Sparteine[a]	20.34	5.34	0.92	0.41	15.32	3.63	2.43	2.11	0.00	0.00	$C_{15}H_{26}N_2$	No $P = 0.086$
Strychnine[a]	13.16	5.72	0.55	0.23	13.07	4.29	0.29	0.08	nd	—	$C_{21}H_{22}N_2O_2$	Yes $P < 0.001$
Stylopine[a]	24.84	2.19	0.43	0.39	7.65	1.0	1.22	0.61	nd	—	$C_{19}H_{17}NO_4$	Yes $P < 0.001$
Thalfoetidine[a]	5.67	1.82	0.07	0.06	3.51	1.46	1.09	0.96	nd	—	$C_{38}H_{42}N_2O_7$	Yes $P = 0.046$
Thalicberine[a]	20.54	5.22	0.07	0.05	3.90	2.59	0.87	0.78	nd	—	$C_{37}H_{40}N_2O_6$	Yes $P < 0.001$
Thaligosidine[a]	4.40	0.79	0.06	0.02	4.75	1.62	0.12	0.08	nd	—	$C_{37}H_{40}N_2O_7$	No $P = 0.641$
Thebaine[a]	17.58	3.84	0.37	0.17	10.92	4.81	0.81	0.57	0.00	0.01	$C_{19}H_{21}NO_3$	Yes $P = 0.024$
Theobromine[a]	0.08	0.11	nd	—	8.58	4.16	0.01	0.01	nd	—	$C_7H_8N_4O_2$	Yes $P = 0.002$
Umbelliferone	nd	—	nd	—	nd	—	nd	—	nd	—	$C_{17}H_{21}NO_4$	—
Yohimbine[a]	16.31	4.32	0.21	0.16	6.35	3.14	1.14	1.01	nd	—	$C_{21}H_{26}N_2O_3$	Yes $P = 0.001$

nd, Signal not detected; STD, standard deviation $n = 6$; MF, molecular formula; MT3P versus CHCA, significance (yes/no) of observed differences between MT3P and CHCA (*t*-test); except for designated molecules, all samples were analyzed in the linear positive mode at a laser energy of 10% (15.6 μJ). [LN10], linear negative mode, laser energy 10% (15.6 μJ); [LP15], linear positive mode, laser energy 15% (19.4 μJ); [LP20], linear positive mode, laser energy 20% (23.2 μJ). Zero numbers indicate that molecular ions were detected but at low intensities.

[a] Alcaloid.

alkaloids present a most interesting analytical target, as these compounds are associated with a wide range of pharmacological and toxic activities. Alkaloids are well known to show dipolar interactions with nitriles, and this particular behavior is successfully exploited for their separation on cyano-bonded phase HPLC columns. Thus, coupling a propanenitrile to a bithiophen derivative that strongly absorbs UV laser light at 337 nm (the typical wavelength of a standard nitrogen laser commonly utilized in MALDI-TOF instruments) yielded 3-[5′-(methylthio)-2,2′-bithiophen-5-ylthio]propanenitrile (MT3P, Figure 5). According to recent experiments (Schinkovitz *et al*., 2012), the compound represents a new matrix molecule exhibiting an unusual selectivity towards alkaloids. So far MT3P has been tested on 55 compounds of various chemical origin and results were compared with those obtained using commercial matrices such as dithranol (DIT), CHCA, 2,2′:5′,2″-terthiophene (TER), and DHB (Figure 5 and Table 1).

Apart from alkaloids, the selection of natural products included coumarins, terpenes, flavonoids, carotenoids, steroids, and peptides. Results from this survey are summarized in Table 1, providing a general overview of the ionization properties of MT3P. Alkaloids almost constantly displayed intense molecular ion signals. Solely nicotine and theobromine were not or weakly detected by MT3P. With regards to ion formation, $[M + H]^+$ was observed for any investigated alkaloid except for claviculine. The latter displayed a quasi-molecular ion of $[M - H]^+$. On first

sight, the formation of $[M - H]^+$ may appear quite unusual, but actually it represents a standard ion originating from photoionization process associated with MALDI (Zenobi and Knochenmuss, 1998).

Beside others, the list of successfully tested alkaloids contained highly biologically active compounds such as codeine, aconitine, L-hyoscyamine, and emetin. Figure 10 displays the spectrum of quinidine and the simultaneous detection of colchicine, L-hyoscyamine, and senecionine (A).

As further shown in Figure 10 and similarly to many MALDI matrices, MT3P exhibits its matrix ions within a mass range 0–600 *m/z*. Most prominent signals were observed at 297 *m/z* (M$^+$) and at 243 *m/z* (fragmentation product resulting from the cleavage of the S–C bond in position 3). Further, two minor signals showed up at 312 and 351 *m/z*, respectively. Nevertheless, the intensity of the observed alkaloid molecular ions was mostly superior and these signals may be used for the accurate calibration of the TOF analyzer. Consequently, alkaloids were easily and precisely detected even if their molecular ions showed up in the vicinity of matrix signals.

Any of the successfully tested alkaloids exhibited strong molecular ions at a relatively low level of laser energy (10% = 15.6 μJ). Furthermore, with respect to these alkaloids, MT3P usually showed superior ionizing properties in comparison to commonly utilized matrices such as DIT, CHCA, TER, and DHB (Figure 5). Solely CHCA was able to ionize as many alkaloids as MT3P, but generally exhibited weaker analyte signals than
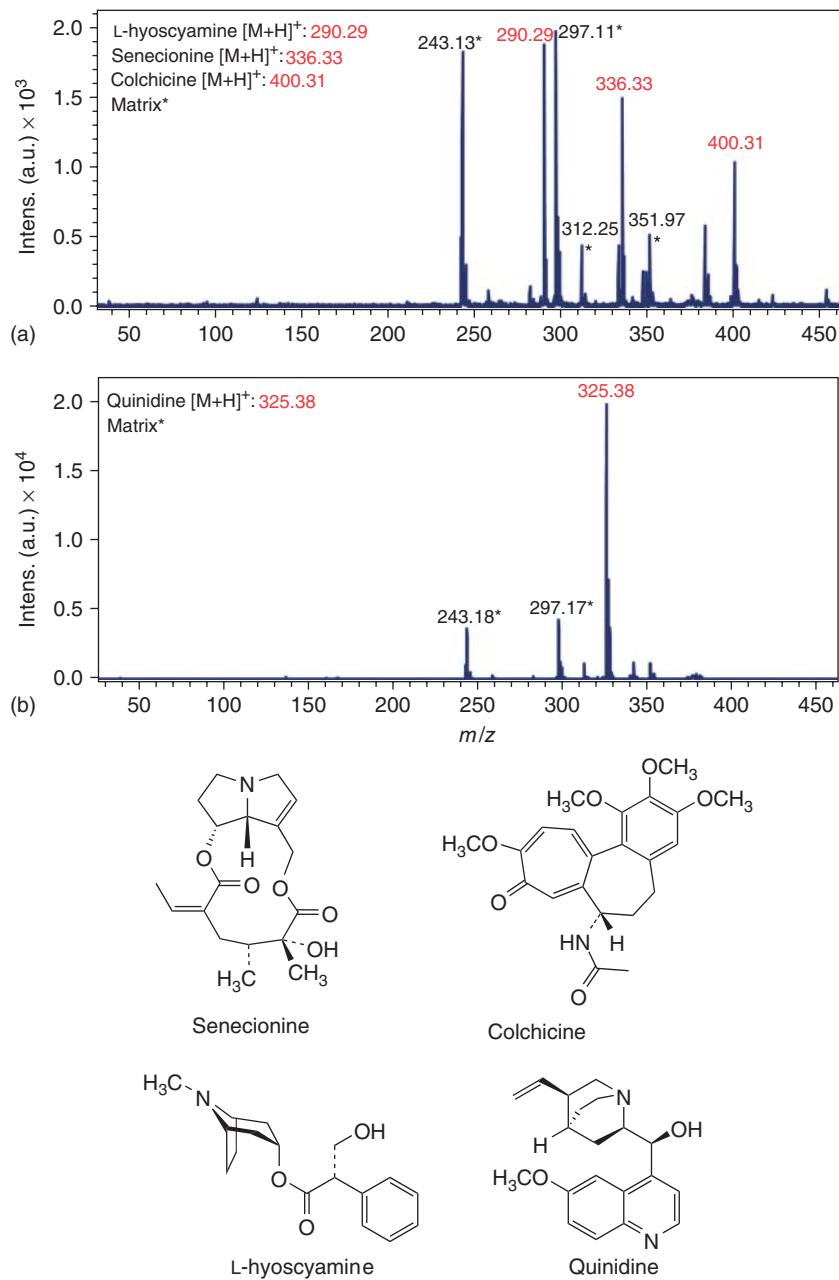
Figure 10 (a) MALDI-TOF mass spectra (positive mode) of selected alkaloids: colchicine, L-hyoscyamine, quinidine, and senecionine using MT3P as the matrix. Simultaneous detection of toxic alkaloids L-hyoscyamine, senecionine, and colchicine (**A**) and spectrum of quinidine as single compound (**B**) (* denotes matrix ions). (b) The structures of the alkaloids senecionine, colchicines, L-hyoscyamine, and quinidine. (*Source*: Anal Bioanal Chem, 403, 2012, 1697-1705, Selective detection of alkaloids in MALDI-TOF: the introduction of a novel matrix molecule, Schinkovitz A., Kenfack G.T., Seraphin D., Levillain E., Dias M. and Richomme P. With kind permission of Springer Science + Business Media.)

MT3P. For the majority of these cases (16 out of 20), the observed difference was significant (*P* values: 0.046 to ≤0.001). In addition, berberine, boldine, pilocarpine, and sparteine showed stronger signal intensities when analyzed by MT3P; however, in comparison to CHCA, the difference was not statistically significant. Six alkaloids namely codeine, L-hyoscyamine, nicotine, thaligosidine, theobromine, and strychnine showed better ionization when using CHCA, but only for L-hyoscyamine, nicotine, and theobromine was the observed difference significant.

Contrary to its excellent ionizing properties of alkaloids, MT3P did barely interact with any compound of different chemical origin. There was apparently strong ionization of pentamethoxyflavone and amentoflavone, but these compounds could also be ionized by simple LDI, without any matrix support, as they both include a strong UV chromophore. A contrario, digitoxine that did not show any LDI could be ionized by MT3P, but required a higher laser energy (20% or 23.2 μJ) and led to a crowded spectrum thus exhibiting a low MSE score (≤0.05). In this respect, the selective setting of laser energy levels provides a useful tool for suppressing the ionization of compounds such as digitoxine that require higher amounts of energy.

The precise mode of action of MT3P still remains unknown but certain structural features may provide useful hints for its elucidation. Considering that solely MT3P and CHCA showed significant interaction with alkaloids, it may be speculated whether the presence of a single free nitrile group enhances the ionization of alkaloids. On the other hand, MT3P unlike CHCA did not show any ionization of peptides, suggesting that the observed selectivity could be due to a more complex mechanism.

Apart from its ionizing properties, an "optimized" MALDI matrix should exhibit strong analyte signals together with reduced formation of matrix ions. Therefore, MT3P and CHCA, representing the best working matrices for alkaloids, were directly compared in terms of their respective MSE scores. These experiments showed that MT3P exhibited cleaner spectra with reduced matrix noise for 18 out of 25 alkaloid samples (Table 2).

Subsequent experiments further revealed that MT3P should ideally be utilized at energy levels between 5% and 15% (11.8–19.4 μJ). Within this energy range, alkaloids displayed strong molecular

Table 2 MSE scores of alkaloids utilizing matrices MT3P and CHCA.

Compound	MSE score			STD$_{CHCA}$
	MT3P	STD$_{MT3P}$	CHCA	
Aconitine	0.83	0.10	0.76	0.23
Atropine	0.30	0.10	0.39	0.13
Berberine	0.64	0.14	0.97	1.86
Boldine	0.46	0.06	0.22	0.19
Claviculine	0.26	0.09	0.18	0.07
Codeine	0.22	0.01	0.14	0.04
Colchicine	0.40	0.05	0.08	0.06
Emetine	0.59	0.08	0.15	0.08
Fumaritine	0.30	0.05	0.22	0.06
Harmine	0.27	0.07	0.35	0.15
L-hyoscyamine	0.58	0.10	0.23	0.06
Limogine	0.27	0.05	0.71	0.10
Nicotine	0.01	0.00	0.06	0.04
Pilocarpine	0.17	0.07	0.12	0.02
Quinidine	0.30	0.04	0.17	0.02
Scopolamine	0.14	0.02	0.12	0.04
Senecionine	0.17	0.01	0.10	0.06
Sparteine	0.32	0.02	0.79	0.52
Strychnine	0.43	0.08	0.44	0.08
Stylopine	0.79	0.12	0.37	0.25
Thalfoetidine	0.56	0.31	0.16	0.10
Thalicberine	0.56	0.16	0.16	0.06
Thaligosidine	0.85	0.13	0.24	0.28
Theobromine	0.00	0.00	0.19	0.06
Yohimbine	0.57	0.28	0.15	0.01

STD, Standard deviation.
All data were acquired in the linear positive mode at a laser energy of 10% (15.6 μJ).

ions, whereas the formation of matrix ions was still limited.

As previously discussed, MSEs are linked to matrix to analyte mixing ratios and laser energy (Knochenmuss *et al.*, 1996; McCombie and Knochenmuss, 2004). Considering that many alkaloids displayed very strong molecular ions when being exposed to a laser energy of 10% (15.6 μJ), a decrease in laser power could result in less matrix noise while still providing sufficient intensities of analyte signals. Therefore, 14 alkaloids of various chemical structures were subsequently analyzed at a reduced level of laser energy of 5% (11.8 μJ) (Figure 11a). As expected, signal intensities decreased with declining laser energy, but any compound was still clearly detectable. In return, MSE scores constantly increased, and for some compounds such as fumarithine (0.3–0.5), thalicberine (0.56–0.79), or emetine (0.59–0.85), the increase was particularly remarkable (Figure 11b).
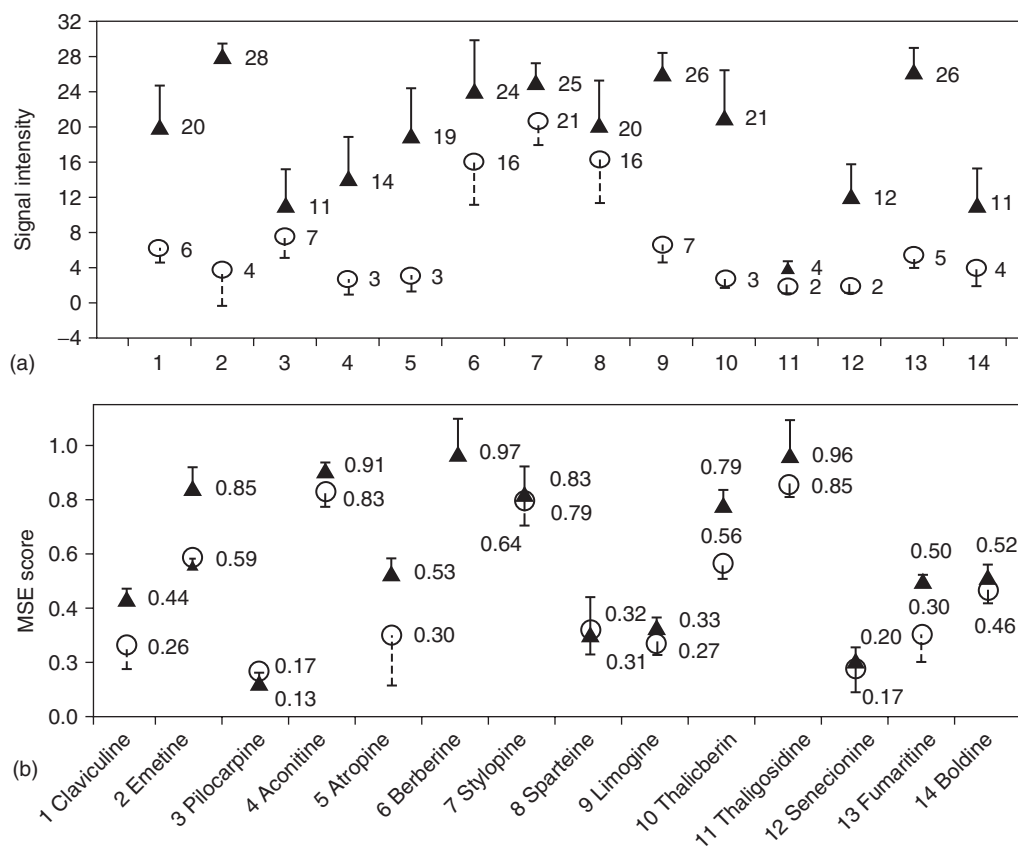
Figure 11 Signal intensity versus MSE scores using MT3P as the matrix in MALDI-TOF MS. Comparison of signal intensities and MSE scores for 14 alkaloids utilizing MT3P at different levels of applied laser energy: (a) ▲, Signal intensity at 10% of laser energy (15.6 μJ); ○ Signal intensity at 5% of laser energy (11.8 μJ). (b) ▲, MSE score at 5% of laser energy (15.6 μJ); ○, MSE score at 10% of laser energy (11.8 μJ). (*Source*: Anal Bioanal Chem, 403, 2012, 1697-1705, Selective detection of alkaloids in MALDI-TOF: the introduction of a novel matrix molecule, Schinkovitz A., Kenfack G.T., Seraphin D., Levillain E., Dias M. and Richomme P. With kind permission of Springer Science + Business Media.)

Exceptionally, two compounds showed unchanged (sparteine) or slightly declining MSE scores (pilocarpine) when being analyzed at 5% of laser energy. Seemingly, these alkaloids responded more sensitively to a reduction of energy than MT3P and should be analyzed at higher levels of laser energy. Overall, no clear correlation between the extent of signal decay and the raise of MSE scores could be observed. A substantial decrease in signal intensities did not necessarily imply a remarkable increase of MSE scores and vice versa. Eventually, it can be concluded that a reduction of laser energy generally improved spectrum qualities, but the extent of the effect was hardly predictable. At this point, it shall be mentioned that all spectra were acquired at a certain analyte and

matrix concentration. As mentioned earlier, MSEs are highly dependent on sample-to-matrix mixing ratios, and changing the latter may have significantly altered observed MSE scores. This also applies to Table 2, in which MTP3 and CHCA were used at concentrations of 25.80 mM and 45.81 mM, respectively. In this respect, the precise evaluation of the impact of sample-to-matrix mixing ratios for MT3P represents an interesting subject of future research.

3 CONCLUSION

The selective detection of alkaloids by MALDI-TOF has been continuously and intensively discussed in

the literature (Sun *et al.*, 1998; Lopez-Legentil *et al.*, 2005; Cheng *et al.*, 2006; Wu *et al.*, 2007a, b; Araoz *et al.*, 2008; Feng and Lu, 2009; Lu *et al.*, 2010), certainly because alkaloids represent highly active natural and/or synthetic substances. Compounds such codeine and its derivatives are commonly used for medicinal applications, whereas others such as strychnine, emetine, and aconitine exhibit severe toxic effects. Therefore, their precise and accurate identification in crude plant material, medicinal preparations, or dietary supplements is highly demanded and subject of contemporary research. While some aromatic alkaloids such as ascididemine can be ionized by simple LDI (Lopez-Legentil *et al.*, 2005), most reports mention CHCA and TER as preferred matrices for MALDI experiments (Cheng *et al.*, 2006; Wu *et al.*, 2007a,b; Liu *et al.*, 2010). In addition, 7-mercapto-4-methylcoumarin has been described for the detection of arecoline and arecaidine, but no further alkaloids were tested (Feng and Lu, 2009). Considering these findings, MT3P represents a most valuable addition to the pool of available matrix molecules. Its enhanced and selective ionizing properties facilitate high quality MALDI spectra of alkaloids at exceptionally low levels of laser energy. Consequently, it may also be particularly suitable for the analysis of unstable compounds. Future studies and experiments may help to further improve the current working protocol allowing a gain in sensitivity and MSE scores. Finally, bithiophene-based molecules may represent an entire new group of matrix molecules. Starting from MT3P, chemical modifications may alter their ionizing profile and create specifically adapted molecules for various MALDI applications. This outlook represents a challenging but most promising perspective for future experiments.

ACKNOWLEDGMENTS

REFERENCES

Araoz, R., Guerineau, V., Rippka, R., *et al.* (2008) *Toxicon*, **51**, 1308–1315.

Ayorinde, F. O., Hambright, P., Porter, T. N., *et al.* (1999) *Rapid Commun. Mass Spectrom.*, **13**, 2474–2479.

Beavis, R. C., Chait, B. T. and Fales, H. M. (1989) *Rapid Commun. Mass Spectrom.*, **3**, 432–435.

Beavis, R. C., Chaudhary, T. and Chait, B. T. (1992) *Org. Mass Spectrom.*, **27**, 156–158.

Bounichou, M., Sanguinet, L., Elouarzaki, K., *et al.* (2008) *J. Mass Spectrom.*, **43**, 1618–1626.

Bounichou, M., Aleveque, O., Dias, M., *et al.* (2009) *J. Mater. Chem.*, **19**, 8032–8039.

Chen, Y. and Vertes, A. (2006) *Anal. Chem.*, **78**, 5835–5844.

Cheng, Z., Guo, Y., Wang, H., *et al.* (2006) *Anal. Chim. Acta*, **555**, 269–277.

Feng, C. and Lu, C. (2009) *Anal. Chim. Acta*, **649**, 230–235.

Fujita, T., Fujino, T., Hirabayashi, K., *et al.* (2010) *Anal. Sci.*, **26**, 743–748.

Guo, Z. and He, L. (2007) *Anal. Bioanal. Chem.*, **387**, 1939–1944.

Guo, Z., Zhang, Q., Zou, H., *et al.* (2002) *Anal. Chem.*, **74**, 1637–1641.

Karas, M. and Hillenkamp, F. (1988) *Anal. Chem.*, **60**, 2299–2301.

Kenfack, G. T., Schinkovitz, A., Babu, S., *et al.* (2011) *Molecules*, **16**, 8758–8774.

Knochenmuss, R. (2006) *Analyst*, **131**, 966–986.

Knochenmuss, R., Dubois, F., Dale, M. J., *et al.* (1996) *Rapid Commun. Mass Spectrom.*, **10**, 871–877.

Komori, Y., Shima, H., Fujino, T., *et al.* (2009) *J. Phys. Chem. C*, **114**, 1593–1600.

Law, K. P. (2010) *Int. J. Mass Spectrom.*, **290**, 72–84.

Law, K. P. and Larkin, J. R. (2011) *Anal. Bioanal. Chem.*, **399**, 2597–2622.

Liu, Z., Lu, L, Song, F, and Liu, S (2010) China Patent, 2009-10067479101644694.

Lopez-Legentil, S., Dieckmann, R., Bontemps-Subielos, N., *et al.* (2005) *Biochem. Syst. Ecol.*, **33**, 1107–1119.

Lu, L., Yue, H., Song, F., *et al.* (2010) *Chem. Res. Chin. Univ.*, **26**, 11–16.

McCombie, G. and Knochenmuss, R. (2004) *Anal. Chem.*, **76**, 4990–4997.

McLean, J. A., Stumpo, K. A. and Russell, D. H. (2005) *J. Am. Chem. Soc.*, **127**, 5304–5305.

Posthumus, M. A., Kistemaker, P. G., Meuzelaar, H. L. C., *et al.* (1978) *Anal. Chem.*, **50**, 985–991.

Ropivia, J., Derbré, S., Rouger, C., *et al.* (2010) *Molecules*, **15**, 6476–6484.

Rostovtsev, V., Green, L., Fokin, V., *et al.* (2002) *Angew. Chem. Int. Ed.*, **41**, 2596–2599.

Sanguinet, L., Aleveque, O., Blanchard, P., *et al.* (2006) *J. Mass Spectrom.*, **41**, 830–833.

Schinkovitz, A., Kenfack, G. T., Levillain, E., *et al.* (2011) *J. Mass Spectrom.*, **46**, 884–890.

Schinkovitz, A., Kenfack, G. T., Seraphin, D., *et al.* (2012) *Anal. Bioanal. Chem.*, DOI: 10.1007/s00216-012-5958-y.

Schuerenberg, M., Dreisewerd, K. and Hillenkamp, F. (1999) *Anal. Chem.*, **71**, 221–229.

Strupat, K., Karas, M. and Hillenkamp, F. (1991) *Int. J. Mass Spectrom. Ion Process.*, **111**, 89–102.

Sun, W., Liu, S., Liu, Z., *et al.* (1998) *Rapid Commun. Mass Spectrom.*, **12**, 821–824.

Suzuki, T., Midonoya, H. and Shioi, Y. (2009) *Anal. Biochem.*, **390**, 57–62.

Tanaka, K., Waki, H., Ido, Y., *et al.* (1988) *Rapid Commun. Mass Spectrom.*, **20**, 151–153.

Vaidyanathan, S., Gaskell, S. and Goodacre, R. (2006) *Rapid Commun. Mass Spectrom.*, **20**, 1192–1198.

Wu, W., Liang, Z., Zhao, Z., *et al.* (2007a) *J. Mass Spectrom.*, **42**, 58–69.

Wu, W., Qiao, C., Liang, Z., *et al.* (2007b) *J. Pharm. Biomed. Anal.*, **45**, 430–436.

Zenobi, R. and Knochenmuss, R. (1998) *Mass Spectrom. Rev.*, **17**, 337–336.

Part Three

Strategies for Selective Classes of Compounds

Analysis of Plant Oligo- and Polysaccharides

Wolfgang Blaschek

Department of Pharmaceutical Biology, University of Kiel, Kiel, Germany

1 INTRODUCTION

Carbohydrates represent the major organic substance on earth. They are mainly built in the process of photosynthesis by cyanobacteria and aquatic and land plants. Carbohydrates are structural elements of nucleic acids (DNA and RNA), some coenzymes (e.g., NAD(P)H and ATP), glycoproteins, and many secondary products (e.g., flavonoid glycosides, anthraquinone glycosides, and cardiac glycosides). Most plants use carbohydrates, especially the disaccharide sucrose to transport energy (in the phloem). Carbohydrates can bear information in molecular or cellular recognition. The bulk of carbohydrates in plants, however, are storage polysaccharides (e.g., starch and fructans) or cell wall polysaccharides (e.g., pectins, hemicelluloses, and cellulose) (Nishinari *et al.*, 2007; Fry, 2011).

Monosaccharides are the basic structural unit of oligo- and polysaccharides. The most common monosaccharides occurring in oligo- and polysaccharides are arabinose and xylose (pentoses); glucose, galactose, mannose, and fructose (hexoses); rhamnose and fucose (6-deoxy hexoses); and galacturonic acid and glucuronic acid (hexuronic acids).

Disaccharides are carbohydrates in which two monosaccharide units are joined by a glycosidic linkage.

Oligosaccharides by definition contain up to 10 monosaccharide residues joined by glycosidic linkages. According to the number of monosaccharide residues, they are trisaccharides, tetrasaccharides, pentasaccharides, and so on and then have a degree of polymerization (DP) of 3, 4, 5, and so on, respectively. Oligosaccharides may be linear, branched, or even cyclic (cyclodextrins). For oligosaccharides, a defined structure can be given.

Some examples for common plant di- and oligosaccharides are given in Table 1.

Polysaccharides are composed of a higher number of monosaccharide units compared to oligosaccharides. Often, they contain some hundred up to some thousand monosaccharide residues (high DP values) and accordingly are of high molecular weight (some thousand up to >1000 kDa). All monosaccharide residues again are joined via glycosidic linkages. Polysaccharides may be linear or branched macromolecules. Polysaccharides are also named glycans. If they are composed of only one type of monosaccharide, they are called *homopolymers* or *homoglycans*. The constituent monosaccharide is used in nomenclature followed by the suffix "an": if it is glucose only the polysaccharide is a glucan (e.g., amylose and cellulose); accordingly agar composed of galactose and 3,6-anhydrogalactose residues only is a galactan. Sometimes, minor constituent monosaccharides are ignored in nomenclature: for example, inulins or phleins that are mainly

Table 1 Some important plant oligosaccharides.

Oligosaccharide	Structure
Sucrose	β-D-Fru$f$-(2↔1)-α-D-Glc$p$
Maltose	α-D-Glc$p$-(1→4)-D-Glc
Maltotriose	α-D-Glc$p$-(1→4)-α-D-Glc$p$-(1→4)-D-Glc
Isomaltose	α-D-Glc$p$-(1→6)-D-Glc
Maltulose	α-D-Glc$p$-(1→4)-D-Fru
Laminaribiose	β-D-Glc$p$-(1→3)-D-Glc
Laminaritriose	β-D-Glc$p$-(1→3)-β-D-Glc$p$-(1→3)-D-Glc
Cellobiose	β-D-Glc$p$-(1→4)-D-Glc
Gentiobiose	β-D-Glc$p$-(1→6)-D-Glc
Gentianose	β-D-Glc$p$-(1→6)-α-D-Glc$p$-(1↔2)-β-D-Fru$f$
α,α-Trehalose	α-D-Glc$p$-(1↔1)-α-D-Glc$p$
β,β-Isotrehalose	β-D-Glc$p$-(1↔1)-β-D-Glc$p$
Kestose	β-D-Fru$f$-(2→6)-β-D-Fru$f$-(2↔1)-α-D-Glc$p$
Isokestose	β-D-Fru$f$-(2↔1)-β-D-Fru$f$-(2↔1)-α-D-Glc$p$
Neokestose	β-D-Fru$f$-(2↔6)-α-D-Glc$p$-(1↔2)-β-D-Fru$f$
Planteobiose	α-D-Gal$p$-(1→6)-D-Fru
Planteose	α-D-Gal$p$-(1→6)-β-D-Fru$f$-(2↔1)-α-D-Glc$p$
Raffinose	α-D-Gal$p$-(1→6)-α-D-Glc$p$-(1↔2)-β-D-Fru$f$
Stachyose	α-D-Gal$p$-(1→6)-α-D-Gal$p$-(1→6)-α-D-Glc$p$-(1↔2)-β-D-Fru$f$
Verbascose	α-D-Gal$p$-(1→6)-α-D-Gal$p$-(1→6)-α-D-Gal$p$-(1→6)-α-D-Glc$p$-(1↔2)-β-D-Fru$f$
Umbelliferose	α-D-Gal$p$-(1→2)-α-D-Gal$p$-(1↔2)-β-D-Fru$f$
Rutinose	α-L-Rha$p$-(1→6)-D-Glc
Sophorose	β-D-Glc$p$-(1→2)-D-Glc

composed of fructose units, but at one end contain one glucose residue, are called *fructans*. If the polysaccharide is composed of two or more different types of monosaccharides, it is called a *heteropolymer* or *heteroglycan*. Then, in nomenclature, the main constituent monosaccharide (often the backbone) is at the end of the word, whereas in the prefix, the additional (often side chains) monosaccharide unit is used: for example, galactomannans are composed of a backbone of mannose units substituted with galactose residues. Polysaccharides often are polydisperse in structure and size. In contrast to nucleic acids or proteins, they are not synthesized via a template-dependent process but by a mixture of (sequentially acting) specific enzymes. Therefore, often, the determination of a precisely defined structure of polysaccharides is impossible. Some examples for common plant polysaccharides are given in Table 2.

Different analytical methods complementing each other have to be used for the determination of the structure of an unknown oligo- or polysaccharide. These methods should give information on

- the constituent monosaccharides (e.g., glucose, arabinose, rhamnose, and galacturonic acid), their ring size (pyranose or furanose), their configuration (D or L), and substituents (e.g., sulfate or methyl or acetyl groups);
- the linkage position between monosaccharide residues, linear or branched molecules, and the anomeric configuration of glycosidic linkages (α or β);
- the sequence of monosaccharide residues, repetitive sequences, and structural heterogeneity;
- the molecular size (DP, degree of polymerization), the molecular size distribution, and heterogeneity in molecular weight;
- conformation of the whole or parts of the molecule (e.g., crystallinity).

The major analytical methods used to obtain the above-mentioned information are various colorimetric assays, complete and partial hydrolysis, treatment with specific glycosidases (enzymatic cleavage), TLC (thin layer chromatography), HPTLC (high performance thin layer chromatography), HPLC (high performance liquid chromatography), HPAEC (high performance anion exchange chromatography), SEC/MALLS (size-exclusion chromatography/multiangle laser light scattering), derivatization, methylation analysis, GC/MS (gas chromatography/mass spectrometry), NMR (nuclear magnetic resonance) spectroscopy, and some others.

2 ISOLATION OF POLYSACCHARIDES

Polysaccharides in plant tissues may be located intracellularly as storage polysaccharides, part of the cell wall as structural elements (some cell wall polysaccharides can also be storage polysaccharides) or excreted to the surrounding or to culture media as extracellular polysaccharides. Most oligosaccharides are water-soluble intracellular molecules. Some polysaccharides are water-soluble too, but others (especially many cell wall polysaccharides) must be solubilized by specific treatment. Other water-soluble material of low or high molecular weight must be removed by various methods.

2.1 General Conditions of Extraction

Polysaccharides should be extracted under conditions avoiding degradation processes. Buffers

Table 2 Some important plant polysaccharides.

Plant Polysaccharides		
Polysaccharide	Main structural features	Function as storage (SP) or cell wall (CW) polysaccharide and further details
Land plants		
Amylose	Linear glucan of α-(1→4) linked D-glucopyranose residues	SP, minor component (approximately 20%) of starch granules
Amylopectin	Branched glucan of α−(1→4)/α−(1→6) linked D-glucopyranose residues	SP, main component (approximately 80%) of starch granules, branching every 20–30 glucose residues
Inulin	Linear fructan of β-(2→1) linked D-fructofuranose residues	SP, for example, in Asteraceae
Levan (Phlein)	Linear fructan of β-(2→6) linked D-fructofuranose residues	SP, for example, in Poaceae
Galacturonan	Linear chain of α-(1→4)-linked D-galacturonic acid residues	CW, structural element of pectin
Rhamno-galacturonan I	Back bone of altering α-(1→4)-linked D-galacturonic acid and α-(1→2)-linked L-rhamnose residues	CW, structural element of pectin, many rhamnose residues branched at C4 with oligosaccharides mainly consisting of D-galactose and L-arabinose (arabinans, galactans, and arabinogalactans)
Arabinogalactan	β-(1→3)- and β-(1→6)-linked β-D-galactose and α-(1→3)-linked L-arabinose residues	As natural gum (gum Arabic) of *Acacia* sp., as main component of arabinogalactan protein (AGP) in cell walls, as sulfated polymers in green seaweeds
Rhamno-galacturonan II	Attached to a backbone of 8–10 galacturonic acids, up to 12 different sugars connected via many different linkages	CW, structural element of pectin, contains rare sugars: 2-O-methyl fucose, 2-O-methyl xylose, apiose, 3-C-carboxy-5-deoxy-L-xylose (aceric acid), 3-deoxy-D-lyxo-2-heptulosaric acid (DHA), and 3-deoxy-D-manno-2-octulosonic acid (KDO)
Xylan	β-(1→4)-linked D-xylopyranose residues, some branching via O-2 or O-3	CW, associated with cellulose, can replace cellulose in some green and red algae, mainly monomeric side chains with D-xylose, L-arabinose, D-glucuronic acid, or 4-O-methyl-D-glucuronic acid
Xyloglucan	β-(1→4)-linked glucose, highly branched with oligosaccharide residues	CW, short side chains with xylose, galactose, glucose, fucose, and arabinose
Glucomannan	β-(1→4)-linked D-mannose and β-(1→4)-linked D-glucose residues in a ratio of 1.5 : 2 : 1	CW and SP, degree of branching < 10% with single β-(1→6)-D-glucose residues, in galactoglucomannans branching with α-(1→6)-D-galactose residues
β-(1→3)(1→4)-Glucan	linear glucans of β-(1→4) and β-(1→3)-linked D-glucose residues	CW and SP, in lichens (Lichenan) and cereals, varying ratio of 1,3 : 1,4-linkages
Cellulose	linear glucan of β-(1→4)-linked D-glucose residues	CW, many hydrogen bonds within and between chains, formation of microfibrils and fibrils, DP of some hundred up to >10,000, and high crystallinity
Marine plants (Algae and Seaweeds)		
Laminaran (Laminarin)	Linear glucan with β-(1→3) and β-(1→6) linked D-glucopyranose residues	SP of brown seaweeds (Phaeophyceae), ratio of 1,3 : 1,6-linkages = 3 : 1
Alginates	Linear blocks of β-(1-4)-linked β-D-mannuronic acid (M) and α-L-guluronic acid (G)	CW of brown seaweeds (Phaeophyceae), blocks of G-residues (G-blocks) or of M-residues (M-blocks) or of alternating M and G-residues (MG-blocks)
Agar	The disaccharide 4-O-β-D-galacto-pyranosyl-3,6-anhydro-L-galactose (agarobiose) as repeating unit in a linear galactan	CW of red seaweeds (Rhodophyta), agarose as main component besides minor amounts of agaropectin
Carrageenan	Linear galactans with differentially sulfated β-(1→3) and α-(1→4)-linked D-galactopyranose residues	CW of red seaweeds (Rhodophyta), three main varieties differing in the degree of sulfation (DS): Kappa-carrageenan with one, Iota-carrageenan with two and lambda carrageenan with three sulfates per disaccharide unit

at pH 5–7.5 can be used for homogenization of plant tissues, as most polysaccharide degrading enzymes eventually present have pH optima of 4–5. One possibility would be to use HEPES at pH 7.5 [4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid], a buffer-stabilizing polysaccharides and not chelating with calcium ions that can be important in the isolation of pectins. Reductive agents such as dithiothreitol or mercaptoethanol (e.g., in concentrations of 10 mM) inhibit the oxidation and covalent binding of phenolic compounds. Normally, high temperature for a long time in the extraction process should be avoided. Short heating may be helpful to increase solubility, to inactivate glycanases, and to precipitate proteins, which then can be removed by centrifugation. Homogenization of plant material can be performed by milling-dried (best freeze-dried) tissue or by direct homogenization in buffers using a mortar (poss. with sea sand), an Ultra-Turrax, an Omni Mixer, ball mills, a French-Press, or by ultrasonication (Aspinall, 1982a; Harris, 1983; Selvendran and O'Neill, 1987; Selvendran, Stevens, and O'Neill, 1985; Fry, 1988).

2.2 Preparation of Cell Wall Material

For the extraction of cell wall polysaccharides, it can be of advantage first to isolate pure cell wall preparations. After a repeated extraction of homogenized plant material with buffer and water to remove any water-soluble compounds (e.g., sugars, amino acids, organic acids, inorganic salts, and proteins), an extraction with ice-cold ethanol (70%) results in the removal of any lipophilic compounds (e.g., fatty oils and phospholipids). Extraction steps are followed by centrifugation (e.g., 2000 g, 10 min) or filtration. The final so-called AIR (alcohol-insoluble residue) mainly contains the cell wall polysaccharides. Alternatively, after extraction of homogenized plant material with buffer and water, the removal of lipophilic compounds can also be performed by washing steps with $CHCl_3$/MeOH (1 : 1) and acetone (Aspinall, 1982a; Harris, 1983; Selvendran and O'Neill, 1987; Selvendran, Stevens, and O'Neill, 1985; Fry, 1988).

2.3 Removal of Starch

If the starting material contained starch, starch may be isolated as well and then has to be removed, for example, by repeated treatment with amylase. The used amylases must be free of other glycanase activities. For long and repeated amylase treatment, a contamination with microorganisms must be avoided. Another method to remove starch is its solubilization and extraction in dimethyl sulfoxide (DMSO) (90%), for example, for 12 h at 25°C. Finally, DMSO can be removed by washing steps with ethanol (70%). Besides starch, however, also some other polysaccharides (e.g., highly acetylated polysaccharides) can be soluble in DMSO at thus may be lost (Aspinall, 1982a; Selvendran, Stevens, and O'Neill, 1985; Fry, 1988).

2.4 Removal of Proteins

Contaminating proteins can be removed from isolated AIR by treatment with proteases, by treatment with SDS solutions, or by extraction with PAW (phenol/acetic acid/water, 2 : 1 : 1, wt/v/v) for 2 h at 25°C. Finally, PAW is removed with ethanol (70%). From solutions of mixtures of water-soluble proteins and polysaccharides, the proteins can be precipitated by heating (some minutes for >90°C) or by precipitation with trichloroacetic acid (10–20%) (Selvendran, Stevens, and O'Neill, 1985; Fry, 1988).

2.5 Extraction of Polysaccharides

If polysaccharides are not soluble in water per se, they must be brought into solution with specific extractants. DMSO, as mentioned earlier, solubilizes amylose and amylopectin (Bouveng and Lindberg, 1965; Selvendran et al., 1985). Chelating agents such as EDTA, EGTA, or CDTA (cyclohexane-trans-1,2-diamine tetra-acetate) used in concentrations of 10–50 mM bind calcium ions and thereby solubilize many pectins but not all of them (Jarvis, 1982; Fry, 1988).

Many water-insoluble polysaccharides can be solubilized in alkali (NaOH or KOH). Increasing alkali concentrations (0.01–4 M) can be used for sequential extraction of different polysaccharides. Specific polysaccharides may be extracted using specific alkali: for example, extracting wheat cell wall material with a saturated barium hydroxide solution selectively extracts water-insoluble arabinoxylans (Gruppen, Hamer, and Voragen, 1991). Substituents

bound via ester-linkage (e.g., O-acetyl groups), however, can be lost by alkali treatment. In order to minimize base-catalyzed hydrolysis, β-elimination, and Lobry de Bruyn–van Ekenstein transformations degrading the reducing end of polysaccharides (peeling) especially in the presence of oxygen, alkaline extraction can be performed under nitrogen atmosphere and/or together with reducing agents (e.g., 0.1–0.5% $NaBH_4$) (Thornber and Northcote, 1962; Whistler and Feather, 1965; Carpita, 1986; Fry, 1988).

Mixtures of MMNO (methylmorpholino-*N*-oxide) and water (e.g., 1:1, mol/mol) can solubilize water-insoluble neutral polysaccharides and even cellulose (at reduced water content), whereas acidic polysaccharides remain almost insoluble. MMNO/water as extractant normally is used at elevated temperatures (80–120°C). Solubilization often is accompanied by reduction in the molecular weight of the polysaccharides (Joseleau, Chambat, and Chumpitazi-Hermoza, 1980; Chanzy, Chumpitazi, and Peguy, 1982).

Cadoxene (1,2-diaminoethan/H_2O/CdO_2, 31:72:10, v/v/wt) is used to solubilize cellulose, and cadoxene/cellulose solutions can be used to determine the molecular weight of cellulose preparations (Brett, 1981; Hon and Srinivasan, 1983).

2.6 Removal of Low Molecular Weight Compounds

Extractants (e.g., their salts) and other low molecular weight compounds have to be removed after extraction which can be done by different methods such as precipitation, dialysis, or crossflow filtration (tangential flow filtration) (Binkley, 1965; Jones and Stoodley, 1965a; Fry, 1988). Membranes with different MWCO (molecular weight cutoff) are available. The specified MWCO values, however, normally are valid for globular molecules, whereas many polysaccharides are of linear shape. Dialysis can be problematic for removal of detergents. To avoid microbial contamination and growth, dialysis should be performed at low temperature. For inhibition of microbial growth, it can also be advantageous to add chlorobutol (0.05%), which after dialysis can easily be removed as being volatile.

Precipitation of polysaccharides with ethanol often is done at a final concentration of 80% at low temperature (4°C). By stepwise increasing the ethanol concentration, it can be possible to achieve sequential precipitation of different polysaccharides. Some polysaccharides (e.g., neutral arabinans) only can be precipitated by higher ethanol concentrations. After ethanol precipitation, the solubility of some polysaccharides can be reduced.

By GPC with Sephadex G-25 (or comparable material), low molecular weight molecules (<3000 Da) can easily be removed. Columns can be run without pumps as gravity is sufficient. As high and low molecular weight markers, dextran blue (0.2%), respectively $CoCl_2$ (0.5%) can be used. Some loss of material can occur by unspecific adsorption, which may become a problem when working with very small amounts of material. Some companies also offer various ready-to-use dialysis devices for large down to very small volumes.

2.7 Storage of Polysaccharides

Polysaccharides for a longer period best are stored as dried or freeze-dried material. For resolubilization, it can be necessary to use buffers at various pH, to add salts (e.g., 0.1% NaCl), to use alkali (e.g., 0.1 *M* NaOH + 0.1% $NaBH_4$), to elevate the temperature, or to use ultrasonication. For pectins, gums, or slime polysaccharides, it can be advantageous first to prepare a suspension in a little volume of ethanol, which then slowly is added to intensively stirred water thereby avoiding formation of clumps.

3 SEPARATION OF MIXTURES OF POLYSACCHARIDES

If specific solubilization or precipitation is not possible for the isolation of individual polysaccharides, separation methods due to differences in their molecular weight and/or their charge becomes necessary. For this purpose, separation is done by SEC and/or IEC. If both methods are used sequentially, it depends on the polysaccharide mixture which method best is performed first.

3.1 Differential Precipitation of Polysaccharides

Differential precipitation may be helpful in the isolation of specific polysaccharides (Fry, 1988).

Most polysaccharides precipitate in the presence of 60–80% of ethanol (final concentration) at low temperature (4°C) and then can be obtained by centrifugation. Xylans can be precipitated already at a final concentration of ethanol of 20%, xyloglucans at one of 50%, whereas arabinans often precipitate only at very high ethanol concentrations of 90% (Whistler and Sannella, 1965).

Acidic polysaccharides can form insoluble complexes with cationic detergents such as CTAB (cethyltrimethylammonium bromide) at a final concentration of 1% + 10 mM Na_2SO_4. After precipitation and centrifugation, the acidic polysaccharides can be sequentially solubilized by increasing amounts of Na_2SO_4. Finally, CTAB can be removed by precipitation of the polysaccharides with ethanol (Scott, 1965).

Ammonium sulfate precipitates proteins, but not most polysaccharides. An exception are β-(1-3)(1-4)-glucans, some acidic pectins, and some glucomannans and galactomannas, which may be precipitated in the presence of high concentrations of $(NH_4)_2SO_4$ (Fry, 1988).

$Ba(OH)_2$ precipitates mannans and some xylans (Meier, 1965).

Copper salts (Fehling's solution) can precipitate some acidic polysaccharides and some mannans and xylans (Jones and Stoodley, 1965b; Aspinall, Molloy, and Craig, 1969).

Calcium salts can be used to precipitate acidic pectins or at least the pectins form gels in the presence of the divalent cations because of the formation of intermolecular junction zones.

Precipitation and gel formation increase with decreasing degree of methylation in pectins (Pittet, 1965).

The solubility of some polysaccharides is pH dependent. If polysaccharides of plant cell wall material have been solubilized with NaOH after neutralization, some hemicelluloses precipitate, whereas others remain in solution (Whistler and Feather, 1965).

3.2 Ion-Exchange Chromatography (IEC)

Various materials for ion-exchange chromatography are known such as DEAE-, QAE-, CM-, SP-Sephadex or DEAE-, CM-Sepharose or Mono Q columns for FPLC as well as IEC columns for HPLC. Acidic polymers bind to the material, whereas neutral and positively charged polymers elute without binding. Bound carbohydrates are eluted by the anions of salts (e.g., Cl^- in NaCl or HCO_3^- in NH_4HCO_3) or bases (OH^- in NaOH). Using continuous gradient elution, weakly charged polysaccharides elute before strongly charged ones. As RI detection is impossible, detection often is performed by photometric carbohydrate tests (e.g., Anthron test). Results from continuous gradient elution can be used to establish a stepwise gradient elution by which larger amounts of polysaccharides can easily be separated (Cooper, 1981; Aspinall, 1982a; Zweig, Sherma, and Churms, 1982; Fry, 1988; Sherma and Shirley, 1991).

3.3 Size-Exclusion Chromatography (SEC) and Determination of Molecular Weight

For SEC (or GPC, gel permeation chromatography), a lot of different materials, prepacked columns, and further equipment (including pumps and detectors) are available. The choice of column material depends on the expected molecular weight of the polysaccharides to be separated. High molecular weight polysaccharides are eluting before low molecular weight ones. Globular molecules elute before linear ones. The basis for separation is the hydrodynamic volume of the polymers and not the absolute molecular weight. For good separations columns should be long with little diameter. Unspecific adsorption can be avoided using salts (e.g., NaCl) or a mixture of pyridine/acetic acid/water (1 : 1 : 23, pH 4.5), which can easily be removed as being volatile. Dextrans, pullulans, and some other polysaccharides are offered by various companies as standards of known molecular weight and their retention time is used to create a calibration curve. Detection can be done by photometric sugar tests (e.g., Anthron assay) or by RI detection (Churms, 1970, 1996; Whistler and Anisuzzaman, 1980; Cooper, 1981; Aspinall, 1982a; John *et al.*, 1982; Zweig, Sherma, and Churms, 1982; Yamasita, Mizuchi, and Kobota, 1982; Sherma and Shirley, 1991).

Using a multiangle laser light scattering (MALLS) detector in combination with a refractive index (RI) detector in SEC, the molecular weight of a polysaccharide can be determined. Light scattering depends on the molecular size and the concentration (measured by RI detection) of the polymer. The laser scattering signals and the concentration data are used to

calculate the molar mass weighted by number (M_n) and by weight (M_w), respectively. SEC/MALLS/RI can also be used for the determination of the molecular weight of oligosaccharides (Wyatt, 1993; Churms, 1996; Knobloch and Shaklee, 1997).

Oligo- and polysaccharides possess only one reducing end group per molecule. The number of reducing end groups is equivalent to the number of molecules. By determination of both, the total amount of monosaccharide units and the amount of reducing end groups, the (average) molecular weight of the molecules can be calculated. Reducing end groups can be determined in colorimetric assays with the help of specific enzymes such as dehydrogenases or by ^{1}H-NMR. Another method is the reduction of the oligo- or polysaccharide followed by acid hydrolysis, acetylation, and quantification of the resulting alditol acetates (reducing end units) and the aldose acetates (all other monosaccharide residues) by GC. The sensitivity of the method can be increased using NaB^3H_4 for reduction and labeling of the reducing end. Molecular weight determination via quantifying the reducing end groups works well for oligosaccharides and low molecular weight polysaccharides but becomes problematic with higher molecular weight polysaccharides (Fry, 1988; Nilsson et al., 1998; Courtin, Van den Beroeck, and Delcour, 2000).

The molecular weight of some polysaccharides can also be determined by measurement of the viscosity of polysaccharide solutions if correlation factors are known. The method has been used for cellulose dissolved in cadoxene (cadmiumethylenediamine) or after nitration of cellulose and dissolution of cellulose nitrate in acetone (Blaschek et al., 1982; Hon and Srinivasan, 1983).

4 CHARACTERIZATION OF OLIGO- AND POLYSACCHARIDES

4.1 Colorimetric Assays

In order to check the success of isolation steps, various photometric tests are used for carbohydrates (and contaminating other material). Concerning carbohydrates, they often are based on the determination of reducing end groups or the formation of chromophores by complexation with specific reagents. Calibration with appropriate standards (eventually

mixtures of monosaccharides) is necessary for accurate quantitative determinations.

The most common assays to determine carbohydrates (and contaminating proteins) are the following ones:

- *Anthron assay*. with anthron in H_2SO_4; measurement at 620 nm; hexoses blue, other sugars greenish; determination of free and polymer-bound hexoses (Dische, 1949, 1962).
- *Orcinol assay*. with orcinol in ethanol and $FeCl_3$ in HCl; measurement at 665 nm; pentoses green to blue; determination of free and polymer-bound pentoses (Dische, 1962).
- *Phenol – sulfuric acid assay*. with phenol and H_2SO_4; measurement at 485 nm; determination of all free and polymer-bound carbohydrates (soluble or insoluble polysaccharides) can also be performed on microplates (Dubois et al., 1956; Masuko et al., 2005).
- *Biphenylol assay*. with hydroxyl-biphenylol in NaOH and borax in H_2SO_4; measurement at 520 nm; uronic acids red to reddish blue; determination of free and polymer-bound uronic acids (Blumenkrantz and Asboe-Hansen, 1973).
- *Cystein/H_2SO_4 assay*. with cysteine–HCl in H_2O and H_2SO_4; measurement at 380, 396, and 427 nm; determination of free and polysaccharide-bound 6-deoxy-hexoses (Dische, 1962).
- *PAHBAH assay*. with *p*-hydroxy-benzoic acid hydrazide in HCl and NaOH; measurement at 410 nm; determination of reducing carbohydrates (Lever, 1972),
- *Updegraff assay*. with HOAc/H_2O/HNO_3 (8 : 2 : 1) in H_3PO_4; determination of cellulose by hydrolysis of all noncellulosic polysaccharides and their removal by centrifugation followed by H_2SO_4 hydrolysis of remaining cellulose and Anthron test for glucose quantification (Updegraff, 1969).
- *Coomassie assay*. with Coomassie Brillant Blue in H_3PO_4; measurement at 595 and 465 nm; determination of soluble proteins (Read and Northcote, 1981).
- *Lowry assay*. with $CuSO_4$/Na-K-tartrate in Na_2CO_3 and Folin/Ciocalteu's reagent; often after TCA precipitation of proteins and their resolubilization in NaOH; measurement at 750 nm; determination of soluble and precipitated proteins (Layne, 1957).

4.2   Hydrolysis of Polysaccharides

For the determination of the monosaccharide constituents of an oligo- or polysaccharide, hydrolytic cleavage of glycosidic linkages is necessary followed by chromatographic separation of the released monosaccharides. By other methods, fragments of polysaccharides can be obtained for further analysis.

4.2.1   Acid Hydrolysis

Hydrolysis with inorganic acids is the most common method for hydrolysis of polysaccharides. Acids are used at different concentrations in different combinations of time and temperature depending on the polysaccharide structure. The hydrolytic conditions should allow cleavage of all glycosidic linkages without decomposition of liberated monosaccharides in hot acids. Often, the chosen hydrolytic conditions are a compromise between both demands. Sometimes, it may become necessary to hydrolyze a given polysaccharide under different conditions and then to compare the results. Using monosaccharide mixtures similar to the composition of the polysaccharide, correction factors for the decomposition of individual monosaccharides under different hydrolytic conditions can be determined. TFA (trifluoroacetic acid) is a preferred acid because it is volatile and can easily be removed after hydrolysis.

The most common conditions for hydrolysis of polysaccharides (Table 3) containing mainly the below-mentioned monosaccharide components are (Adams, 1965; Aspinall, 1976; Fry, 1988; Biermann and McGinnis, 1989) as follows:

- very easily hydrolyzed are Api_f and KDOp; conditions: 0.1 M TFA at 50°C for 24 h.
- easily hydrolyzed is Ara_f; conditions: 0.1 M TFA at 100°C for 1 h.
- easily hydrolyzed are 6-deoxy-hexoses (e.g., Fuc_p and Rha_p); conditions: 1 M TFA at 100°C for 1 h.
- easily hydrolyzed are furanosidic monosaccharides (e.g., Frc_f in fructans); conditions can be 2% oxalic acid at 80°C for 30 min; oxalic acid then can be precipitated as calcium oxalate by addition of calcium acetate; other conditions may be 2 M TFA at 80°C for 1 h.
- standard hydrolytic conditions for many polysaccharides containing high amounts of Ara_p, Xyl_p, Man_p, Gal_p, Glc_p, $GlcNAc_p$, and $GalNAc_p$ are:

Table 3   General conditions for hydrolysis of polysaccharides mainly containing specific monosaccharide residues.

Hydrolysis of Polysaccharides with Acids

Furanoses > Pyranoses
Pentoses > Hexoses
α-Glycosidic > β-glycosidic linkage
1,4- > 1,3- > 1,2- > 1,6-linked monosaccharides

Polysaccharide composition	Conditions for hydrolysis
Api*f*, KDO*f*	0.1 M TFA, 50°C, 24 h
Ara*f*	0.1 M TFA, 100°C, 1 h
Fuc*p* and Rha*p*	1 M TFA, 100°C, 1 h
Fru*f*	2% oxalic acid, 80°C, 30 min
	2 M TFA, 80°C, 60 min
Glc*p*, Gal*p*, Man*p*, Ara*p*, and Xyl*p*	2 M TFA, 120°C, 1 h
	4% H_2SO_4, 120°C, 1 h
GalA*p* and GlcA*p*	2 M TFA, 120°C, 1 h
	70–90% formic acid, 100°C, 1 h
	+ 2 M TFA, 120°C, 1 h
GlcNAc*p* and GalNAc*p*	2 M TFA, 120°C, 1 h
	0,5–6 N HCl, 100°C, 1 h
GlcN*p*, GalN*p*	2 M TFA, 120°C, 1 h
	4 M HCl, 100°C, 9 h
Cellulose	72% H_2SO_4, 25°C, 1 h
	+ 4% H_2SO_4, 120°C, 1 h

2 M TFA at 120 for 1 h; no sufficient hydrolysis of cellulose or chitin because of their crystallinity.

- more acid stable are $GalA_p$ and $GlcA_p$; conditions as 2 M TFA at 120°C for 1 h result in incomplete hydrolysis of uronic acids, but more drastic conditions can cause decomposition (decarboxylation) of them; if uronic acids are linked to neutral sugars, the resulting aldobiuronic acids (e.g., GalA-Rha from rhamnogalacturonans) are also very stable to hydrolysis and the yield of the appropriate neutral sugar will also be lower; other conditions for hydrolysis of uronic acid containing polysaccharides can be 70–90% formic acid at 100°C for 4–24 h (uronic acids are relatively stable in formic acid and formyl esters are formed) followed by hydrolysis with 2 M TFA at 100°C for 1 h.
- more acid stable are $GlcNAc_p$ and $GalNAc_p$; conditions: 2 M TFA at 100°C for 1 h; as deacetylation of GlcNAc or GalNAc results in incomplete hydrolysis after reacetylation hydrolysis can be repeated; other conditions can be 0.5–6 N HCl at 100°C for 1 h.

- more acid stable also are GalN$_p$ and GlcN$_p$; conditions: 2 M TFA at 120°C for 1 h; other conditions can be 4 m HCl at 100°C for 9 h.
- very acid stable are (partially) crystalline polymers such as cellulose or chitin; for example, cellulose can be dissolved in 72% H_2SO_4 at 25°C for 1 h (swelling of microfibrils) and after dilution to 3–4% of H_2SO_4 hydrolysis is performed at 120°C for 1 h; by addition of Ba(OH)$_2$ and BaCO$_3$, thereafter H_2SO_4 can be precipitated and removed as BaSO$_4$.

4.2.2 Partial Hydrolysis of Polysaccharides

Using mild hydrolytic conditions, specific sugar residues can be removed from a polysaccharide without total hydrolysis of the polymer (e.g., terminal arabinose residues from an arabinogalactan resulting in the remaining galactan) or the polysaccharide is fragmented into oligosaccharides, which then can be analyzed separately. Thereby, in a kind of puzzle, the structure of the polysaccharide can be analyzed (Adams, 1965; Painter, 1965; Wolfrom and Franks, 1965; Aspinall, 1982b; Fry, 1988).

4.2.3 Acetolysis

Acetolysis can complement (partial) acid hydrolysis. In acid hydrolysis, 1,4-, 1,3-, or 1,2-glycosidic bonds are more easily hydrolyzed compared to 1,6-linkages, whereas, in acetolysis, 1,6-glycosidic bonds are the most labile ones and preferentially cleaved. Glycosidic linkages in α-configuration are more susceptible to acetolysis than those in β-configuration. Only little degradation takes place for 6-deoxy-sugars (e.g., Rha or Fuc). Acetylosis can be performed in acetic anhydride, acetic acid, and H_2SO_4 (10 : 10 : 1) at 25°C for 1–10 h. Resulting mono- and oligosaccharide acetates can be recovered by extraction with organic solvents and separated directly by chromatography or deacetylated by treatment with methanolic barium methoxide before further chromatographic or other analysis. (Aspinall, 1976, 1982b; Lindberg, Lönngren, and Svensson, 1975; Fry, 1988; Biermann and McGinnis, 1989).

4.2.4 Smith Degradation

The Smith degradation is a combination of periodate oxidation, reduction, and mild acid hydrolysis

of a polysaccharide. The aim is a structure-dependent fragmentation and analysis of the fragments in order to gain more knowledge on structural elements of the polymer. In a first step, the polysaccharide is oxidized with sodium periodate (e.g., 0.05 M sodium periodate at 20°C for 48 h). Vicinyl (two neighboring) hydroxyl groups in monosaccharide units are oxidized to two aldehyde groups accompanied by cleavage between the C-atoms bearing the hydroxyl groups. If there exist three adjacent hydroxyl groups, a double cleavage of the carbon chain results in the formation of two aldehyde groups and release of formic acid. Excess periodate is destroyed by addition of ethylene glycol. The aldehyde groups are then reduced to primary alcohol groups with sodium borohydride. By mild acid hydrolysis (e.g., 0.5 M TFA at 20°C for 15 h), the oxidized sugar residues can be cleaved without hydrolyzing glycosidic linkages. The optimum conditions for oxidation and hydrolysis can vary with the polysaccharide investigated. Depending on the individual monosaccharide residues and the type of linkage different fragments are released such as erythritol, glycerol, glycerinaldehyde, glycolaldehyde, or formic acid (Table 4). These products can be analyzed by HPLC or after derivatization by GC. The remaining structure of the polysaccharide, respectively, the produced oligosaccharides can also be analyzed in detail by methylation analysis. The method can give information concerning the type of linkage, the existence of pyranosidic or furanosidic rings, or the position of O-acetyl, O-methyl, or N-acetyl groups, which are protective groups against periodate oxidation. Sometimes, the backbone of a polysaccharide can be obtained if side chains are removed by Smith degradation; for example, backbone units consisting of 1,3-linked hexose residues are not degraded (Goldstein *et al*., 1965; Hay, Lewis, and Smith, 1965; Lindberg, Lönngren, and Svensson, 1975; Aspinall, 1976, 1982b; Biermann and McGinnis, 1989).

4.2.5 Enzymatic Hydrolysis

Various glycosidases can be used for specific cleavage of glycosidic linkages in oligo- and polysaccharides. The enzymes should be of high quality just attacking one type of linkage and free of traces of other glycosidases. Some of them are endoglycanases (e.g., α-amylase, EC 3.2.1.1; internal hydrolysis of α-(1→4)-glucans), others

Table 4 Products of Smith degradation using as example glucose residues with different glycosidic linkage in a polysaccharide.

Type of linkage in the polysaccharide	Smith degradation			
	After oxidation with $NaJO_4$ and reduction with $NaBH_4$		After mild acid hydrolysis	
	From	Product	From	Product
1-Glucose	C3	Formic acid	C1-C2	Glycolaldehyde
			C4-C5-C6	Glycerol
1,2-Glucose	–	–	C1-C2-C3	Glyceraldehyde
			C4-C5-C6	Glycerol
1,3-Glucose	–	–	–	–
1,4-Glucose	–	–	C1-C2	Glycolaldehyde
			C3-C4-C5-C6	Erythritol
1,6-Glucose	C3	Formic acid	C1-C2	Glycolaldehyde
			C4-C5-C6	Glycerol

are exoglycanases releasing disaccharides (e.g., β-amylase, EC 3.2.1.2; terminal release of maltose from α-(1→4)-glucans), or monosaccharides (e.g., glucoamylase, EC 3.2.1.3; terminal release of glucose from α-(1→4)-glucans). Some other examples are endo-β-xylanase (EC 3.2.1.8; internal hydrolysis of β-(1→4)-xylans), α-galactosidase (EC 3.2.1.22; terminal hydrolysis of α-linked galactose), pullulanase (EC 3.2.1.37; terminal hydrolysis of α-linked L-arabinose at reducing ends), and so on. They are offered by various companies. The enzymatic degradation products can be mixtures of mono-, oligo-, and polysaccharides that can be separated and characterized (after derivatization) by chromatographic methods in order to get additional information on the structure of the polysaccharide (Wood *et al.*, 1994; Henrissat and Bairoch, 1996; Grishutin *et al.*, 2004).

4.3 Separation of Mono- and Oligosaccharides

Mono- and oligosaccharides obtained by cleavage of glycosidic bonds can be analyzed by different methods.

4.3.1 Paper (PC) and Thin Layer Chromatography (TLC)

ptThe separation of mixtures of mono- and oligosaccharides can be achieved by paper chromatography (PC) or thin-layer chromatography (TLC), which have been the first methods to analyze carbohydrate mixtures (See **Thin-layer Chromatography, with Chemical and Biological Detection Methods**). Meanwhile, also HPTLC (high performance thin layer chromatography) plates are available from various suppliers. A lot of different methods concerning stationary and mobile phases, conditions, and detection reagents are described. The methods are relatively simple, quick, inexpensive, and sensitive and can be performed with numerous samples simultaneously. Exact quantification of individual carbohydrates, however, is a problem (Smith, 1960; Hais and Macek, 1963; Stahl, 1967; Wing, 1972; Ghebregzabher *et al.*, 1976; Zweig, Sherma, and Churms, 1982; Touchstone and Dobbins, 1983; Fry, 1988; Jork *et al.*, 1989; Sherman and Shirley, 1991).

4.3.2 HPLC of Mono- and Oligosaccharides

For the separation of mixtures of monosaccharides, often gas–liquid chromatography (GLC) analysis is preferred, whereas for that of mixtures of oligosaccharides, often HPLC is used. The most common HPLC columns for carbohydrates are packed with amino-bonded silica-based materials or metal-loaded polymeric substrates such as cross-linked styrene–divinylbenzene. Both can be used for various modes of separation such as reversed-phase, ion exchange, hydrophilic and hydrophobic interaction, or size-exclusion chromatography. UV detection normally is not used

as sugars must be detected at wavelengths below 200 nm causing high background interference. RI detection is possible although sensitivity is relatively low, whereas temperature sensitivity is high, and RI detection is impossible for gradient elution because of extreme sensitivity to changes in mobile-phase concentrations. Postcolumn derivatization of carbohydrates is possible for their UV–Vis detection but is laborious. In addition, various fluorophores can be used to enable detection and to improve detection sensitivity in the analysis of reducing mono- and oligosaccharides. Pulsed amperometric detection (PAD) is another convincing method in HPLC detection of carbohydrates on anion exchange resins with NaOH as eluent. In alkali, carbohydrates are weakly acidic and interact with strong anion exchange resins. In PAD, carbohydrates are electrochemically oxidized at the surface of a gold or platinum electrode and the produced and measured current is proportional to the concentration of the particular carbohydrate. The relative detector response changes with the DP of an oligosaccharide. Gradient elution in HPAEC/PAD (high performance anion exchange chromatography coupled with pulsed amperometric detection) is no problem. The method is often used for the analysis of di- and oligosaccharides, for example, to control partial hydrolysis of starch. Separation of malto-oligo-/polysaccharides up to a DP of 40 is possible (McGinnis and Fang, 1980; Verhaar and Kuster, 1981; Wells *et al*., 1982; Yamashita *et al*., 1982; Zweig, Sherma, and Churms, 1982; Blaschek, 1983; Rocklin and Pohl, 1983; Pecina *et al*., 1984; Hicks, Lim, and Haas, 1985; Rajakyla, 1986; Kobota, Yamashita, and Takasaki, 1987; Fry, 1988; Biermann and McGinnis, 1989; Koizumi, Okada, and Fukuda, 1991; Sherma and Shirley, 1991; De Ruiter *et al*., 1992; Corradini, 1994; Churms, 1996; Lee, 1996; Suortti, 1997; Zalyalieva *et al*., 1999; Okatch, 2003; Shilova and Bovin, 2003).

4.3.3 Reduction or Oxime Formation

As monosaccharides are not sufficiently volatile, GC analysis of mixtures of them obtained after hydrolysis of an oligo- or polysaccharide requires derivatization to increase volatility. In order to avoid formation of too much multiple peaks per monosaccharide often, first of all a derivatization at the anomeric C-atom is performed,

A widely used method is reduction of the carbonyl group of the monosaccharides with sodium borohydride in ammonium hydroxide or in DMSO resulting in the formation of alditols. Reduction of an aldose results in the formation of one single alditol, whereas reduction of ketoses yields two of them; for example, reduction of fructose produces nearly equal amounts of glucitol and mannitol.

Another often-used method is the formation of oximes with hydroxylamine in pyridine producing the corresponding *syn*(E) and *anti*(Z) forms from each reducing sugar. Another possibility is the formation of *O*-methyl-oximes using methyl-hydroxylamine also resulting in the two isomeric forms (Aspinall, 1976, 1982b; Bradbyury *et al*., 1981; Klok *et al*., 1981; Zweig, Sherma, and Churms, 1982; Blakeney, Harris, and Stone, 1983; Guerrant and Moss, 1984; Harris, Bacic, and Clarke, 1985; Biermann and McGinnis, 1989; Sherma and Shirley, 1991; Ruiz-Matute *et al*., 2011).

4.3.4 Acetylation or Silylation

The production of volatile derivatives mainly is performed by acetylation or silylation.

Acetylation can be done with acetic anhydride in pyridine or in DMSO, and 1-methylimidazole can be used as catalyst that renders removal of borate unnecessary.

For silylation in pyridine or in DMSO, various silylation reagents and mixtures of them can be used such as HMDS (hexamethyldisilizane), TMCS (trimethylchlorosilane), BSA [*N*,*O*-bis(trimethylsilyl)acetamide], TMSI (*N*-trimethylsilylimidazole), or BSTFA [*N*,*O*-bis(trimethylsilyl)trifluoroacetamide]. The reaction mixtures should be free of water; in the presence of traces of water, an excess of silylation reagents should be used. Owing to the high volatility of trimethylsilyl oximes (TMSO), also oligosaccharides normally up to a DP of 5 can be analyzed as TMSO derivatives by GC. TMS derivatives of uronic acids are problematic; to avoid multiple peaks, they can be reduced to aldonic acids, converted into aldono-1,4-lacotones, and then silylated.

The volatile sugar derivatives are separated by capillary gas chromatography with an FID. For quantification, response factors have to be used, which can be determined with the help of corresponding

sugar mixtures. Often, inositol (or another carbohydrate) is used as internal standard for retention times of individual peaks and for quantification; internal standards can already be added before hydrolysis (if of acceptable stability). In case of very similar retention times of two compounds, their identification should be confirmed by GC/MS (Brobst, 1972; Loewus and Shah, 1972; Sloneker, 1972; Bradbury, Halliday, and Medcalf, 1981; Klok *et al*., 1981; Oshima, Yoshikawa, and Kumanotani, 1981; Zweig, Sherma, and Churms, 1982; Blakeney, Harris, Stone, 1983; Garcia-Raso *et al*., 1987; Guerrant and Moss, 1987; Biermann and McGinnis, 1989; Sherma and Shirley, 1991; Doco, O'Neill, and Pellerin, 2001; Ruiz-Matute *et al*., 2011).

4.4 Methylation Analysis

Methylation analysis has been and still is the most powerful method in structural analysis of oligo- and polysaccharides.

4.4.1 Formation of PMAA (Partially Methylated Alditol Acetates)

In a first step in methylation analysis, all free hydroxyl groups of an oligo- or polysaccharide by methylation are converted into methoxyl groups. Thereafter, acidic hydrolysis of the poly-methyl ethers cleaves glycosidic linkages without attacking methyl-ether bonds, resulting in monomers with hydroxyl groups at former linkage positions. The monosaccharides (pyranosidic or furanosidic rings) by reduction then are converted into alditols with an additional hydroxyl group originating from ring opening. Acetylation of all hydroxyl groups finally results in volatile partially methylated alditol acetates (PMAA), which can be identified and quantified by GLC combined with mass spectrometry (GC-MS). The substitution pattern of the O-acetyl-groups in the individual PMAA reflects the linkage positions and ring sizes of the corresponding sugar in the original polymer. No information, however, on the anomeric configuration of the glycosidic linkages (α or β) and the sequence of specific monosaccharide residues can be given.

Mainly dry DMSO is used as solvent for the oligo- or polysaccharides. Solubility may be increased by ultrasonication (at elevated temperature); partial insolubility results in undermethylation. The poly-anion is formed in the presence of a strong base such as dimsyl sodium (sodium methylsulfinylmethanide), dimsyl potassium, or dry powdered sodium hydroxide. Dimsyl sodium or dimsyl potassium is prepared by addition of sodium or potassium hydride to DMSO; as the DMSO anion reacts rapidly with water, carbon dioxide, and oxygen, the procedure is performed under nitrogen. Methylation of the alkoxide groups then is done with methyl iodide. If the solubility of the polysaccharide in DMSO is not satisfactory, a (repeated) premethylation or a preceding acetylation of the polysaccharide can be helpful. Alkali-labile groups such as acetyl groups are lost in the course of methylation; this is also the case for genuine *O*-acetylated polysaccharides. Methylation is stopped by addition of water and the permethylated polysaccharide is extracted, for example, with methylene chloride. Methylated polysaccharides also can be purified with the help of small C18-columns (Sep-Pak) and elution with acetonitrile. Subsequent hydrolysis can be done with hydrochloric or sulfuric acid, but often is carried out with TFA (e.g., 2 *M* TFA at 120°C for 1–2 h) as it can easily be removed by evaporation under a stream of nitrogen. The monosaccharides are then reduced to their alditols with sodium borohydride ($NaBH_4$) or sodium borodeuteride ($NaBD_4$); the latter introduces deuterium labeling at C1 and thus allows to distinguish it from the C5- or C6-end of the alditol. Acetylation with acetic anhydride in pyridine then finally results in PMAA that are analyzed by GC/MS. Retention times, mass spectra, and peak areas are used to identify and quantify monosaccharides and their linkage pattern. An example for the methylation analysis of a xyloglucan oligosaccharide and the resulting PMAA is provided in Table 5 (Hakamori, 1964; Björndal, Lindberg, and Svensson, 1967; Björndal *et al*., 1970; Sweet, Shapiro, and Albersheim, 1975; Aspinall, 1976, 1982b; Valent *et al*., 1980; Phillips and Fraser, 1981; Lomax and Conchie, 1982; McNeil *et al*., 1982; Zweig, Sherma, and Churms, 1982; Waeghe *et al*., 1980; Ciucanu and Kerek, 1984; Harris *et al*., 1984; Harris, Bacic, and Clarke, 1985; Blakeney and Stone, 1985; Sweeley and Nunez, 1985; Biermann and McGinnis, 1989; Sherma and Shirley, 1991; Jay, 1996).

If the polysaccharide contains uronic acids, which show unusual stability to acid hydrolysis, before

Table 5 PMAA obtained by methylation analysis of (a section of) a xyloglucan.

Methylation analysis of a xyloglucan oligosaccharide

```
    E          D          C          B          A
  4 Glc (1→4) Glc (1→4) Glc (1→4) Glc (1→4) Glc 1
               6                     6
               ↓                     ↓
               1                     1
               Xyl    F              Xyl   G
                                     2
                                     ↓
                                     1
                                     Gal    H
```

PMAA	OAc	OMe	Type of linkage
A	1,4,5	2,3,6	Reducing end Glc*p*
B	1,4,5,6	2,3	1,4,6-Glc*p*
C	1,4,5	2,3,6	1,4-Glc*p*
D	1,4,5,6	2,3	1,4,6-Glc*p*
E	1,5	2,3,4,6	Terminal Glc*p*
F	1,5	2,3,4	Terminal Xyl*p*
G	1,2,5	3,4	1,2-or 1,4-Xyl*p*
H	1,5	2,3,4,6	Terminal Gal*p*

methylation analysis, a chemical reduction of carboxyl groups is recommended. Usually, carbodiimide activation of carboxyl groups is performed with *N*-cyclohexyl-*N*′-(2-morpholinoethyl)carbodiimide metho-*p*-toluenesulfonate (CMC) at pH 4.75 followed by reduction with sodium borohydride or sodium borodeuteride at pH 7.0. The polysaccharide thereafter containing only neutral sugars is then subjected to normal methylation analysis. Use of NaBD$_4$ enables to distinguish between genuine uronic acids (deuterized) and genuine neutral sugars in PMAA analysis (Perry and Hulyalkar, 1965; Taylor and Conrad, 1972; Zweig, Sherma, and Churms, 1982; Biermann and McGinnis, 1989; Sherma and Churms, 1991).

4.4.2 GC/MS Analysis of PMAA

Capillary GLC coupled with a mass spectrometer (EI-MS) is used for separation and quantification of PMAA. Fragmentation of PMAA follows certain rules:

- Molecular ions are rarely detected.
- Loss of an acetylium ion (CH$_3$CO$^+$) with *m/z* 43 predominates and is the base peak.
- Primary fragments result from cleavage of the alditol backbone.
- Fragmentation between two methoxy-bearing carbon atoms is favored.
- Fragmentation between a methoxy- and an acetoxy-bearing carbon atom is less frequently.
- The charge resides on the fragment with a methoxy-bearing carbon atom besides cleavage.
- Fragmentation between two acetoxy-bearing carbon atoms is rare.
- Secondary fragments originate from primary fragments by loss of methanol (*m/z* 32), acetic acid (*m/z* 60), ketene (*m/z* 42), and formaldehyde (*m/z* 30)
- Loss of methanol or acetic acid from the primary fragment is primarily from the β-carbon to the charge of the carbonium ion.
- A PMAA with deuterium labeling at C1 produces C1-containing fragments with even *m/z*, whereas

without deuterium labeling, these fragments have odd mass-to-charge ratio.

In Figure 1, some examples for MS fragmentation patterns and mass spectra of PMAA of differently linked galactose residues are shown (Kochetkov and Chizhov, 1966; Björndal, Lindberg, and Svensson, 1967; Björndal *et al.*, 1970; Aspinall, 1976, 1982b; Jansson *et al.*, 1976; Lindberg and Lönngren, 1978; DeJongh, 1980; Phillips and Fraser, 1981; Lomax and Conchie, 1982; McNeil *et al.*, 1982; Zweig, Sherma, and Churms, 1982; Sweeley and Nunez, 1985; Biermann and McGinnis, 1989).

Stereoisomers of PMAA (e.g., 1.4-Glc and 1.4 Gal) have very similar mass spectra but different retention times in GC. Some pairs of PMAA (e.g., 3- and 4-OMe-hexoses, 1.2- and 1.4-pentoses, and 1.2- and 2.6-fructose) result in the same mass spectra. By introducing deuterium labeling at C1 (reduction with $NaBD_4$ in methylation analysis), asymmetry is introduced and they can be distinguished (Biermann and McGinnis, 1989). Some examples are shown in Figures 2 and 3.

If a polysaccharide contains individual monosaccharide residues with methoxy groups, its hydrolysis, reduction to alditols, and acetylation results in specific PMAA with single OMe groups only (Biermann and McGinnis, 1989). For example, a genuine terminal 3-OMe-fucose will produce the PMAA 1,2,3,5-tetra-O-acetyl-3-O-methyl-fucitol.

The localization of genuine O-acetyl groups (because of their alkaline hydrolysis in normal methylation analysis) requires a different strategy. By treatment of the polysaccharide with methyl vinyl ether, all free hydroxyl groups can be protected. Thereafter, in methylation analysis by alkaline hydrolysis, the O-Ac groups are lost and replaced by O-Me groups, whereas later on by hydrolysis and acetylation, the protecting groups are replaced by O-Ac groups. For example, a genuine 1,4-linked 3-OAc-mannose will produce the PMAA 1,2,4,5,6-penta-O-acetyl-3-O-methyl-mannitol (De Belder and Normann, 1969; Lomax, Gordon, and Chesson, 1983).

4.5 Mass Spectrometry

For MS analysis of polysaccharides, they first must be degraded by partial acid hydrolysis or treatment with specific enzymes (glycosidases) into oligosaccharides that then are separated by chromatographic methods such as HPLC. Various MS techniques can be used to analyze individual oligosaccharides and to get information on linkages, branching, monosaccharide sequence, and substitutions (methyl, acetyl, and sulfate groups) from characteristic fragment patterns. Methods such as MS/MS, FAB (fast atom bombardment), MALDI (matrix-assisted laser desorption ionization), MALDI-TOF (MALDI coupled to a time-of-flight analyzer), or ESI (electrospray ionization) are used. By cleavage of glycosidic linkages (and by cross-ring fragmentation), ions are generated. Depending on the ionization technique and the structure of the oligosaccharide specific series of fragments can be formed and used to give information on the sequence of sugar units. Rules for the frequency of different linkages and resulting fragments have been established. Derivatization (permethylation) of oligosaccharides is also applied to help facilitated identification of branching (Dell *et al.*, 1994; Viseux, de Hoffmann, and Domon, 1998; Chai, Piskarev, and Wawson, 2001; Kabel, Schols, and Voragen, 2001; Tang, Mechref, and Novotny, 2005; Bauer *et al.*, 2006; Böcker, Kehr, and Rasche, 2011; Bauer, 2012).

4.6 Nuclear Magnetic Resonance Spectroscopy

By NMR, additional valuable information on the structure of an oligo- or polysaccharide can be obtained. NMR is particularly useful for the determination of the anomeric configuration of glycosidic linkages (α or β), the ring size (pyranose or furanose), and detection of substituents (O-methyl, O-acetyl, N-acetyl, methyl, and carboxyl groups). Both ^{1}H and ^{13}C-NMR can be applied (See **NMR of Large Molecules**). ^{1}H-NMR usually is run in D_2O in order to eliminate exchangeable hydroxyl protons and the excess of H_2O protons. Owing to the low natural abundance of ^{13}C (1,1%), for ^{13}C-NMR, higher amounts of sample material (10–30 mg) have to be used compared to ^{1}H-NMR (1–2 mg). On the other hand, ^{13}C-NMR spectra normally are less complex, although both ^{1}H-NMR- and ^{13}C-NMR spectra especially of polysaccharides consisting of various sugars in different types of linkage can become extremely complex. As the anomeric carbon is bonded to two oxygen atoms, the C1 and its protons show downfield shifts (to higher parts per million) and can be
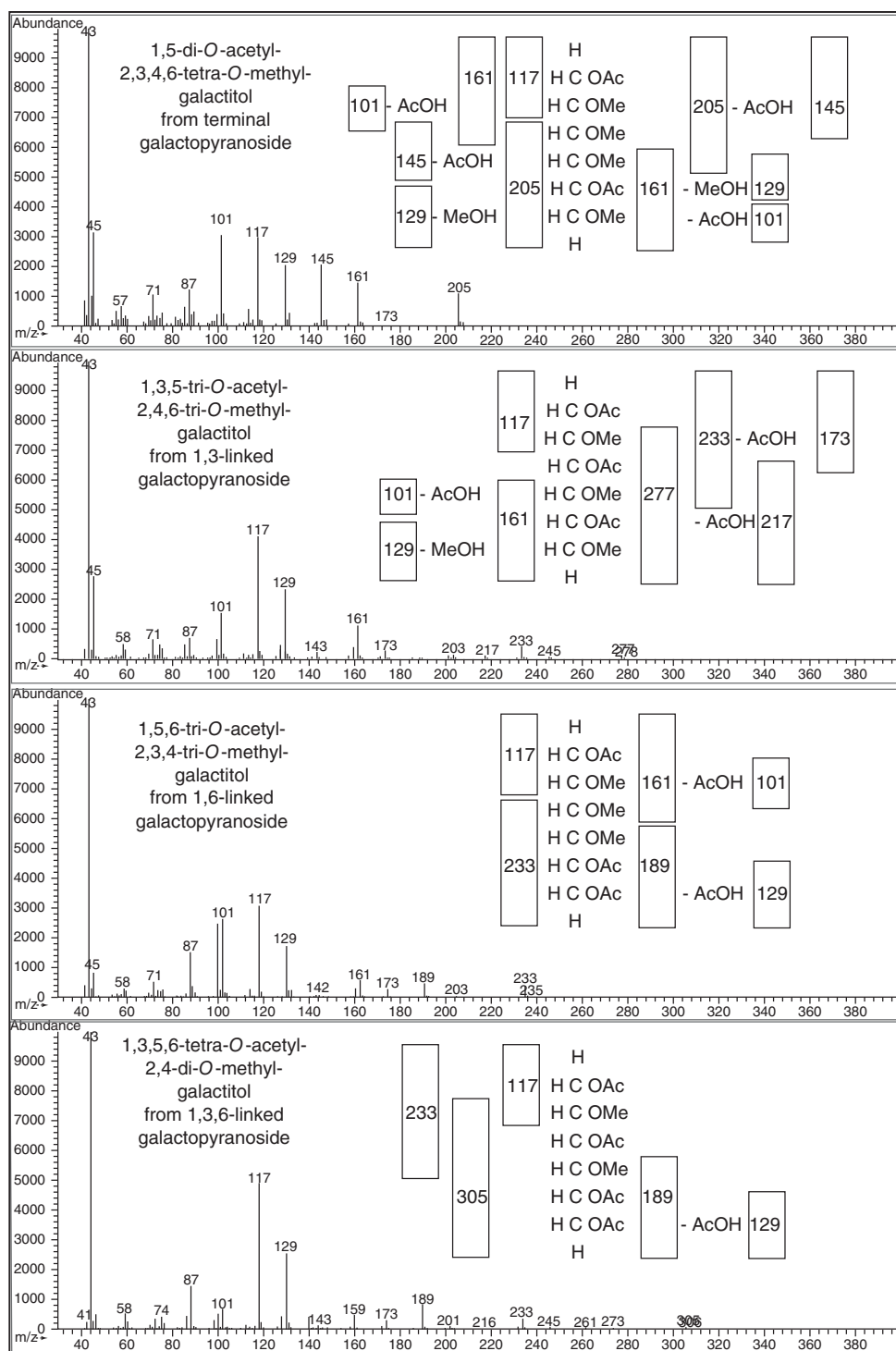
Figure 1 Mass spectra of PMAA of differently linked galactose residues in an arabinogalactan type 2 and main features of fragmentation patterns.

```
            H                     H
        D / H C OAc         D / H C OAc
190 / 189   H C OAc             H C OMe   118 / 117
            H C OMe         ------------------------
------------------------------    H C OMe
            H C OMe             H C OAc       189
117         H C OAc             H C OAc
            H                       H
          1,2-Xyl                 1,4-Xyl
```

Figure 2 Differentiation of PMAA from 1,2-linked and 1,4-linked xylose residues by deuterium labeling at C1 in methylation analysis.

```
            H                     H
        D / H C OAc         D / H C OMe
190 / 189   H C OAc             H C OAc   161 / 162
            H C OMe             H C OMe
------------------------------    ------------------------
            H C OMe             H C OMe
161         H C OAc             H C OAc       189
            H C OMe             H C OAc
            H                       H
          1,2-Frc                 2,6-Frc
```

Figure 3 Differentiation of PMAA from 1,2-linked and 2,6-linked fructose residues by deuterium labeling at C1 in methylation analysis.

Table 6 Characteristic chemical shifts in ^{13}C-NMR of polysaccharides.

Chemical shifts (ppm) in ^{13}C-NMR of polysaccharides	
C1 (β-glycosidic linkage)	103–106
C1 (α-glycosidic linkage)	98–103
Not anomeric C-atom involved in glycosidic linkage	80–87
C2–C5	65–75
CH$_2$OH	60–65
C=O	175–180
COOH	170–175
O-Me	55–61
O-Ac	20–23
CH$_3$	15–18

well distinguished from the other signals. The carboxylate group in uronic acids is downfield, whereas O-acetyl or N-acetyl groups have upfield-shifted methyl groups. Methyl protons and carbon atoms in deoxy sugars are upfield shifted as well. Some typical ranges of chemical shifts in ^{13}C-NMR- and ^{1}H-NMR spectra of oligo- or polysaccharides are given in Tables 6 and 7.

Table 7 Characteristic chemical shifts in ^{1}H-NMR of polysaccharides.

Chemical shifts (parts per million) in ^{1}H-NMR of polysaccharides	
H1 (β-glycosidic linkage)	5,3–5,8
H1 (α-glycosidic linkage)	4,5–4,8
H5	4,5–4,6
H2–H6	3,5–4,5
COOH	9–13
O-Me	3,3–3,5
O-Ac	2,0–2,2
CH$_3$	1,4–1,6

High molecular weight of polysaccharides eventually in combination with high viscosity of their solutions may result in reduced sensitivity and broad lines in NMR spectra. Improvement can be achieved by heating (up to 80°C) or by partial hydrolysis (best with specific endo-glycosidases, if available). NMR spectra of oligosaccharides often can be interpreted in more detail and often allow also the determination of the sequence of monosaccharide units (Lemieux and Stevens, 1966; Aspinall, 1976; Perlin, 1976; Jennings and Smith, 1980; Hall, 1980; Gorin, 1981; Barker *et al.* 1982; Perlin and Casu, 1982; Sweelley and Nunez, 1985; Jansson, Kenne, and Widmalm, 1989; Sheng, Cherniak, and van Halbeek, 1998; Duus, Gotfredsen, and Bock, 2000; Bubb, 2003).

REFERENCES

Adams, G. A. (1965) Complete acid hydrolysis, in *Methods in Carbohydrate Chemistry*, ed. R. L. Whistler, Academic Press, New York, vol. **5**, pp. 269–275; u. 285–287.

Aspinall, G. O. (1976) Polysaccharide methodology, in *Organic Chemistry, Series Two*, ed. G. O. Aspinall, Butterworths, London, vol. **7**, pp. 201–222.

Aspinall, G. O. (1982a) Isolation and fractionation of polysaccharides, in *The Polysaccharides*, ed. G. O. Aspinall, Academic Press, New York, vol. **1**, pp. 19–34.

Aspinall, G. O. (1982b) Chemical characterization and structure determination of polysaccharides, in *The Polysaccharides*, ed. G. O. Aspinall, Academic Press, New York, vol. **1**, pp. 35–131.

Aspinall, G. O., Molloy, J. A. and Craig, J. W. T. (1969) *Can. J. Biochem.*, **47**, 1063–1070.

Barker, R., Nunez, H. A., Rosevear, P., *et al.* (1982) *Methods Enzymol.*, **83**, 58–69.

Bauer, S. (2012) *Front. Plant Sci.*, **3**, DOI: 10.3389/fpls.2012.00045.

Bauer, S., Vasu, P., Persson, S., *et al.* (2006) *Proc. Natl. Acad. Sci. USA*, **103**, 11417–11422.

Biermann, J. and McGinnis, G. D. (1989) *Analysis of Carbohydrates by GLC and MS*, CRC Press, Boca Raton.

Binkley, S. B. (1965) Dialysis, in *Methods in Carbohydrate Chemistry*, eds. R. L. Whistler, J. N. BeMiller and M. L. Wolfrom, Academic Press, New York, vol. **V**, pp. 54–55.

Björndal, H., Lindberg, B. and Svensson, S. (1967) *Carbohydr. Res.*, **5**, 433–440.

Björndal, H., Hellerquvist, C. G., Lindberg, B., *et al.* (1970) *Angew. Chem.*, **16**, 643–674.

Blakeney, A. B., Harris, P. J. and Stone, B. A. (1983) *Carbohydr. Res.*, **113**, 291–299.

Blakeney, A. B. and Stone, B. A. (1985) *Carbohydr. Res.*, **140**, 319–324.

Blaschek, W. (1983) *J. Chromatogr.*, **256**, 157–163.

Blaschek, W., Koehler, H., Semler, U., *et al.* (1982) *Planta*, **154**, 550–555.

Blumenkrantz, N. and Asboe-Hansen, G. (1973) *Anal. Biochem.*, **54**, 484–489.

Böcker, S., Kehr, B. and Rasche, F. (2011) *Trans. Comput. Biol. Bioinform.*, **8**, 976–986.

Bouveng, H. O. and Lindberg, B. (1965) Native acetylated wood polysaccharides. Extraction with dimethyl sulfoxide, in *Methods in Carbohydrate Chemistry*, ed. R. L. Whistler, Academic Press, New York, vol. **5**, pp. 147–150.

Bradbury, A. G. W., Halliday, D. J. and Medcalf, D. G. (1981) *J. Chromatogr.*, **213**, 146–150.

Brett, C. T. (1981) *J. Exp. Bot.*, **32**, 1067–1077.

Brobst, K. M. (1972) Gas-liquid chromatography of trimethylsilyl derivatives, in *Methods in Carbohydrate Chemistry*, eds. R. L. Whistler and J. N. BeMiller, Academic Press, New York, vol. **6**, pp. 3–8.

Bubb, W. A. (2003) *Concepts Magn. Reson. A*, **19A**(1), 1–19.

Carpita, N. C. (1986) *Plant Physiol.*, **80**, 660–666.

Chai, W., Piskarev, V. and Wawson, A. M. (2001) *Anal. Chem.*, **73**, 651–657.

Chanzy, H., Chumpitazi, B. and Peguy, A. (1982) *Carbohydr. Polym.*, **2**, 35–42.

Churms, S. C. (1970) Gel chromatography of carbohydrates, in *Advances in Carbohydrate Chemistry and Biochemistry*, eds. R. S. Tipson and D. Horton, Academic Press, New York, vol. **25**, pp. 13–26.

Churms, S. C. (1996) *J. Chromatogr. A*, **720**(1-2), 75–91.

Ciucanu, I. and Kerek, F. (1984) *Carbohydr. Res.*, **131**, 209–217.

Cooper, T. G. (1981) *Biochemische Arbeitsmethoden*, Walter de Gruyter, Berlin, Germany.

Corradini, C. (1994) *Ann. Chim.*, **84**(9-10), 385–396.

Courtin, C. M., Van den Beroeck, H. and Delcour, J. A. (2000) *J. Chromatogr. A*, **866**(1), 97–104.

De Belder, A. N. and Normann, B. (1969) *Carbohydr. Res.*, **8**, 1–6.

De Ruiter, G. A., Schols, H. A., Voragen, A. G., *et al.* (1992) *Anal. Biochem.*, **207**(1), 176–85.

DeJongh, D. C. (1980) Mass spectrometry, in *The Carbohydrates: Chemistry and Biochemistry*, eds. W. Pigman, D. Horton and J. Wander, Academic Press, New York, vol. **IB**, pp. 1327–1353.

Dell, A., Reason, A. J., Khoo, K.-H., *et al.* (1994) *Methods Enzymol.*, **230**, 108–132.

Dische, Z. (1949) *J. Biol. Chem.*, **181**, 379–392.

Dische, Z. (1962) Color reactions of carbohydrates, in *Methods in Carbohydrate Chemistry*, eds. R. L. Whistler and M. L. Wolfrom, Academic Press, New York, vol. **1**, pp. 475–514.

Doco, T., O'Neill, M. A. and Pellerin, P. (2001) *Carbohydr. Polym.*, **46**, 249–259.

Dubois, M., Gilles, K. A., Hamilton, J. K., *et al.* (1956) *Anal. Chem.*, **28**, 350–356.

Duus, J. Ø., Gotfredsen, C. H. and Bock, K. (2000) *Chem. Rev.*, **100**, 4589–4614.

Fry, S. C. (1988) *The growing plant cell wall: chemical and metabolic analysis*, Essex, Longman Scientific and Technical.

Fry, S. C. (2011) *Ann. Plant Rev.*, **41**, 1–42.

Garcia-Raso, A., Martinez-Castro, I., Paez, M. L., *et al.* (1987) *J. Chromatogr.*, **398**, 9–20.

Ghebregzabher, M., Rufini, S., Monaldi, B., *et al.* (1976) *J. Chromatogr.*, **127**, 133–162.

Goldstein, I. J., Hay, G. W., Lewis, B. A., *et al.* (1965) Controlled degradation of polysaccharides by periodate oxidation, reduction, and hydrolysis, in *Methods in Carbohydrate Chemistry*, eds. R. L. Whistler, J. N. BeMiller and M. L. Wolfrom, Academic Press, New York, vol. **V**, pp. 361–370.

Gorin, P. A. (1981) *Adv. Carbohydr. Chem. Biochem.*, **38**, 13–104.

Grishutin, S. G., Gusakov, A. V., Markov, A. V., *et al.* (2004) *Biochim. Biophys. Acta*, **1674**, 268–281.

Gruppen, H., Hamer, R. J. and Voragen, A. G. J. (1991) *J. Cereal Sci.*, **13**(3), 275–290.

Guerrant, G. O. and Moss, C. W. (1984) *Anal. Chem.*, **56**, 633–638.

Hais, I. M. and Macek, K. (1963) *Paper Chromatography, a Comprehensive Treatise*, 3rd edn, New York, Academic Press.

Hakamori, S. (1964) *J. Biochem.*, **55**, 205–208.

Hall, L. D. (1980) High-resolution NMR spectroscopy, in *The carbohydrates: Chemistry and Biochemistry*, eds. W. Pigman, D. Horton and J. Wander, Academic Press, New York, vol. **IB**, pp. 1299–1326.

Harris, P. J. (1983) Cell walls, in *Isolation of Membranes and Organelles from Plant Cells*, eds. J. L. Hall and A. L. Moore, Academic Press, London, pp. 25–53.

Harris, P. J., Henry, R. J., Blakeney, A. B., *et al.* (1984) *Carbohydr. Res.*, **127**, 59–73.

Harris, P. J., Bacic, A. and Clarke, A. E. (1985) *J Chromatogr.*, **350**, 304–306.

Hay, G. W., Lewis, B. A. and Smith, F. (1965) Periodate oxidation of polysaccharides: general procedures, in *Methods in Carbohydrate Chemistry*, eds. R. L. Whistler, J. N. BeMiller and M. L. Wolfrom, Academic Press, New York, vol. **V**, pp. 357–361.

Henrissat, B. and Bairoch, A. (1996) *Biochem. J.*, **316**, 695–696.

Hicks, K. B., Lim, P. C. and Haas, M. J. (1985) *J. Chromatogr.*, **319**, 159–171.

Hon, S. and Srinivasan, K. S. V. (1983) *J. Appl. Polym. Sci.*, **28**, 1–10.

Jansson, P.-E., Kenne, L., Liedgren, H., *et al.* (1976) *Chem. Commun.*, **8**, 2–74.

Jansson, P.-E., Kenne, L. and Widmalm, G. (1989) *Carbohydr. Res.*, **188**, 169–191.

Jarvis, M. C. (1982) *Planta*, **154**, 344–346.

Jay, A. (1996) *J. Carbohydr. Chem.*, **15**(8), 897–923.

Jennings, H. J. and Smith, I. C. P. (1980) Determination of polysaccharide structures with ^{13}C-NMR, in *Methods in Carbohydrate Chemistry*, eds. R. L. Whistler and J. N. BeMiller, Academic Press, New York, vol. **VIII**, pp. 97–105.

John, M., Schmidt, J., Wandrey, C., *et al.* (1982) *J. Chromatogr.*, **247**, 281–288.

Jones, J. K. N. and Stoodley, R. J. (1965a) Fractionation by ultrafiltration, in *Methods in Carbohydrate Chemistry*, eds. R. L. Whistler, J. N. BeMiller and M. L. Wolfrom, Academic Press, New York, vol. **V**, pp. 47–48.

Jones, J. K. N. and Stoodley, R. J. (1965b) Fractionation using copper complexes, in *Methods in Carbohydrate Chemistry*, ed. R. L. Whistler, Academic Press, New York, vol. **5**, pp. 36–38.

Jork, H., Funk, W., Fischer, W., *et al.* (1989) *Dünnschicht – Chromatographie: Reagenzien und Nachweismethoden*, Weinheim, Wiley-VHC.

Joseleau, J.-P., Chambat, G. and Chumpitazi-Hermoza, B. (1980) *Carbohydr. Res.*, **90**, 339–344.

Kabel, M. A., Schols, H. A. and Voragen, A. G. J. (2001) *Carbohydr. Polym.*, **44**, 161–165.

Klok, J., Nieberg-Van Velzen, E. H., De Leeuw, J. W., *et al.* (1981) *J. Chromatogr.*, **207**, 273–275.

Knobloch, J. E. and Shaklee, P. N. (1997) *Anal. Biochem.*, **245**, 231–241.

Kobota, A., Yamashita, K. and Takasaki, S. (1987) *Methods Enzymol.*, **138**, 84–94.

Kochetkov, N. K. and Chizhov, O. S. (1966) *Adv. Carbohydr. Chem.*, **21**, 185–240.

Koizumi, K., Okada, Y. and Fukuda, M. (1991) *Carbohydr. Res.*, **215**(1), 67–80.

Layne, E. (1957) *Methods Enzymol.*, **3**, 447–454.

Lee, Y. C. (1996) *J. Chromatogr.*, **720**(1–2), 137–149.

Lemieux, R. U. and Stevens, J. D. (1966) *Can. J Chem.*, **44**, 249–262.

Lever, M. (1972) *Anal. Biochem.*, **47**, 273–279.

Lindberg, B. and Lönngren, J. (1978) *Methods Enzymol.*, **50**, 3–32.

Lindberg, B., Lönngren, J. and Svensson, S. (1975) *Adv. Carbohydr. Chem. Biochem.*, **31**, 185–241.

Loewus, F. and Shah, R. H. (1972) Gas-liquid chromatography of trimethylsilyl ethers of cyclitols, in *Methods in Carbohydrate Chemistry*, eds. R. L. Whistler and J. N. BeMiller, Academic Press, New York, vol. **VI**, pp. 14–20.

Lomax, J. A. and Conchie, J. (1982) *J. Chromatogr.*, **236**, 385–394.

Lomax, J. A., Gordon, A. H. and Chesson, A. (1983) *Carbohydr. Res.*, **122**, 11–22.

Masuko, T., Minami, A., Iwasaki, N., *et al.* (2005) *Anal. Biochem.*, **339**, 69–72.

McGinnis, G. D. and Fang, P. (1980) High performance liquid chromatography, in *Methods in Carbohydrate Chemistry*, eds. R. L. Whistler and J. N. BeMiller, Academic Press, New York, vol. **VIII**, pp. 33–43.

McNeil, M., Darvill, A., Aman, A. G., *et al.* (1982) *Methods Enzymol.*, **83**, 3–45.

Meier, H. (1965) Fractionation by precipitation with barium hydroxide, in *Methods in Carbohydrate Chemistry*, ed. R. L. Whistler, Academic Press, New York, vol. **5**, pp. 45–46.

Nilsson, G. S., Anderson, M., Ruzgas, T., *et al.* (1998) *Anal. Biochem.*, **265**, 151–156.

Nishinari, K., Takemasa, M., Zhang, H., *et al.* (2007) Storage plant polysaccharides: xyloglucans, glactomannans, glucomannans, in *Comprehensive Glycoscience*, ed. J. P. Kamerling, Elsevier, Amsterdam, The Netherlands, vol. **2**, pp. 613–652.

Okatch, H. (2003) *Afr. J. Biotechnol.*, **2**(12), 636–644.

Oshima, R., Yoshikawa, A. and Kumanotani, J. U. (1981) *J. Chromatogr.*, **213**, 142–145.

Painter, T. J. (1965) Partial acidic and enzymic hydrolysis, in *Methods in Carbohydrate Chemistry*, eds. R. L. Whistler, J. N. BeMiller and M. L. Wolfrom, Academic Press, New York, vol. **V**, pp. 280–285.

Pecina, R., Bonn, G., Burtscher, E., *et al.* (1984) *J. Chromatogr.*, **287**, 245–258.

Perlin, A. S. (1976) Carbon-13 NMR-spectroscopy of carbohydrates, in Organic Chemistry*, Series Two, Carbohydrates*, ed. G. O. Aspinall, Butterworth, London, vol. **7**, pp. 1–34.

Perlin, A. S. and Casu, B. (1982) Spectroscopic methods, in *The Polysaccharides*, ed. G. O. Aspinall, Academic Press, New York, vol. **1**, pp. 133–193.

Perry, M. B. and Hulyalkar, R. K. (1965) *Can. J. Biochem.*, **43**, 573–584.

Phillips, L. R. and Fraser, B. A. (1981) *Carbohydr. Res.*, **90**, 149–152.

Pittet, A. O. (1965) Dissolution of polysaccharides, in *Methods in Carbohydrate Chemistry*, eds. R. L. Whistler, J. N. BeMiller and M. L. Wolfrom, Academic Press, New York, vol. **V**, pp. 3–5.

Rajakyla, E. (1986) *J. Chromatogr.*, **353**, 1–12.

Read, S. M. and Northcote, D. H. (1981) *Anal. Biochem.*, **116**, 53–64.

Rocklin, R. D. and Pohl, C. A. (1983) *J. Liq. Chromatogr.*, **6**, 1577–1590.

Ruiz-Matute, A. I., Hernández-Hernández, O., Rodríguez-Sánchez, S., *et al.* (2011) *J. Chromatogr. B*, **879**, 1226–1240.

Scott, J. E. (1965) Fractionation by precipitation with quaternary ammonium salts, in *Methods in Carbohydrate Chemistry*, eds. R. L. Whistler, J. N. BeMiller and M. L. Wolfrom, Academic Press, New York, vol. **V**, pp. 38–44.

Selvendran, R. R. and O'Neill, M. A. (1987) Isolation and analysis of cell walls from plant material, in *Methods of Biochemical Analysis*, ed. D. Glick, John Wiley & Sons, London, vol. **32**, pp. 25–153.

Selvendran, R. R., Stevens, B. J. H. and O'Neill, M. A. (1985) Developments in the isolation and analysis of cell walls from edible plants, in *Biochemistry of Plant Cell Walls*, eds. C. T. Brett and J. R. Hillman, Cambridge University Press, pp. 39–78.

Sheng, S., Cherniak, R. and van Halbeek, H. (1998) *Anal. Biochem.*, **256**, 63–66.

Sherma, J. and Shirley, C. C. (1991) *CRC Handbook of Chromatography: Carbohydrates*, CRC-Press, Boca Raton, vol. **II**.

Shilova, N. V. and Bovin, N. V. (2003) *Russ. J. Bioorg. Chem.*, **29**(4), 309–324(translated from Bioorganicheskaya Khimiya, 2003, **29** (4): 339–355, original Russian Text Copyright © 2003 by Shilova, Bovin.

Sloneker, J. H. (1972) Gas-liquid chromatography of alditol acetates, in *Methods in Carbohydrate Chemistry*, eds. R. L. Whistler and J. N. BeMiller, Academic Press, New York, vol. **VI**, pp. 20–24.

Smith, I. (1960) *Chromatographic and Electrophoretic Techniques, Vol 1, Chromatography*, 2nd edn, London, Heinemann.

Stahl, E. (1967) *Dünnschicht-Chromatographie: Ein Laboratoriumshandbuch*, 2nd edn, New York, Springer.

Suortti, T. (1997) *J. Chromatogr. A*, **763**(1-2), 331–335.

Sweeley, C. C. and Nunez, H. A. (1985) *Annu. Rev. Biochem.*, **54**, 765–801.

Sweet, D. P., Shapiro, R. H. and Albersheim, P. (1975) *Carbohydr. Res.*, **40**, 217–225.

Tang, H., Mechref, Y. and Novotny, M. V. (2005) *Bioinformatics*, **21**(Suppl. 1), 431–439.

Taylor, R. L. and Conrad, H. E. (1972) *Biochemistry*, **11**, 1383–1388.

Thornber, J. P. and Northcote, D. H. (1962) *Biochem. J.*, **82**, 340–346.

Touchstone, J. C. and Dobbins, M. F. (1983) *Practice of thin layer chromatography*, 3rd edn, New York, John Wiley& Sons.

Updegraff, D. M. (1969) *Anal. Biochem.*, **32**, 420–424.

Valent, B. S., Darvill, A. G., McNeil, M., *et al.* (1980) *Carbohydr. Res.*, **79**, 165–192.

Verhaar, L. A. T. and Kuster, B. F. M. (1981) *J. Chromatogr.*, **220**, 313–328.

Viseux, N., de Hoffmann, E. and Domon, B. (1998) *Anal. Chem.*, **70**, 4951–4959.

Waeghe, T. J., Darvill, A. G., McNeil, M., *et al.* (1980) *Carbohydr. Res.*, **123**, 281–304.

Wells, G. B., Kontoyiannidou, V., Turco, S. J., *et al.* (1982) *Methods Enzymol.*, **83**, 132–137.

Whistler, R. L. and Anisuzzaman, A. K. M. (1980) Gel permeation chromatography, in *Methods in Carbohydrate Chemistry*, eds. R. L. Whistler and J. N. BeMiller, Academic Press, New York, vol. **VIII**, pp. 45–53.

Whistler, R. L. and Feather, M. S. (1965) Hemicellulose. extraction from annual plants with alkaline solution, in *Methods in Carbohydrate Chemistry*, ed. R. L. Whistler, Academic Press, New York, vol. **5**, pp. 144–145.

Whistler, R. L. and Sannella, J. L. (1965) Fractional precipitation with ethanol, in *Methods in Carbohydrate Chemistry*, eds. R. L. Whistler, J. N. BeMiller and M. L. Wolfrom, Academic Press, New York, vol. **V**, pp. 34–36.

Wing, R. E. (1972) Thin-layer chromatography, in *Methods in Carbohydrate Chemistry*, eds. R. L. Whistler and J. N. BeMiller, Academic Press, New York, vol. **VI**, pp. 42–53; 54–59; 60–64.

Wolfrom, M. L. and Franks, N. E. (1965) Partial acid hydrolysis, in *Methods in Carbohydrate Chemistry*, eds. R. L. Whistler, J. N. BeMiller and M. L. Wolfrom, Academic Press, New York, vol. **V**, pp. 276–280.

Wood, P. J., Erfle, J. D., Teather, R. M., *et al.* (1994) *J. Cereal Sci.*, **19**, 65–75.

Wyatt, P. J. (1993) *Anal. Chim. Acta*, **272**, 1–40.

Yamasita, K., Mizuochi, T. and Kobata, A. (1982) *Methods Enzymol.*, **83**, 105–126.

Zalyalieva, S. V., Kabulov, B. D., Akhundzhanov, K. A., *et al.* (1999) *Cehm Nat Comp*, **35**, 1–13.

Zweig, G., Sherma, J. and Churms, S. C. (1982) *CRC Handbook of Chromatography: Carbohydrates*, CRC-Press, Boca Raton, vol. **I**.

Analytical Strategies for Multipurpose Studies of a Plant Volatile Fraction

Cecilia Cagliero, Barbara Sgorbini, Chiara Cordero, Erica Liberto, Carlo Bicchi and Patrizia Rubiolo

Dipartimento di Scienza e Tecnologia del Farmaco, Università di Torino, Torino, Italy

1 INTRODUCTION

The introduction of metabolomics into plant science and its evolution over recent decades have led to its being definitively accepted that each secondary metabolite plays a well-defined role in the plant's life cycle, that all compounds together concur to the plant's metabolism, and that they are all directly or indirectly connected to one another, independently of their structures. This obviously also concerns the plant volatile fraction, making it an important biosensor that is not only diagnostic of plant metabolism and changes but also, quite often, an effective marker of other fractions, including nonvolatile metabolites.

Any discussion of analysis strategies, as that which follows, cannot leave the above-mentioned considerations out of the picture, and as a preliminary, some general definitions must be introduced, which will be appropriate to frame the examination of the volatile fraction and its analysis:

(i) the term *metabolomics* was introduced at the end of the 1990s and is defined as the systematic study of the unique chemical metabolite fingerprints (the *metabolome*) resulting from specific cellular processes (Oliver *et al*., 1998);

(ii) metabolomics concerns several thousand metabolites that cannot all be analyzed at the same time; this means that reliable *profiling*

and *fingerprinting* methods are required for diagnostic fractions or extracts to measure absolute or relative amounts of metabolites involved in the biological phenomenon investigated (Fiehn *et al*., 2000; Glassbrook and Ryals, 2001; Harrigan and Goodacre, 2003; Sumner, Mendes, and Dixon, 2003; Halket *et al*., 2005). *Profiling* in general produces a detailed analytical profile of the plant sample obtained by combining a separation technique with a spectroscopic technique (e.g., GC/MS, LC/MS, CE/MS etc.); components can be identified and quantified in this profile. *Fingerprinting* refers to more general and rapid high throughput screenings and is mainly applied to discriminate and classify samples. In general, it is carried out with a nonseparative approach such as headspace-mass spectrometry(HS-MS), direct infusion-MS, NMR, and FT-IR (MIR, NIR, etc.) combined with statistical data processing (Halket *et al*., 2005.);

(iii) more recently, Maffei and coworkers defined the *plant volatilome* " ….. as the complex blend of essential oils (EOs) and volatile organic compounds (VOCs) fed by different biosynthetic pathways and produced by plants, constitutively and/or after induction, as a defense strategy against biotic and abiotic stress." (Maffei, Gertsch, and Appendino, 2011); the volatilome can be adopted as an approach to studying biological phenomena mainly on the basis of

monitoring the volatile fraction (Bicchi and Maffei, 2012).

Analysis of the volatile fraction therefore requires methods and technologies whose performance is suitable for monitoring variations in the fraction's composition. More specifically, they must sample, detect, identify, and, when necessary, quantify the volatile component(s) (in general, secondary metabolites)

responsible for a plant's biological property or characteristic or that can be taken as marker(s) of a given biological phenomenon and/or that are appropriate to follow the reaction(s) of the vegetable organism when its metabolism is altered.

The above-mentioned considerations on the extended role played by the volatile fraction have greatly contributed to increasing interest in studying it, in terms not only of chemical composition but also of significance and function.
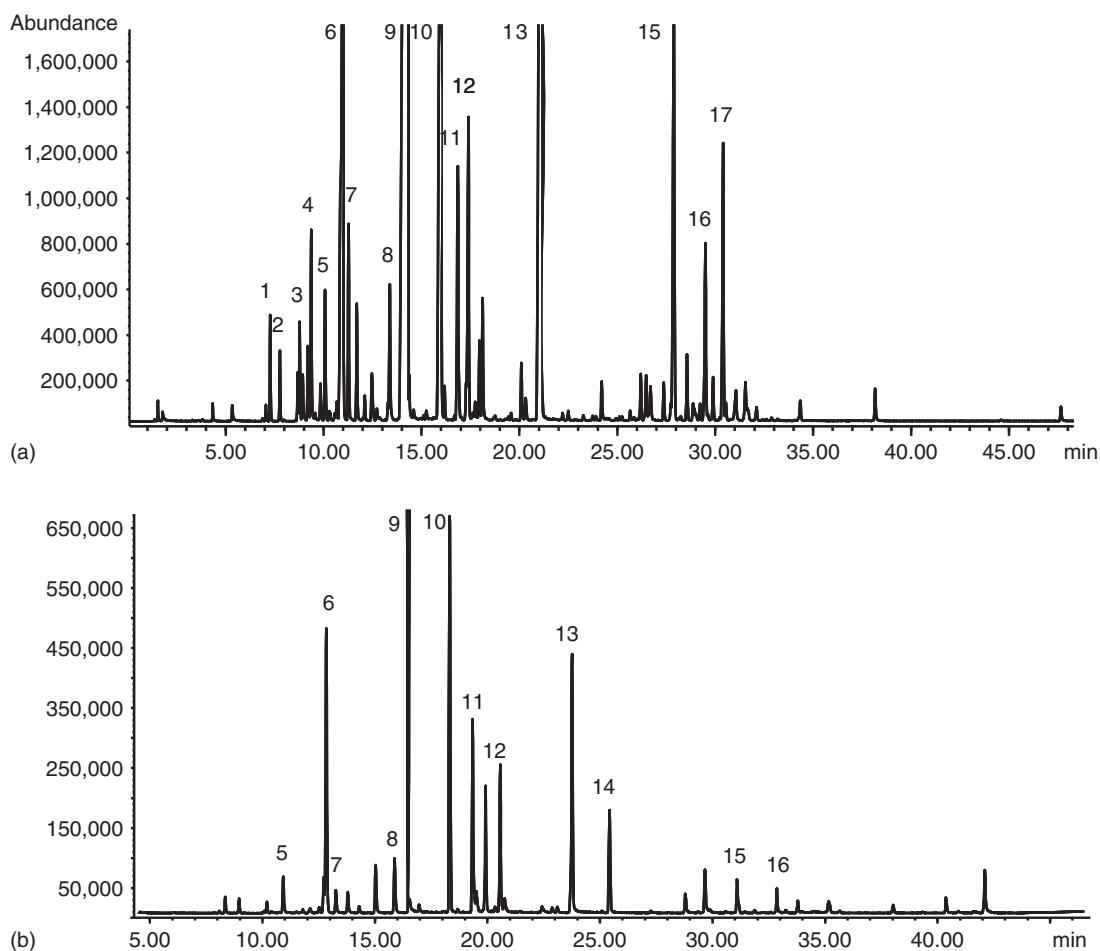
Figure 1 Comparison between HS-SPME (a) and essential oil (b) GC-MS profiles of the same sample of lavender (*Lavandula angustifolia* Mill.). Analysis system: Agilent 6890 GC-5973 MSD system (Agilent, Little Falls, DE, USA). Sample preparation: SPME fiber: CAR/PDMS/DVB, 2 cm; sample amount: 20 mg; vial volume: 20 mL; sampling time: 30 min; and sampling temperature: 60°C. GC-MS conditions: column: HP5 (30 m, d_c 0.25 mm, d_f 0.25 μm); temperature program: 50°C (1 min)//3°C/min//250°C (5 min). Desorption temperature: 230°C; desorption time: 5 min; injection mode: split; ratio 1 : 20; carrier gas: He; flow rate: 1 mL min^{-1}. Essential oil obtained by hydrodistillation of 100 g of dried plant. Analysis conditions: injection volume: 1 μL of essential oil diluted 1 : 200 in cyclohexane. List of identified compounds: (1) α-pinene, (2) camphene, (3) β-pinene, (4) myrcene, (5) Δ-3-carene, (6) limonene+1,8-cineole, (7) *cis*-β-ocymene, (8) α-terpinolene, (9) linalool, (10) camphor, (11) borneol, (12) 4-terpineol, (13) lynalyl acetate, (14) lavandulyl acetate, (15) β-caryophyllene, (16) *trans*-β-farnesene, and (17) germacrene D.

As already defined in the chapter titled **Headspace Sampling and Gas Chromatography: A Successful Combination to Study the Composition of a Plant Volatile Fraction** of this handbook, in chemical terms, the *volatile fraction* of a matrix of vegetable origin can generally speaking be defined as a mixture of volatiles that can be sampled because of their ability to vaporize, either spontaneously, and/or under suitable conditions, or by adopting appropriate techniques (Bicchi and Maffei, 2012; Rubiolo et al., 2010a, 2010b). The term *volatile fraction analysis* therefore includes a range of approaches and/or techniques, which produce samples that are all representative of the volatiles characterizing the vegetable matrix but that, conversely, may have different and unrelated or noncomparable compositions: headspace (HS), essential oils, and extracts prepared with specific techniques, as well as flavors, fragrances, and aromas, can all be involved in the volatile fraction framework. Some other fundamental and complementary considerations, related to the volatile fraction of a plant, or more in general to other matrices or living organisms, must additionally be considered.

The terms *headspace* and *essential oil* are still, but erroneously, quite often confused, although their definitions clearly distinguish them. HS is defined as the gaseous or vapor phase in equilibrium (or not) with a (plant) matrix that can be sampled to characterize its composition (Kolb and Ettre, 1997), whereas an essential oil is officially defined as the "odorous product, usually of complex composition, obtained from a botanically defined plant raw material by steam distillation, dry distillation, or a suitable mechanical process without heating" (European Pharmacopoeia, 2014). Figure 1 reports a comparison between HS sampled by HS-SPME and essential oil GC-MS profiles of the same sample of lavender (*Lavandula angustifolia* Mill.), showing the difference between the abundance of the main components in consequence of the adopted sample preparation techniques (analysis conditions and peak identification are included in the figure caption).

Aroma, flavor, and fragrance are also worth mentioning because they are directly or indirectly related to the volatile fraction and, as a consequence, to the odorous properties of a plant. By this, we mean the interaction of volatile compounds coming into contact with the olfactory tissue situated in a specific area of the nasal cavity, known as *regio olfactoria*, or else with an insect's sensitive and/or "olfactory"

organs. Although the three terms are not defined by an international standard, the concepts behind them are clear; according to Belitz, Grosch, and Schieberle (2009), (i) "aroma substances" are volatiles that can be perceived orthonasally and retronasally by the odor receptor sites of the smell organ (the term aroma is often confused with essential oil or HS); (ii) "fragrance" may be defined as a sweet, pleasant scent perceived orthonasally from a nonfood material; and (iii) flavor is a broad and comprehensive term characteristic of a food, generally involved with the overall sensation provided by the interaction of taste, odor, and the sensation of texture when food is consumed.

2 STRATEGIES OF ANALYSIS

2.1 An Introduction

The strategy discussed in the following text first focuses on profiling the volatile fraction, taking as reference analysis technique GC with mass spectrometry (MS) or FID as detectors, then examining fingerprinting with a nonseparative MS approach. The development of an analysis method naturally implies adequate knowledge of the goal(s), and it is hoped to achieve, so that the first, and decisive, choice between targeted or nontargeted approaches can be made appropriately (Pierce *et al.*, 2008).

A nontargeted approach consists of one (or more) analyses in which as many chromatographic peaks as possible (or those above a selected intensity) are located in the chromatogram and characterized by their GC retention indices and mass spectra. This approach aims principally (i) to define the composition of the sample(s) under investigation, including the main classes and/or groups of compounds and/or components and their relative abundance and (ii) to build a (reference) profile for use as a parameter providing reliable recognition or comparison of the samples within the set investigated.

A targeted approach comprises the quantitative analysis (see **Headspace Sampling and Gas Chromatography: A Successful Combination to Study the Composition of a Plant Volatile Fraction**) of a limited number of analytes, selected from the most significant groups characterizing the sample(s) investigated, and adopted as representative markers of the biological phenomenon or quality (see in the following discussion) describing the sample(s) investigated (Rubiolo *et al.*, 2013).

The two approaches are absolutely complementary (for example, a targeted method is often derived from the results of a nontargeted experiment) and are currently used in combination in *dereplication* studies (i.e., studies in which plant extracts containing nuisance or known secondary metabolites are differentiated from those containing new/interesting compounds) (Lang *et al.*, 2008).

The everyday strategy can vary depending on the sample composition or on whether a screening or an exhaustive study is required. In the case of screening, the analytical method must provide a significant profile, by either profiling or fingerprinting, which contains sufficient information to characterize the sample(s) under investigation and representative of the problem to be resolved. An exhaustive study implies a conventional strategy that aims to identify as many components as possible in the volatile fraction investigated, so as to characterize it not only through analysis but also by fractionation of the sample, so as to isolate unknown components (or specific subfractions) or to suitably enrich them, and then submit them to spectroscopic investigation (NMR, MS, IR, etc.) to elucidate their structure. An exhaustive investigation is not indispensable, but it must be done when minor components or significant unknowns are to be enriched or identified, for example, because they are responsible for one of the biological activities of the sample investigated.

The strategy to adopt, both for screening the plant volatile fraction with separative (profiling) and non-separative (fingerprinting) methods and for exhaustive studies, is discussed in greater detail in the following section.

2.2 Profiling of the Volatile Fraction

This section deals with sample preparation for HS and conventional techniques, analysis with GC techniques for separation, in combination with FID and MS for identification and quantitation, and statistical data processing (e.g., chemiometrics) of the resulting qualitative and quantitative profiles.

The strategy of profiling the volatile fraction has changed radically over the past 15–20 years, not only because of the rapid evolution of analytical technology, but also spurred by the dramatic increase in the number of chemical analyses; a routine laboratory is required to run for control and safety assessment purposes. This evolution has affected all the main steps

of an analytical procedure, that is, sample preparation, analysis (separation and detection techniques), and data processing. In particular, data processing has now become an active and complementary part of the entire analytical process, offering a further level of information in particular when complex datasets have to be processed, and results must be delivered rapidly.

Much effort has been focused on developing automatic systems, the general thrust aimed at merging the "*sample preparation–analysis–data elaboration*" sequence into a single step, in what are known as *total analysis systems* (TAS) (Manz, Graber, and Widmer, 1990; Dittrich, Tachikawa, and Manz, 2006). These systems are important not only for routine use, when large numbers of relatively similar samples must be processed daily, but also in research in the plant field (e.g., metabolomics) in which the chemical definition of a biological phenomenon requires a statistically significant number of reliable results.

Analysis of the plant volatile fraction is perfectly compatible with the TAS approach, as it requires single-step and highly effective sample preparation techniques (HS and high concentration capacity headspace techniques) that are easy to automate and combine with GC or GC-MS, providing highly reliable data that is suitable for on-line processing.

2.2.1 Sample Preparation

The isolation and/or recovery of the volatile fraction from a plant can be achieved through a limited number of complementary approaches, that is, HS techniques, essential oil distillation, and, more seldom, extraction with suitable solvents and/or techniques combined on-line or off-line with distillation or vacuum distillation of the resulting extract(s).

The most widely used and traditional representative of the plant volatile fraction is the essential oil, not least because it enables volatiles to be isolated in well-defined products that can be used as an ingredient in several industrial fields (mainly the cosmetics, food, and pharmaceutical fields). With the exception of those of the citrus fruits, which are in general obtained by cold mechanical processes, essential oils are prepared by steam- or hydro-distillation, in agreement with the definition given earlier (European Pharmacopoeia, 2014) with dedicated apparatuses based on that first described by Clevenger (1928a, 1928b): Figure 2 shows a scheme
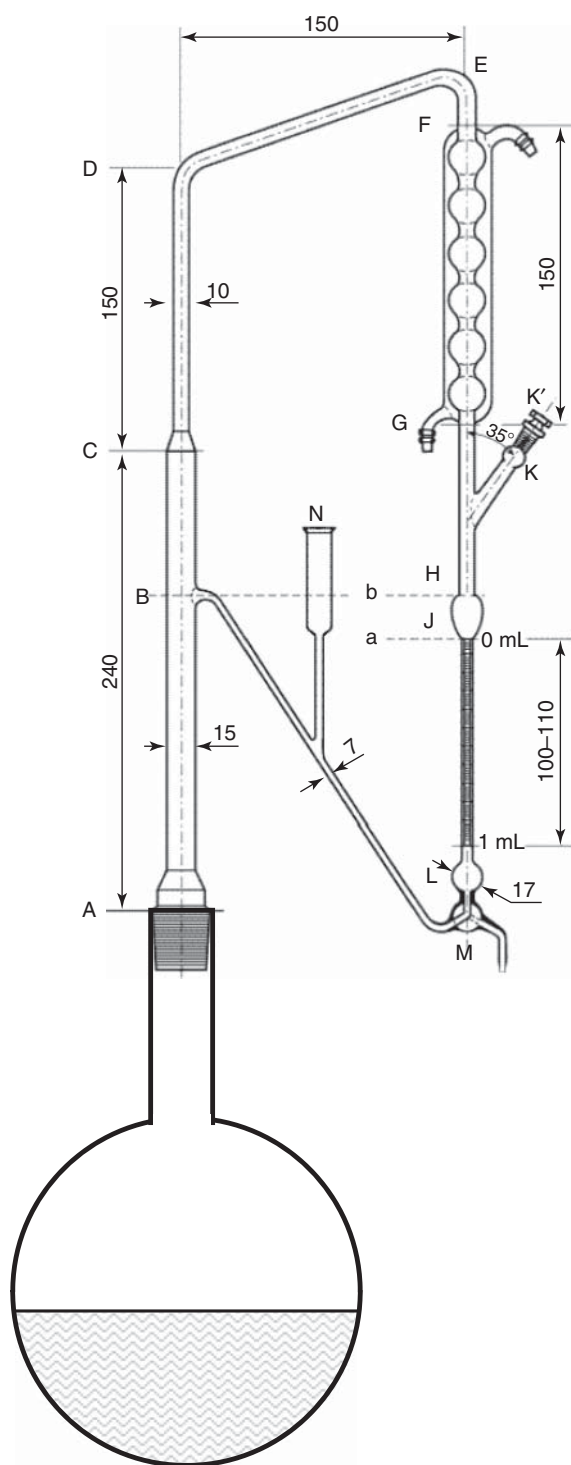
Figure 2 Scheme of the apparatus to obtain essential oils reported in the European Pharmacopoeia.

of the apparatus to obtain essential oils described in the European Pharmacopoeia. Essential oils are representative of the volatile fraction, as they are prepared at relatively mild temperature: hydro- or steam-distillation never exceeds 100°C, being based on the theory of immiscible liquid distillation, and has the advantage that it can easily and reliably be scaled up to the industrial level. Conversely, it has some limitations, including, in particular, (i) the very high volatility components are lost because they are difficult to trap, (ii) water-soluble components are lost to an extent that depends on their solubility, and (iii) artifacts may be formed because of heating the vegetable material for at least 2 h at a temperature close to 100°C, frequently in a mildly acidic aqueous medium. In addition, it obviously cannot be combined on-line to a GC-MS system, making profiling not suitable to be automatized and time-consuming.

The static headspace (S-HS) profile probably represents the volatile fraction of a plant more accurately because, in general, it can be carried out at lower temperature than hydro- or steam-distillation, and is done on the matrix as such, without any other processes or chemical treatments. It thus gives a true picture of the compounds released from a plant in function of their volatility. However, it is limited by its relatively low sensitivity, and the resulting sample is not precisely defined. At the laboratory scale, this last limitation can only be overcome by dynamic headspace sampling (D-HS) combined with analyte cryotrapping or adsorption or sorption on suitable material, followed by solvent elution; at the industrial scale, more sophisticated approaches must be applied (e.g., "enfleurage"). High concentration capacity-headspace techniques (HCC-HS) (see **Headspace Sampling and Gas Chromatography: A Successful Combination to Study the Composition of a Plant Volatile Fraction**) and D-HS techniques can provide plant volatile fractions that are enriched by up to several orders of magnitude compared to S-HS; their main limit is that component ratios in the resulting sample can differ from those in the original plant matrix, being conditioned by their interactions (sorption or adsorption) with the accumulating material, whose chemical nature influences their recovery in function of their structure and polarity.

Extraction with suitable solvents can also be very useful to provide samples from which a representative profile of the volatile fraction can be obtained by HS sampling or by distillation or vacuum distillation (see in the following discussion). Yield

and throughput of extraction and/or distillation can be improved by microwave assistance (MAE, MAHD, etc.) (Ferhat, Meklati, and Chemat, 2007, 2008; Sahraoui *et al*., 2008) or ultrasound assistance [ultrasound-assisted extraction (UAE) and ultrasound assisted hot air drying (UAHD)] (Chemat, Zill-e-Huma, and Kamran, 2011; Mason, Chemat, and Vinatoru, 2011); these procedures act as an aid to improve the process performance, without modifying extraction principles and mechanisms, nor the composition. The case of supercritical fluid extraction (SFE) is different in that it is a technique exploiting the peculiar characteristics of a fluid (generally CO_2) in the supercritical state, whose polarity and selectivity can be varied either physically, through its density (i.e., temperature and pressure), or chemically, by means of a solvent used in percentages that keep the fluid in the supercritical state. Beyond this limit, the extraction can only be defined as liquid CO_2 extraction (Fornari *et al*., 2012; Herrero *et al*., 2010; Zougagh, Valcarcel, and Rios, 2004).

2.2.2 Analysis

It is at the analysis step that the methods and results of profiling differ from those of fingerprinting (see in the following discussion). As already mentioned, profiling requires as many components as possible to be separated in a single run and in as short as possible a time, thus making GC the technique of choice to separate the components of a volatile fraction. The renewed interest in the volatile fraction, principally in profiling applications (metabolomics), has enhanced the importance of chromatographic data such as retention indices (I^Ts) (D'Acampora Zellner *et al*., 2008) and normalized retention times (retention time locking, RTL) (Giarrocco, Quimby, and Klee, 1997; Blumberg and Klee, 1998) combined on-line or off-line with mass spectra of plant volatile fraction components, as parameters for automatic identification (in some cases only their location in the chromatogram). I^Ts concepts and applications are described in detail in the chapter titled **Headspace Sampling and Gas Chromatography: A Successful Combination to Study the Composition of a Plant Volatile Fraction**, paragraph 3.5. Profiling applications imply the complementary (although still fundamental) step of component identification, as these applications offer the possibility of comparing analytical profiles, if possible automatically. This means that reproducible and reliable peak location in the GC or GC-MS pattern, and alignment through chromatographic/mass spectral data, are indispensable for correct sample comparison, whether or not there will be a subsequent statistical processing step. This is a general approach whereby volatiles can be identified, and it is possible to discriminate samples within a set, or characteristics or properties of samples, on the basis of their qualitative and quantitative chromatographic profiles; in turn, this makes it possible, for instance, to propose an interpretation of the biological phenomenon being investigated or to detect sample origin or adulteration. All gas chromatographic techniques can therefore be involved in profiling: conventional and fast GC, enantioselective GC, and multidimensional GC with both heart-cut 2DGC and comprehensive GC (GCxGC) combined with FID or, better, with MS. The choice among these techniques naturally depends on the analysis aim and on the complexity of the sample investigated. The appropriate choice is (i) fast GC when a large number of samples must be analyzed, if possible adopting conditions determined by applying method translation, to make profiles comparable over time to those of conventional GC, and also provided that sample preparation techniques taking compatible times are adopted (Klee and Blumberg, 2002; Bicchi *et al*., 2010a); (ii) enantioselective GC, mainly with cyclodextrin derivatives as chiral selectors (see **Headspace Sampling and Gas Chromatography: A Successful Combination to Study the Composition of a Plant Volatile Fraction**), when chirality, or better enantiomeric ratio (er) or excess (ee), are mandatory to discriminate between samples or to detect fraud (Liberto *et al*., 2008; Rubiolo *et al*., 2010b; Bicchi *et al*., 2010b); (iii) heart-cut 2DGC when one part, or a small number of only partially separated complex parts, of a chromatogram is automatically transferred from a first column (first dimension) to a second one coated with a different stationary phase, for more effective separation (Dugo, Dugo, and Mondello, 2002); and (iv) GCxGC when a very complex sample has to be analyzed simultaneously on columns coated with orthogonal stationary phases to obtain a significant and diagnostic profile (Tranchida *et al*., 2011). These techniques are described in detail in the chapter titled **Headspace Sampling and Gas Chromatography: A Successful Combination to Study the Composition of a Plant Volatile Fraction**. Figure 3 compares the Fast-GC (a) and GCxGC (b) profiles of a Vetiver

essential oil (*Chrysopogon zizanioides* (L.) Roberty) originating from Haiti showing how the choice of analysis technique can influence the resulting profiles and how it should be determined by the expected or required results (e.g., screening vs exhaustive studies) (for analysis conditions and peak identification, see the figure caption).

Detection plays a crucial role in the analysis of volatile fractions. The most widely used detector in routine work is MS because it not only provides a signal indicating the analyte's retention in the GC column but also adds further information orthogonal to the GC data, by giving spectral data suitable for structural characterization and identification of components. The reliability and high productivity of bench-top quadrupole mass spectrometric detectors, together with their acceptable cost, have greatly contributed to replacing conventional GC detectors in most applications and, in particular, in analysis of the plant volatile fraction. Conventional GC detectors [both universal (FID and TCD) and selective (NPD, FPD, and ECD)] are now seldom used in this field, except for FID, which is used for quantitation (see **Headspace Sampling and Gas Chromatography: A Successful Combination to Study the Composition of a Plant Volatile Fraction**; Cordero *et al*., 2012a) and FPD and sulfur chemiluminescence detectors (SCD) used on sulfureted compounds for specific applications.

A discussion of modern MS would require a dedicated paragraph, so this section will only give a brief mention of the additional opportunities that it can offer for the analysis of the volatile fractions. Other chapters in this book deal with this topic in greater detail. Modern MS detectors adopt mass analyzers based on different physical principles, depending on the applications, performance, and results required [quadrupole, ion trap, time of flight, and combinations of these in MS-MS (mass spectrometry/mass spectrometry)] or hybrid systems). Conventional and fast GC-MS and heart-cut 2DGC analyses can reliably be carried out with the latest generation of fast quadrupole or magnetic sector mass analyzers (Rubiolo *et al*., 2008). The latest generation of fast quadrupole analyzers can achieve frequencies up to 50 Hz (i.e., number of recorded spectra per second) that can successfully be used as detectors for GCxGC, although with some limitations in the mass range (Debonneville and Chaintreau, 2004; Cordero *et al*., 2007; Purcaro *et al*., 2010). Higher acquisition speeds can only be obtained with time-of-flight

mass spectrometry (TOFMS) analyzers with orthogonal acceleration or with array detectors, which can acquire up to one hundred spectra per second (Watson *et al*., 1990) over a wide spectral mass range and/or, at a lower scan speed, can operate at very high mass resolution. Thanks to their high instrumental precision (5 ppm) and dedicated software, these provide a very useful parameter for analyte identification, that is, the molecular formula. Similar results can be obtained with high performance ion trap analyzers (e.g., Orbitrap).

Tandem mass spectrometry or MS/MS (MS^n) is a valid option, which is achieved by coupling more than one analyzer. This approach involves the fragmentation of selected precursor ions, deriving from the initial ionization of an analyte in the first analyzer, and separation of the resulting fragments in the second analyzer; this process provides further structure-specific information and/or, at the same time, high selectivity and sensitivity, making it very useful for trace analyses (Sparkman, Penton, and Kitson, 2011; Koesukwiwat *et al*., 2010). MS/MS can be achieved in a spatial or a temporal domain: "tandem-in-space" is generally carried out with triple quadrupole (QqQ) mass analyzers, whereas "tandem-in-time" is based on ion-trap (IT-MS) analyzers. More recently, tandem-MS has dramatically increased its performance with the adoption of high resolution TOF-MS analyzers as a second MS unit in the so-called hybrid mass spectrometers; the resulting MS-MS system consists of the combination of a quadrupole or an ion-trap analyzer with a high resolution TOF-MS analyzer (Q-TOF or IT-TOF, respectively). The resulting hybrid system thus combines the advantages of tandem MS with those of high resolution analyzers. Low or high resolution tandem-MS can also be used with reduced fragmentation, by applying soft ionization ion sources (such as chemical ionization) to obtain a limited number of diagnostic fragments (e.g., quasi-molecular ions $[M - H]^+$ or $[M - H]^-$) in the first ionization step.

2.2.3 Data Processing

The routine use of chemiometrics and statistics for postacquisition data processing has contributed greatly to increasing the amount of information that can be obtained from an analytical procedure. The GC profile of a plant volatile fraction submitted to appropriate statistical processing can
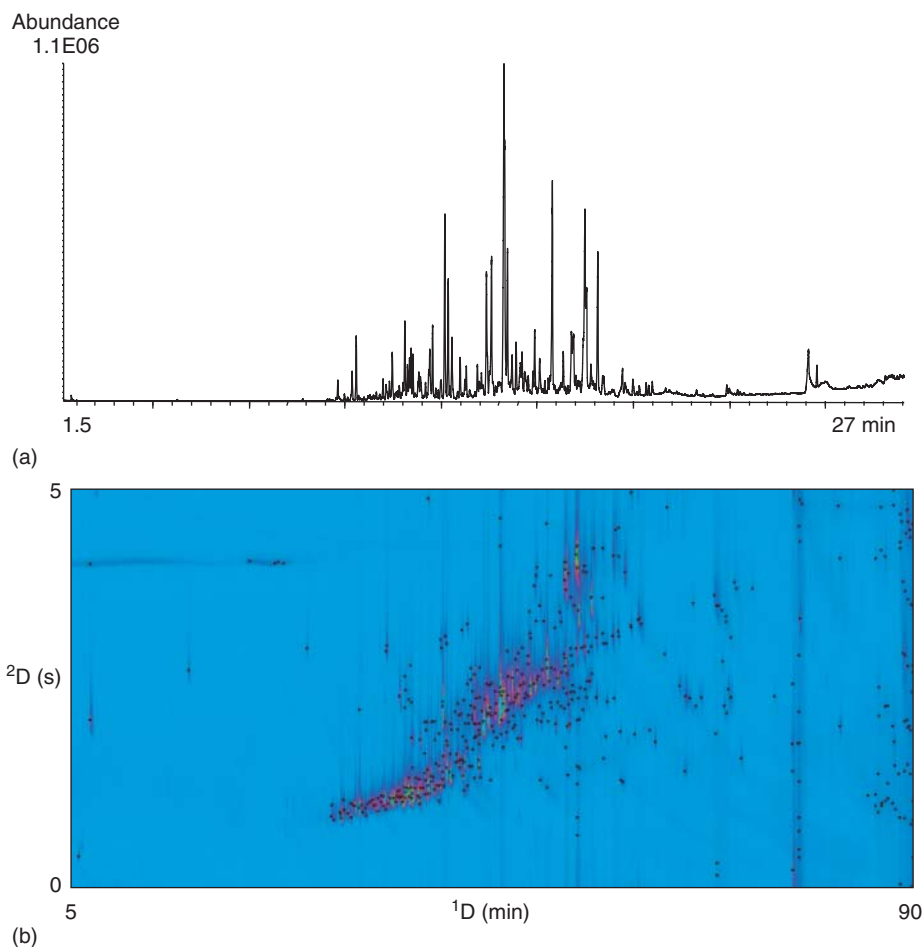
Figure 3   GC profiles of Vetiver (*Chrysopogon zizanioides* (L.) Roberty) essential oil from Haiti: (a) monodimensional fast-GC-TIC-MS profile, (b) two-dimensional comprehensive GC-MS contour plot. Black bubbles indicate 650 separated peaks. Analysis conditions – instrumentation: Agilent 6890N GC – Agilent 5975 MSD system (Agilent Little Falls, DE USA); sample volume: 2 µL of Vetiver essential oil diluted in cyclohexane (10 mg mL^{-1}). (a) One-dimensional fast-GC-MS analysis – column: SE54, l: 10 m, d_c: 0.10 mm, d_f: 0.10 µm; injection mode: split; ratio: 1 : 200, temperature: 270°C; carrier gas: helium; flow rate: 0.4 mL min^{-1}; temperature program 60°C (3 min)//8.2°C/min//260°C (3 min). (b) GCxGC-MS analysis – column set: 1D: SE54 (l: 30 m, d_c: 0.25 mm, d_f: 0.25 µm); 2D: OV1701 [l: 1 m, d_c: 0.10 mm, d_f: 0.10 µm (MEGA, Legnano (Milan), Italy]; thermal modulator: KT2004 loop, coolant liquid nitrogen (Zoex Corporation, Houston, TX, USA), modulation time: 5 s; carrier gas: helium; temperature program 60°C (1 min)//2.5°C/min//260°C (10 min). MSD conditions – ionization mode: EI 70 eV, transfer line temperature: 280°C; scan range *m/z* 35–250 in fast scanning mode (10,000 amu s^{-1}).

provide reliable sample discrimination and characterization or "identification" of diagnostic compounds, peaks, or even fragments (see in the following discussion), with which it is possible to differentiate between samples within a group. Moreover, a suitable GC profile can also offer information of a higher level, that is, information directly defining a (biological) characteristic of the matrix or plant sample(s) under investigation (see Section 2.3.1);

this includes correlation and/or prediction of variables or properties that are difficult to establish with other approaches, such as for instance volatile fraction *versus* metabolism, sensory characteristic, or biological activity. This paragraph reports on techniques applicable to a relatively large number of samples, ranging from a few dozen to a few hundred; other approaches are necessary for decidedly larger numbers of samples (e.g., neural network). GC or GC-MS profile(s) can be used in targeted

or untargeted methods, the former by selecting specific markers for the statistical processing, the latter operating on the whole chromatographic profile, upon which noise subtraction, peak alignment, and normalization are generally automatically carried out directly by the software. The total chromatographic profile (or MS profile in fingerprinting), expressed by the abundances of a selected number of markers (targeted) or by all peaks (untargeted) characterizing the vegetable matrix under investigation, after suitable treatment, can therefore be taken as a further distinctive parameter characterizing a sample, in view of its statistical processing, in particular for samples consisting of several uncorrelated components.

The most widely used approach is multivariate analysis, in particular principal component analysis (PCA), that is, an unsupervised method that can detect the differences between samples, within a set characterized by a suitable number of components (variables) through their linear combination, and can then concentrate the information on those that best explain the variability (Brereton, 2007). Not only can this method classify and differentiate group(s) of samples within a set, it can also identify those variables (components or MS fragments) with the highest statistical significance in sample(s) discrimination. In the above-mentioned context, supervised statistical methods can be applied as a complement to PCA. For instance, linear discriminant analysis (LDA) or soft-independent model of class analogy (SIMCA) are used when the information about classes is known *a priori* and the final goal is to build a model, which will then be used to classify new samples. Beside these tools, regression methods such as orthogonal partial least square analysis (OPLS) can be used to predict a characteristic of the matrix investigated from its GC-MS or MS profile or to correlate different characteristics of it. This can be useful to evaluate the intensity of association between the profile (or the fingerprint) of the volatile fraction of a plant and its origin or to detect the plant when it is used as such or in complex mixtures and/or any adulteration (Brereton, 2007; Berenson and Levine, 2002). Correct development of a classification and/or prediction model requires a significant number of samples that represent as widely as possible the sample variability, because it implies processing a training set of samples, to demonstrate a correlation between the two properties investigated (one being the matrix analytical profile) and a test set, to validate the proposed model.

Figure 4a reports the GC profile of a *Santalum album* essential oil from Sri Lanka together with (Figure 4b) the PCA biplot (loadings and scores plot) discrimination of a set of essential oil samples of other related species in particular on the first two principal components using profile components (for analysis conditions, see the figure caption). These results clearly show that profiling can successfully be used to detect possible sandalwood essential oil adulteration.

2.3 Fingerprinting of the Volatile Fraction

The possibility to use data processing with chemometrics as an active analysis tool has not only emphasized the role of data processing in the entire analytical procedure but also influenced analysis strategy and stimulated the introduction of new approaches, including fingerprinting. The term *fingerprint* generally applies to forensic science and is the "*impression of a fingertip on any surface … an ink impression of the lines upon the fingertip taken for the purpose of identification and/or something that identifies: as (a) a trait, trace, or characteristic revealing origin or responsibility; (b) analytical evidence (as a spectrogram) that characterizes an object or substance; in particular the chromatogram or electrophoretogram obtained by cleaving a protein by enzymatic action and subjecting the resulting collection of peptides to two-dimensional chromatography or electrophoresis*" (Merriam Webster Dictionary and Thesaurus). As mentioned earlier, fingerprinting in the chromatographic sense consists of rapid high throughput nonseparative screening, producing diagnostic profiles suitable to discriminate and classify samples. These methods are mainly carried out with HS-MS, direct infusion-MS, NMR, and FT-IR (MIR, NIR, etc.) combined with appropriate data mining. This paragraph discusses in detail HS-MS, also known as *mass-spectrometry-based electronic nose* (MS-EN) (Fenaille *et al.*, 2003; Pilar Marti, Busto, and Guasch, 2004; Pavón *et al.*, 2006; Karlshøj, Nielsen, and Larsen, 2007; Jelen, Ziolkowska, and Kaczmarek, 2010). HS and HCC-HS techniques, directly combined with non-separative MS methods, were introduced by Marsili in 1999 to study off-flavors in milk (Marsili, 1999).

HS-MS methods involve the direct nonseparative analysis of the volatile fraction of a sample to provide
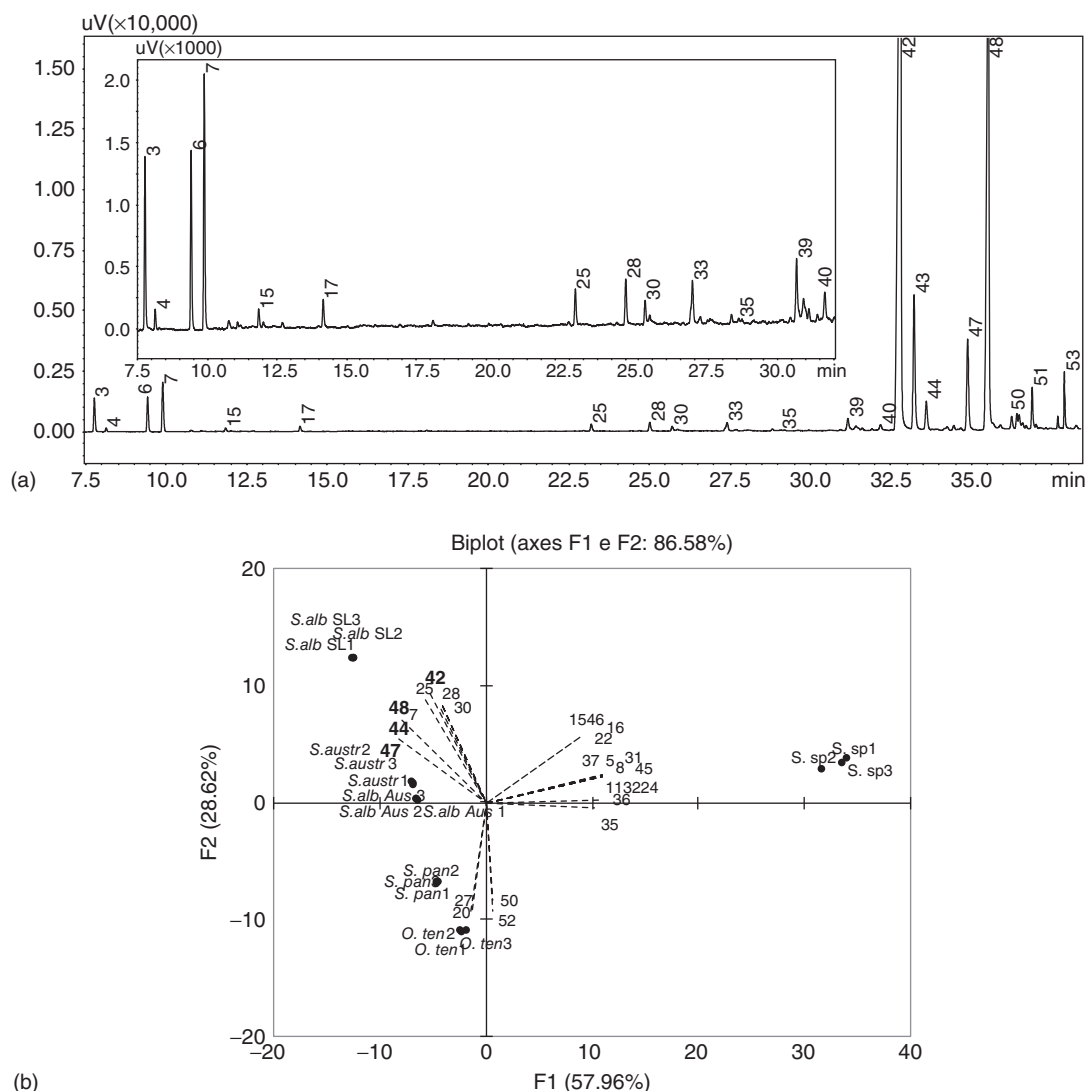
Figure 4 (a) GC-MS profile of a *Santalum album* L. essential oil from Sri Lanka. Analysis conditions – instrumentation: Shimadzu QP2010-PLUS GC-MS system/GC-MS Solution 2.51 software (Shimadzu Italia, Milan, Italy); column: SLB-IL60 ionic liquid stationary phase (*l*: 15 m, d_c: 0.10 mm, d_f: 0.08 μm) (Supelco, Bellafonte, PA, USA); injection mode: split; ratio: 1 : 20, temperature: 250°C; carrier gas: helium; flow rate: 0.4 mL min^{-1}; temperature program: 70°C//1.5°C/min//150°C//10°C/min//280°C; Marker identification: 42 *(Z)*-α-santalol, 43 *(Z)*-α-*trans*-bergamotol, 44: *(E)*-α-santalol; 47: epi-β-santalol; 48: *(Z)*-β-santalol; and 51: *(E)*-β-santalol. (b) PCA biplot (loadings' and scores' plots of different Santalaceae essential oil samples). Samples (legend): *Santalum album* L. Australia (S.alb Aus), *S. album* L. Sri Lanka (S.alb SL), *Santalum spicatum L.* (R.Br) A. DC. (S.sp.), *Santalum paniculatum* Hook. & Arn. (S.pan.), *Santalum austrocaledonicum* Vieill (S.austr.), and *Osyris tenuifolia* Hengler (O.ten.).

a representative and diagnostic MS fingerprint in which, in general, each *m/z* peak acts as a "sensor" whose intensity derives from the contribution of each compound producing that fragment. The resulting MS profiles can then be submitted to suitable chemometric processing to rapidly characterize

and discriminate samples within a set, to monitor biological phenomena when the volatile fraction can be taken as a marker, and/or to correlate samples with a technological process. These methods were immediately successful in the food field (Gardner and Philip, 1999; Marcos Lorenzo *et al.*, 2002)

in particular to characterizing several matrices but are now also of considerable interest in plant metabolic fingerprinting. Specific instrumentation provided with dedicated software has therefore been developed (Pirouette, 2010).The most widely applied HCC-HS technique in MS-EN is headspace solid-phase microextraction (HS-SPME) because of its flexibility, compatibility, its ability to be coupled on-line with MS, and ease of automation (Liberto *et al.*, 2013; Nicolotti *et al.*, 2013). MS-EN can also be used to monitor and quantify target compounds in a group of samples, provided that specific and diagnostic ions are obtained with a compatible ion generation mode (EI, CI, APCI, PTR, etc.) (Pérez Pavon *et al.*, 2004; Liberto *et al.*, 2013; Bicchi *et al.*, 2011). Mass spectral fingerprint or diagnostic ion abundance(s) can also be used as analytical decision maker (ADM) (Sandra, David, and Tienpont, 2004) for fast screening of a large number of samples and to decide which samples must be submitted to conventional separative analysis, because they do not produce unequivocal results. This approach, first developed for environmental analysis, significantly reduced the number of time-consuming conventional analyses and has been applied successfully in the authors' laboratory for aromatic and medicinal plant characterization (Ruosi *et al.*, 2012).

Compared to a separative GC-MS profile, MS-EN patterns provide a fast response, but the resulting TIC is not particularly significant, and it appears that less information can be obtained from the MS profile. The nature of the signals making up the fingerprint (i.e., fragments (m/z) whose intensity is the result of the contributions of all sample components generating that ion in its ionization process) makes it mandatory to adopt chemometrics "to extract" data from the MS profile that can give significant information. However, fingerprint precision over a long period is one of the main limits when a mathematical model is generated and used to classify or correlate sample composition or characteristics and the fingerprint's resources of hidden information. In HS-SPME-MS, this variability is mainly due to SPME fiber performance over time and to MS signal instability, due in turn to ion source contamination, aging of electron multiplier, and/or filament electron emission. These limits can be overcome by (i) constantly monitoring SPME fiber performance (Bicchi *et al.*, 2007), (ii) checking electron multiplier response over time, and (iii) standardizing the MS profile through internal normalization.

Figure 5 reports the HS-SPME-GC-MS (a) and HS-SPME-TIC-MS (b) profiles together with mass spectral fingerprint (c) of a thyme sample (*Thymus vulgaris* L.), whereas Figure 6 reports PCA score plot (a) and loadings plot (b) on the first two principal components using the significant ions in the MS profiles to discriminate between samples belonging to different chemotypes; samples 25–27 contain carvacrol in unusually high amounts. The loading plot clearly shows that the discriminating peaks are those characteristic of thymol and carvacrol (m/z 91, 135, and 150) (for analysis conditions and peak identification, see the figure caption).

2.3.1 Fingerprinting Approaches for GCxGC-MS Data Processing and Interpretation

This paragraph is dedicated to fingerprinting applied to multidimensional chromatographic techniques (in particular to GCxGC) because of these methods' increasing success, mainly due to the huge amount of information they can provide. This has stimulated the development of new approaches for data mining, capable of exploiting the full informative potential of these techniques. Two-dimensional (2D) separation patterns have an intrinsic potential for sample fingerprinting, making the diagnostic distribution of the analyte across the chromatographic plane highly informative for sample characterization, differentiation, discrimination, and classification and the technique analogous to biometric fingerprinting processes.

Two-dimensional separation patterns are highly diagnostic, not only because of the well-known high practical GCxGC peak capacity, sensitivity, and specificity, but also because of chemically related groups of substances give useful structure retention patterns, affording profitable group characterization of samples.

The improved informative content, however, produces large and complex data sets, consisting of bidimensional retention data, detector responses, and MS spectra; this data requires suitable data mining processes (i) to extract useful and consistent data on sample compositional characteristics and (ii) to make evident the higher level of information "hidden" in them.

Different approaches have been attempted to link separation data to samples' compositional characteristics, and their effectiveness has been demonstrated
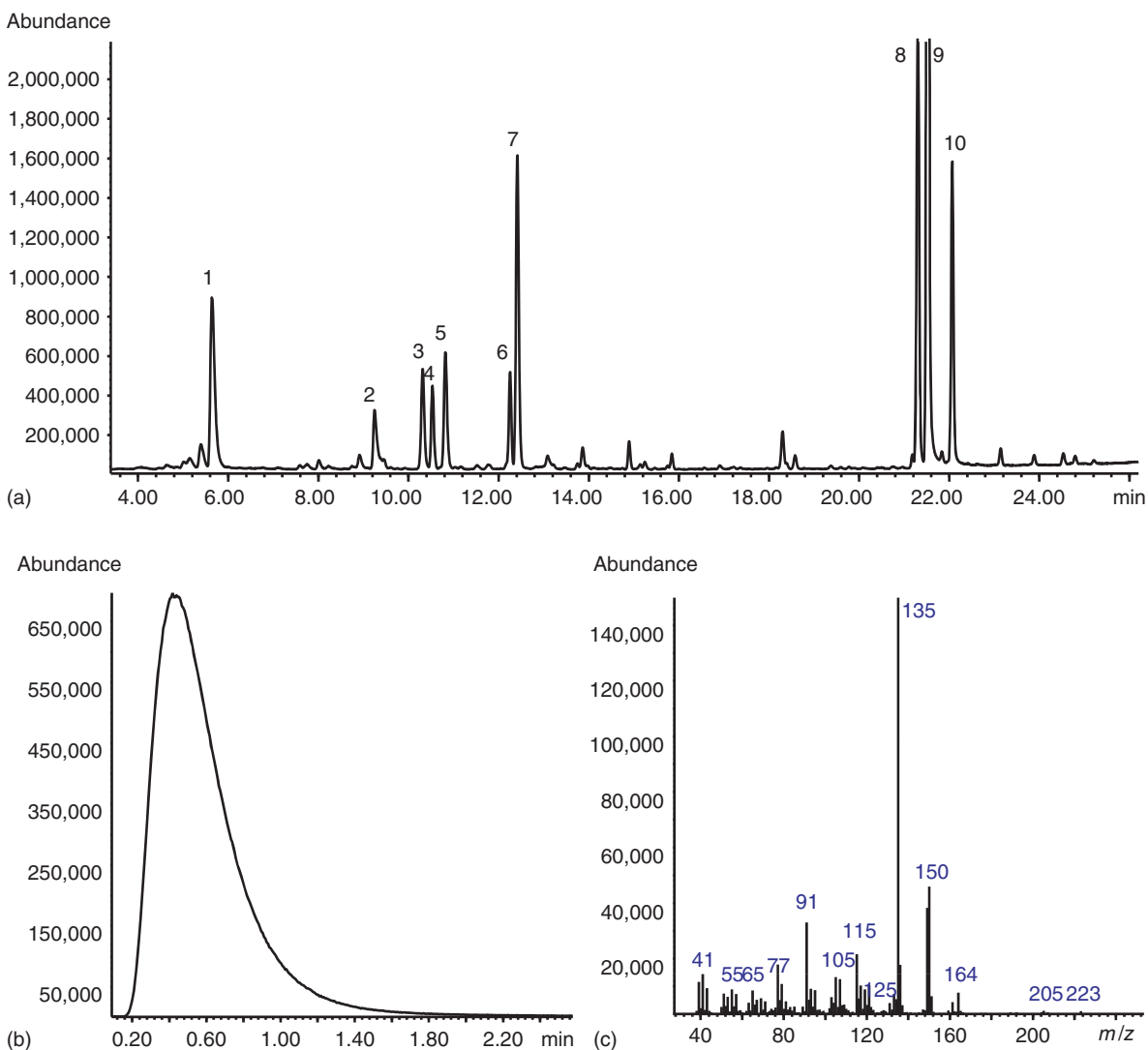
Figure 5   (a) HS-SPME-GC-MS, (b) HS-SPME-TIC-MS, and (c) corresponding spectral fingerprint of an Italian thyme sample (*Thymus vulgaris* L.). Analysis conditions – instrumentation: Agilent 6890 GC-5973 MSD system (Agilent, Little Falls, DE, USA). Sample preparation – SPME fiber: PDMS, 1 cm; sample amount: 10 mg; vial volume: 20 mL; sampling time: 30 min; sampling temperature: 50°C; desorption temperature: 250°C; desorption time: 5 min. GC-MS conditions – column Megawax 20M (*l*: 50 m, d_c: 0.20 mm, d_f: 0.20μm) [MEGA, Legnano (Milan) Italy); injection mode: split; ratio: 1 : 5; temperature: 250°C; carrier gas: helium; flow rate: 1.0 mL min^{-1}; temperature program: 50°C (1 min)//3°C/min//250°C (5 min). MS analysis conditions – transfer column: deactivated fused silica tubing (*l*: 6.70 m, d_c: 0.10 mm); injection mode: split; ratio: 1 : 5, temperature: 250°C; carrier gas: helium; flow rate:.0.4 mL min^{-1}; oven temperature: 250°C. Component identification: (1) *p*-cymene, (2) linalool, (3) thymol methyl ether, (4) carvacrol methyl ether, (5) β-caryophyllene, (6) α-terpineol, (7) borneol, (8) eugenol, (9) thymol, and (10) carvacrol.

in several application fields (Pierce *et al*., 2008; Vial *et al*., 2009; Schmarr and Bernhardt, 2010; Hoggard *et al*., 2010; Wang *et al*., 2010). GCxGC data handling approaches are also commonly classified

into the *targeted* and *nontargeted* methods (Pierce *et al*., 2008). In GCxGC, nontargeted methods are often based on chemometrics or on image processing procedures and must be able to exploit
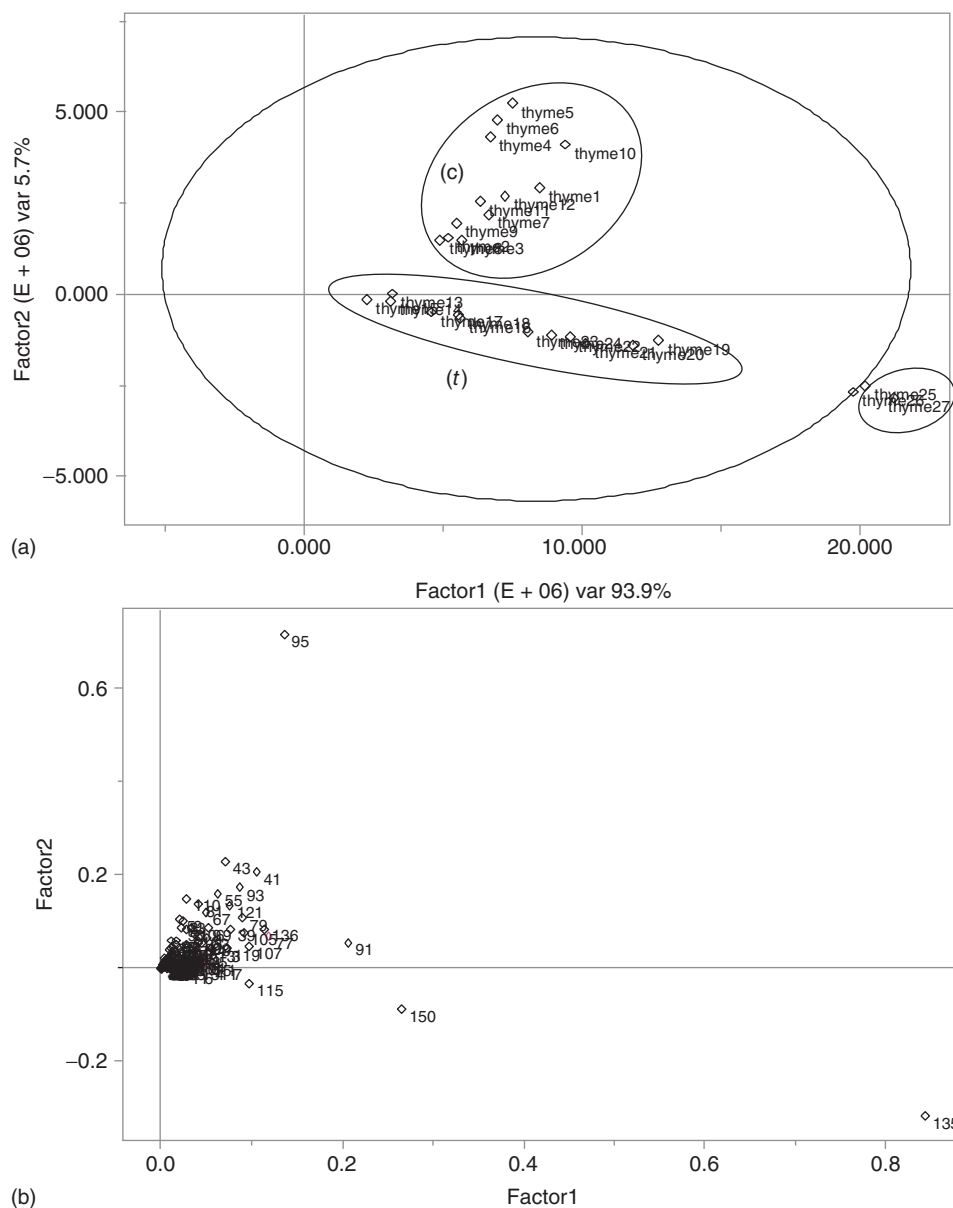
Figure 6 Discrimination of 27 samples of thyme (*Thymus vulgaris* L.) belonging to two different chemotypes, thymol (t) and carvacrol (c). (a) PCA score plot and (b) loadings plot of the first two principal components using significant ions from the MS profiles. For sampling and analysis conditions, see captions of Figure 5.

the multidimensionality of a separation in full, including the fragmentation patterns, if MS detection is involved.

Different approaches to classify and characterize GCxGC results have been studied in the authors' laboratory (Cordero, Bicchi, and Rubiolo, 2008; Cordero *et al.*, 2010a, 2010b; Kiefl *et al.*, 2012; Reichenbach *et al.*, 2012; Cordero *et al.*, 2012b), in particular by developing *nontargeted* methods aimed at finding compositional differences among samples in their whole complexity, that is, through their multidimensional profile (fingerprint), regardless

of their chemical composition. A new, effective, specific, and reliable *nontargeted* analysis approach for complex samples was developed (Cordero *et al.*, 2010a, 2010b) in the authors' laboratory, in collaboration with other research groups, for comparative evaluation of two-dimensional chromatographic data of complex aroma samples from roasted hazelnuts (*Corylus avellana* L.) and coffee and from plant volatile fraction (e.g., *Juniperus communis*). The approach, called *comprehensive template-based fingerprinting*, was inspired by biometric fingerprinting and does not rely on sample chemical speciation but on the information provided by the GCxGC separation *in toto* (and, in this case, taken as a sample characterizing parameter). As the GCxGC separation pattern comprises a number of 2D peaks spread over the 2D chromatographic plane, each peak (generally corresponding to a single compound) can be treated as a separate minutia for comparative pattern analysis, as for individual fingertip features. In biometry, minutiae are local ridge characteristics that occur at either a ridge bifurcation or a ridge ending and features are qualitative differences within samples pattern or part of them. Figure 7 reports the HS-SPME GCxGC contour plots of (a) volatiles from *Mentha* x *piperita* L. leaf and (b) in the herbivore mint bug (*Chrysolina herbacea*) feces on feeding. Fingerprinting by advanced template matching revealed several known (2-α-hydroxy-, 3-α-hydroxy-, 3-β-hydroxy-, and 9-hydroxy-1,8-cineole) and unknown oxidized monoterpenes resulting from the multitrophic interaction. White bubbles indicate 2D peaks corresponding to identified oxidation products arising from *Chrysolina* metabolism (Cordero *et al.*, 2012b).
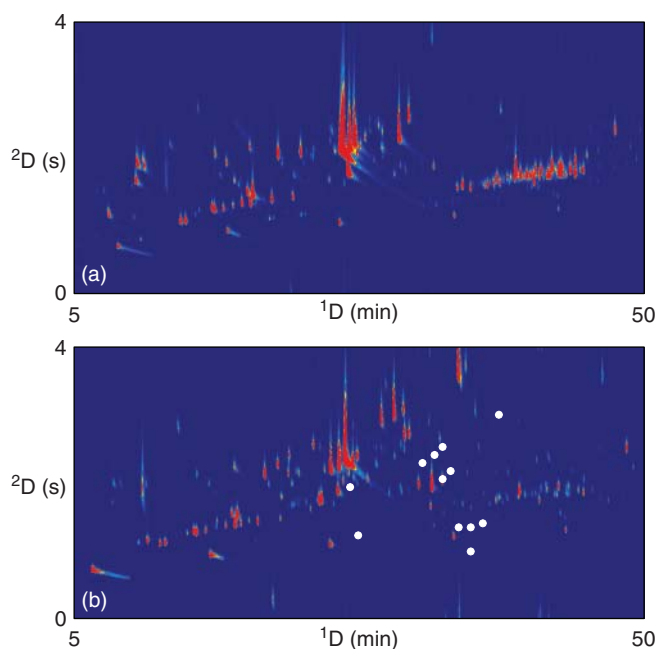
Figure 7   GCxGC contour plots of (a) volatiles from *Mentha x piperita* L. leaf and (b) in the herbivore mint bug feces (*Chrysolina herbacea*) on feeding. Analysis conditions: intact leaves (about 50 mg) from living plants of *M. x piperita* L. were carefully picked-off and immediately placed in 2.0 mL vials and hermetically sealed for HS-SPME sampling. Feces fluid (3–5) mg collected from insects feeding on *M. x piperita* L. were transferred in a 2.0-mL headspace vials, exactly weighted and hermetically sealed for HS-SPME sampling. HS-SPME sampling conditions: fiber: 2 cm Stableflex 50 : 30 μm DVB-Carboxen-PDMS (Supelco, Bellafonte, USA); sampling time: 40 min, temperature: 30°C. GCxGC analyses: instrumentation: GCxGC-MS system: Agilent 6890 GC – Agilent 5975 MSD. GCxGC interface: KT 2004 loop modulator (Zoex Corporation, Houston, TX, USA), modulation time: 4 s. Column set: 1D: SE54 column (30 m, d_c: 0.25 mm, d_f: 0.25 μm), 2D OV1701 column (1 m, d_c: 0.1 mm, d_f: 0.10 μm) [MEGA – Legnano (Milan)-Italy]. Analysis conditions: injection mode: split; ratio: 1 : 10, temperature: 250°C; carrier gas: helium; temperature program: 45°C (1 min)//2.5°C/min//260°C (10 min). MSD conditions: ionization mode: EI 70 eV (Agilent, Little Falls, DE, USA); transfer line temperature: 280°C, scan range: *m/z* 35–250 in fast scanning mode (10,000 amus^{-1}).

Briefly, the approach follows this logic path:

(a) 2D peaks from a *source* chromatogram are detected and a template created, by recording retention times, detector responses, and MS fragmentation patterns;

(b) the template is compared to the peak pattern of the next sample (*analyzed* chromatogram) in the set to establish correspondences between peaks of the same analyte;

(c) if necessary, the template is transformed (or aligned), in the retention times plane, to compensate for retention time shifts;

(d) positive matches are constrained by evaluating retention time coherence and mass spectral match factors (NIST algorithm similarity or identity).

(e) peaks in the *analyzed* chromatogram without correspondences are then added to the template and the "updated" template adopted for the successive comparative steps.

This series of operations is applied to the entire set of sample chromatograms to generate a *consensus template* of nontargeted peaks useful for the cross comparison of samples. *Comprehensive template-based fingerprinting* can be used for classification purposes, based on pattern similarity, or for pairwise comparisons to reveal qualitative and quantitative differences in sample chemical composition between pairs.

3 SCREENING AND EXHAUSTIVE STUDIES

Screening and exhaustive analysis are two highly complementary approaches involved in the study of the qualitative and quantitative compositions of a plant volatile fraction.

Screening can be carried out on an essential oil analyzed off-line by GC-MS, or with TAS systems usually consisting of one HCC-HS sampling technique connected on-line to a GC-MS system operating automatically. Component identification, or at least characterization, is preferentially carried out automatically, by an on-line or off-line combination of GC and MS data, by comparison with I^Ts and MS spectral data stored in database(s) included in the GC-MS software (Costa *et al*., 2007; FFNSC MS Library; NIST 05 – NIST/EPA/NIH Mass Spectral Library; AMDIS, 2007).

As mentioned earlier, for many applications, identification of a component is not necessary; it has only to be correctly located in the chromatogram through its retention index and MS spectrum. Screening can also be carried out with other approaches, all based on fingerprinting with nonseparative MS methods, or with spectroscopic techniques such as NMR or NIR combined with statistical data processing (see above). The increased NMR field frequency (with the consequent increased instrument resolution) has also been successfully exploited not only for screening but also for component identification by selective spectral matching between the spectral profile of the total sample and the spectra of component standards stored in a dedicated database. This approach was first investigated in-depth for essential oils by Formacek and Kubeczka (2002) in the 1980s and then by Casanova and Tomi with ^{13}C NMR spectroscopy (Tomi *et al*., 1995; Tomi and Casanova, 2007).

When the volatile fraction of a plant contains unknown, or apparently unknown, component(s), their identification requires an exhaustive strategy to be put in place. Unfortunately, only a few research groups still run such studies, which underlie phytochemistry and which were very common until 10–15 years ago. Several reasons have produced this change of attitude, among others: (i) the increasing difficulty of finding unknown odorous plants of interest for industrial applications; (ii) the ever stricter regulations affecting the biological activity of cosmetic and food ingredients (mainly toxicity) that leads to the tendency to use well-known chemical compounds; (iii) the spread of green chemistry approaches, which limit the consumption of reagents and solvents; (iv) the increase in controls aimed at avoiding biopiracy, which limit the amount of plant to be processed; and (v) the high potential of analytical instrumentation and informatics, which reduces the need for complex and time-consuming structure elucidation procedures of unknown compounds, unless they are closely related to the biological phenomenon investigated.

Exhaustive studies require two main steps. The first consists of identifying as many components as possible, in general through direct GC-MS analysis of the volatile fraction of the plant investigated, exactly as for screening, by interactively combining retention index and mass spectral libraries on-line or off-line. This step is fundamental to detect unknown component(s) to be identified and locate them in the total chromatogram. The effectiveness of this step is destined to increase as the recent and powerful GC-MS

systems are routinely adopted; these have MS detectors based on high resolution IT or TOF analyzers, either as such or in a hybrid combination (q-TOF or IT-TOF). They can provide molecular formulae and, through dedicated software, can propose hypotheses of structures, thus dramatically increasing the possibility of instrumental component identification (Kind and Fiehn, 2007). The second step aims to identify minor or marker components not identifiable in the screening step, through the conventional approach by isolating them in pure or enriched form, through fractionation, and then elucidating or confirming their structures, via spectroscopic techniques (mainly NMR, MS, and IR).

Exhaustive studies also include the identification of compounds responsible for biological phenomena involved in the plant volatile fraction. An approach first developed to screen the pharmacological activity of plant extracts, but that can successfully be transferred to increase the effectiveness of an exhaustive study, is the bioguided strategy (Pieters and Vlietinck, 2005). In this approach to study the chemical composition of a distillate, an extract, or a fraction, the sequence of operative steps is chosen from the results of the biological test of interest (i.e., sensory properties and plant–plant or plant–insect interaction), thus enabling attention to be focused on the fractions or components responsible for the activity itself. This strategy is complementary to dereplication (Lang *et al.*, 2008).

The first mandatory condition in this connection is the availability of a suitable sample amount (at least 10 mL) in a liquid form; in general, the sample therefore consists of (i) most often, the essential oil of the plant to be investigated; (ii) the volatile fraction, obtained by normal or vacuum distillation, or by steam- or hydro-distillation, from a plant extract obtained in turn through different techniques [solvent extraction assisted (or not) by ultrasound, microwaves, pressurized solvents, SFE, etc.]; (iii) less frequently, it consists of samples resulting from preparative purge and cryotrapping HS systems; and (iv) very seldom, of samples resulting from preparative purge combined with adsorption or sorption on specific materials, from which they are recovered by solvent elution.

Figure 8 shows a model process for a bioassay-oriented exhaustive study of the volatile fraction. The strategy that will be described in the following text is very general because the operative sequences depend on the composition of the volatile fraction resulting from the initial screening step. In spite of the very considerable technological evolution, and the introduction of several new technologies, exhaustive studies of the volatile fraction are still based on conventional protocols. In general, the volatile fraction is mostly lipophilic, thus affording a first rough fractionation based on simple chemical or physical properties, for instance, the separation of acidic, neutral, or basic compounds by liquid–liquid extraction, or through solutions of increasing basicity to separate carboxylic acid or volatile phenols, or with acidic solution to separate volatile alkaloids. This operation is useful not only to separate basic and acid compounds but also to restrict the range of polarity of the so-called neutral fraction (in general, the most abundant in an essential oil). The next step implies two operations (vacuum distillation or column chromatography) whose order depends on the composition of the sample investigated. One of the advantages of plant volatile fractions, compared to other fractions of a plant extract, is that differences in component volatility can be exploited for an effective fractionation. Differently volatile components can successfully be separated by a fractionated distillation under vacuum to avoid artifact formation and thermolabile compound degradation. For instance, monoterpenoids can easily be separated from sesquiterpenoids and, among them, hydrocarbons from oxygenated compounds, depending on both their relative volatility and the efficiency of the distillation system. The result is also a highly simplified residue of components undistilled under the conditions adopted. Sophisticated systems to control vacuum and temperature, and providing a high number of theoretical plates, are now available [e.g., a semimicro distillation unit provided with the Fischer® Spaltrohr® column (Fischer HMS 500)] to provide simplified fractions or, in some cases, even pure compounds. Similar considerations can be made for preparative liquid chromatographic fractionation: column chromatography, flash chromatography, low and medium pressure liquid chromatography, preparative thin layer chromatography, or solid-phase extraction are generally used as such, or in various combinations, depending on sample complexity and amount. Preparative liquid chromatography is in general carried out in the normal-phase mode because of the chemical nature of the volatile fraction components, that is, with silica as stationary phase and eluents of increasing polarity as mobile
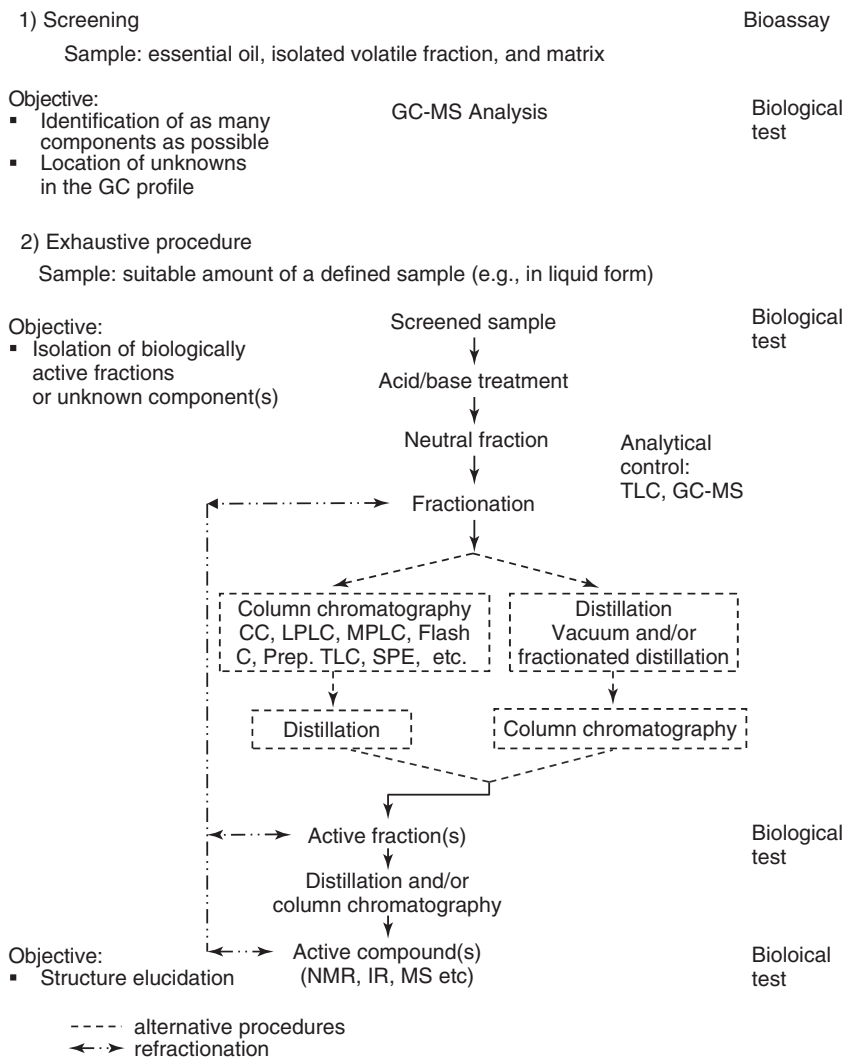
Figure 8 Diagram of a process for a bioassay-oriented exhaustive study of the volatile fraction. C, chromatography.

phase. These operations must usually be repeated, varying stationary or eluting phase selectivity, until a fraction or fractions, selectively enriched with the component(s) to be identified, is obtained in sufficient amounts (about 1 mg) to provide unequivocal MS, IR, and NMR spectra and, if possible, to apply two-dimensional NMR techniques, so as to obtain sufficient reliable information for correct structure elucidation. The final confirmation of the new structure can only legitimately be obtained by comparison

with the synthesized standard. The success of the fractionation/isolation procedure can be monitored either by simple TLC control or by analyzing (e.g., by GC-MS) each fraction after each fractionation step. Flash chromatography has evolved considerably over recent few years, in terms of new stationary phases and of automation and detection methods: in particular, on-line combinations with PDA, evaporative light scattering detectors (ELSDs), and very recently MS are now commercially available (Isolera™

Dalton mass-directed flash chromatography, 2013; www.biotage.com). Liquid chromatography and distillation can be combined complementarily in different orders, or applied several times, depending on the polarity, volatility, and analyte abundances of the fraction(s) to be processed.

An interesting and relatively little considered method for exhaustive studies of the volatile fraction is preparative GC. Over recent 10–15 years, this technique has regained interest because of the introduction of technological improvements. Two approaches are now available: (i) systems operating with packed GC columns of inner diameter that is ideally a good compromise between efficiency and column loading, so as to obtain sufficient separation of the fraction from which enriched or pure component(s) are to be isolated and, at the same time, to obtain sufficient of it (them) in a few chromatographic runs. This system operates through a bypass at the column exit, automatically actuated on a time-window basis, which directs the flow at the column exit to one or more cryotraps where the analytes of interest are condensed; the process can also fully be automated, provided that constant automatic peak alignment control is operated; (ii) systems based on capillary columns (possibly coated with a relatively thick film of stationary phase) provided with sophisticated systems of peak alignment and time-window actuators that direct the carrier gas eluting from the chromatographic column to a number of cold or cryotraps, in order to accumulate appropriate analyte amounts for further spectroscopic analysis (NMR, IR, and MS) through a large number of highly repeatable automatic injections. These systems have successfully been used to isolate analytes for structure elucidation; pure enantiomer(s) to establish *ee* or *er* correctly, so as to evaluate the genuineness of an essential oil; or pure analytes (enantiomers) for sensory evaluation. This approach is technologically more complex and expensive than the packed columns systems and requires rigorous control of the chromatographic pattern so as to avoid cross contamination of the analytes in the cryotraps (http://www.gerstel.com; Sciarrone *et al*., 2012).

4   CONCLUSIONS

The increasing importance of the volatile fraction in plant biology and applications has changed analysis strategies to a remarkable extent. The approaches introduced with metabolomics have been successful not only because of the new technologies (hybrids TOF-MS and IT-MS detectors, dedicated software, highly effective GC techniques, etc.) but also because of the increasingly integrated and harmonized development of the three main steps characteristic of an analytical procedure. This has involved sample preparation techniques especially designed for on-line coupling to GC, in its turn combined with fast and/or high resolution MS detectors, so as to provide data that is increasingly diagnostic and in amounts sufficient to be transformed into information at a higher level, through automatic and appropriate data mining techniques. This characteristic is essential for the high throughput required in metabolomics and, at the same time, to achieve results that are suitable for successful comparison of fingerprinting with data mining methods and to pinpoint biomarkers. The present and future trends in studies on the role and composition of the plant volatile fraction all aim at fully exploiting the potential of the approaches and techniques that are now available, whose limits are not yet fully known, in the attempt to automate existing procedures, increasing data quality and improving processing, and developing new tools for data mining. This trend will also quickly modify the role of the operator, extending the demands made on his or her expertise from knowledge of a specific technology to the ability to interpret biological phenomena: thus his or her knowledge of analytical chemistry will be integrated with that relating to plant biochemistry, phytochemistry, and data mining, so that he or she can contribute to "reading" the huge amount of data obtainable with the new technologies, and transforming it into results of biological interest.

REFERENCES

AMDIS (2007) http://chemdata.nist.gov/mass-spc/amdis/, Version 2.65. (accessed January 2014).

Belitz, H. D., Grosch, W., and Schieberle, P. (2009) *Food Chemistry*, 4th edn, Springer-Verlag, Berlin Heidelberg, vol. **5**, pp. 340–400.

Chapters 13-17, Berenson, M. L., and Levine, D. M. (2002) *Basic Business Statistics: Concepts and Applications*, 8th edn, Prentice Hall, USA, pp. 487–718.

Bicchi, C., and Maffei, M. (2012) The plant volatilome: methods of analysis, in *High-Throughput Phenotyping in Plants: Methods and Protocols – Series: Methods in Molecular Biology*, Humana PressVol, New York, vol. **918**, pp. 289–310.

Bicchi, C., Cordero, C., Liberto, E., *et al.* (2007) *J. Chromatogr. A*, **1152**, 138–149.

Bicchi, C., Blumberg, L., Cagliero, C., *et al.* (2010a) *J. Chromatogr. A*, **1217**, 1530–1536.

Bicchi C, Cagliero C, Cordero C, Liberto E, Sgorbini B, and Rubiolo P. 2010b. New trends in the analysis of the volatile fraction of matrices of vegetable origin. In *Proceedings of the 9th Wartburg Symposium on Flavor Chemistry & Biology* - April 13-16, 2010. P. Schieberle, T. Hofmann, W. Meyerhof (Eds). Eisenach, Germany.

Bicchi, C., Ruosi, M. R., Cagliero, C., *et al.* (2011) *J. Chromatogr. A*, **1218**, 753–762.

Blumberg, L. M., and Klee, M. S. (1998) *Anal. Chem.*, **70**, 3828–3839.

Brereton, R. G. (2007) *Applied Chemometrics for Scientists*, John Wiley & Sons Ltd.Chapter V, Chichester, pp. 147–167.

Chemat, F., Bergamelli, F., and Visinoni, F. (2008) *Planta Med.*, **74**, 917–918.

Chemat, F., Zill-e-Huma, K., and Kamran, M. (2011) *Ultrason. Sonochem.*, **18**, 813–835.

Clevenger, J. F. (1928a) *J. Am. Pharm. Assoc.*, **17**, 346–349.

Clevenger, J. F. (1928b) *J. Assoc. Off. Agric. Chem*, **11**, 332–335.

Cordero, C., Bicchi, C., Joulain, D., *et al.* (2007) *J. Chromatogr. A*, **1150**, 37–49.

Cordero, C., Bicchi, C., and Rubiolo, P. (2008) *J. Agric. Food Chem.*, **56**, 7655–7666.

Cordero, C., Liberto, E., Bicchi, C., *et al.* (2010a) *J. Chromatogr. Sci.*, **48**, 251–261.

Cordero, C., Liberto, E., Bicchi, C., *et al.* (2010b) *J. Chromatogr. A*, **1217**, 5848–5858.

ChapterCordero, C., Liberto, E., Sgorbini, B., *et al.* (2012a) Gas chromatography, in *Chemical Analysis of Food: Techniques and Applications*, ed. Y. Pico, Academic Press, San Diego, CA, vol. **11**, pp. 311–374.

Cordero, C., Zebelo, S. A., Gnavi, G., *et al.* (2012b) *Anal. Bioanal. Chem.*, **402**, 1941–52.

Costa, R., De Fina, M., Valentino, M. R., *et al.* (2007) *Nat. Prod. Commun.*, **2**, 413–418.

D'Acampora Zellner, B. D., Bicchi, C., Dugo, P., *et al.* (2008) *Flavour Fragr. J.*, **23**, 297–314.

Debonneville, C., and Chaintreau, A. (2004) *J. Chromatogr. A*, **1027**, 109–115.

Dittrich, P. S., Tachikawa, K., and Manz, A. (2006) *Anal. Chem.*, **78**, 3887–3908.

Dugo, G., Dugo, P., and Mondello, L. (2002) Multidimensional chromatography: foods, flavours and fragrances applications, in *Multidimensional Chromatography*, eds. L. Mondello, A. C. Lewis and K. D. Bartle, John Wiley & Sons Ltd, Chichester pp. 215–249.

European Pharmacopoeia Online (2014) 8th Edition (8.0) http://online.edqm.eu/EN/entry.htm (accessed January 2014).

Fenaille, F., Visani, P., Fumeaux, R., *et al.* (2003) *J. Agric. Food Chem.*, **51**, 2790–2796.

Ferhat, M. A., Meklati, B. Y., and Chemat, F. (2007) *Drug Future*, **32**(Suppl.: A), 89–90.

FFNSC 2.0 – Flavors and Fragrances of Natural and Synthetic Compounds – Mass Spectral Database (2011) Chromaleont, Messina, Italy.

Fiehn, O., Kopka, J., Dormann, P., *et al.* (2000) *Nat. Biotechnol.*, **18**, 1157–1161.

Formacek, V., and Kubeczka, K. H. (2002) *Essential Oils Analysis by Capillary Gas Chromatography and Carbon 13-NMR Spectroscopy*, Chichester, John Wiley & Sons Ltd.

Fornari, T., Vicente, G., Vazquez, E., *et al.* (2012) *J. Chromatogr. A*, **1250**, 34–48.

Gardner, J. W., and Philip, P. N. (1999) *In* Electronic Noses: Principles and Applications, Oxford University Press, New York.

Giarrocco, V., Quimby, B., and Klee, M. (1997) *Retention Time Locking: Concepts and Application;* Gas Chromatography Application Note, Agilent technologies Inc., Santa Clara CA, USA.

Glassbrook, N., and Ryals, J. (2001) *Curr. Opin. Plant Biol.*, **4**, 186–190.

Halket, J. M., Waterman, D., Przyborowska, A. M., *et al.* (2005) *J. Exp. Bot.*, **56**, 219–243.

Harrigan, G. G., and Goodacre, R. (2003) *Metabolic profiling: its role in biomarker discovery and gene function analysis*, Kluwer Academic Publishers, Boston.

Herrero, M., Mendiola, J. A., Cifuentes, A., *et al.* (2010) *J. Chromatogr. A*, **1217**, 2495–2511.

Hoggard, J. C., Wahl, J. H., Synovec, R. E., *et al.* (2010) *Anal. Chem.*, **82**, 689–698.

Jelen, H. H., Ziolkowska, A., and Kaczmarek, A. (2010) *J. Agric. Food Chem.*, **58**, 12585–12591.

Karlshøj, K., Nielsen, P. V., and Larsen, T. O. (2007) *J. Agric. Food Chem.*, **55**, 4289–4298.

Kiefl, J., Cordero, C., Nicolotti, L., *et al.* (2012) *J. Chromatogr. A*, **1243**, 81–90.

Kind, T., and Fiehn, O. (2007) *BMC Bioinformatics*, **8**, 105.

Klee, M. S., and Blumberg, L. M. (2002) *J. Chromatogr. Sci.*, **40**, 234–247.

Koesukwiwat, U., Lehotay, S. J., Miao, S., *et al.* (2010) *J. Chromatogr. A*, **1217**, 6692–6703.

Kolb, B., and Ettre, L. S. (1997) *Static Headspace-Gas Chromatography, Theory and Practice*, Wiley-VCH, New York.

Lang, G., Mayhudin, N. A., Mitova, M. I., *et al.* (2008) *J. Nat. Prod.*, **71**, 1595–1599.

Liberto, E., Cagliero, C., Sgorbini, B., *et al.* (2008) *J. Chromatogr. A*, **1195**, 117–126.

Liberto, E., Ruosi, M. R., Cordero, C., *et al.* (2013) *J. Agric. Food Chem.*, **61**, 1652–1660.

Maffei, M. E., Gertsch, J., and Appendino, G. (2011) *Nat. Prod. Rep.*, **28**, 1359–1380.

Manz, A., Graber, N., and Widmer, H. M. (1990) *Sensor Actuat. B-Chem.*, **1**, 244–248.

Marcos Lorenzo, I., Pérez Pavón, J. L., Fernandez Laespada, M. E., *et al.* (2002) *J. Chromatogr. A*, **945**, 221–30.

Marsili, R. T. (1999) *J Agric. Food Chem.*, **47**, 648–654.

Mason, T. J., Chemat, F., and Vinatoru, M. (2011) *Curr. Org. Chem.*, **15**, 237–247.

Nicolotti, L., Cordero, C., Bicchi, C., *et al*. (2013) *Food Chem.*, **138**, 1723–33.

NIST 05 – NIST/EPA/NIH Mass Spectral Library. (2005) (National Institute of Standards and Technology, Gaithersburg, MD, U.S.A.).

Oliver, S. G., Winson, M. K., Kell, D. B., *et al*. (1998) *Trends Biotechnol.*, **16**, 373–378.

Pavón, J. L. P., Sanchez, M., Nogal, D., *et al*. (2006) *TRAC: Trends Anal. Chem.*, **25**, 257–266.

Pérez Pavon, J. L., Del Nogal, S. M., Garcia Pinto, C., *et al*. (2004) *J. Chromatogr. A*, **1048**, 133–139.

Pierce, M., Hoggard, J. C., Mohler, R. E., *et al*. (2008) *J. Chromatogr. A*, **1184**, 341–352.

Pieters, L. and Vlietinck, A. J. (2005) *J. Ethnopharmacol.*, **100**, 57–60.

Pilar Marti, M., Busto, O., and Guasch, J. (2004) *J. Chromatogr. A*, **1057**, 211–217.

Pirouette (2010) http://www.infometrix.com/software/pirouette.html. Infometrix Inc. Bothell, WA 98011-3759. (accessed January 2014).

Purcaro, G., Tranchida, P. Q., Ragonese, C., *et al*. (2010) *Anal. Chem.*, **82**, 8583–8590.

Reichenbach, S. E., Tian, X., Cordero, C., *et al*. (2012) *J. Chromatogr. A*, **1226**, 140–148.

Rubiolo, P., Liberto, E., Sgorbini, B., *et al*. (2008) *J. Sep. Sci.*, **31**, 1074–1084.

Rubiolo, P., Sgorbini, B., Liberto, E., *et al*. (2010a) Analysis of the plant volatile fraction, in *The Chemistry and Biology of Volatiles*, ed. A. Herrmann, Wiley, Chichester, pp. 49–93.

Rubiolo, P., Sgorbini, B., Liberto, E., *et al*. (2010b) *Flavour Fragr. J.*, **25**, 282–290.

Rubiolo, P., Casetta, C., Cagliero, C., *et al*. (2013) *Anal. Bioanal. Chem.*, **405**, 1223–1235.

Ruosi MR, Bicchi C, Cagliero C, Cordero C, Griglione A, Liberto E, Rubiolo P, Sgorbini B. (2012) Quantitative analysis

of vegetable solid matrices by multiple headspace solid phase microextraction: determination of biologically active compounds in spices. In *36th International Symposium on Capillary Chromatography and 9th GCxGC Symposium*. Mondello, L. and Sandra, P. (Eds.). Riva del Garda, Italy; C24.

Sahraoui, N., Vian Abert, M., Bornard, I., *et al*. (2008) *J. Chromatogr. A*, **1210**, 229–233.

Sandra, P., David, F., and Tienpont, B. (2004) *Chromatographia*, **60**, S299–S302.

Schmarr, H. G., and Bernhardt, J. (2010) *J. Chromatogr. A*, **1217**, 565–574.

Sparkman, O. D., Penton, Z., and Kitson, F. G. (2011) *Gas Chromatography and Mass Spectrometry: A Practical Guide*, 2nd edn, San Diego, Academic Press.

Sumner, L. W., Mendes, P., and Dixon, R. A. (2003) *Phytochemistry*, **62**, 817–836.

Tomi, F., and Casanova, J. (2007) *Acta Horticult*, **723**, 185–192.

Tomi, F., Bradesi, P., Bighelli, A., *et al*. (1995) *J. Magn. Reson. Anal.*, **1**, 25–34.

Tranchida, P. Q., Mondello, L., Poynter, S. D. H., *et al*. (2011) Comprehensive two-dimensional gas chromatography combined with mass spectrometry, in *Comprehensive Chromatography Combination with Mass Spectrometry*, ed. L. Mondello, John Wiley & Sons Ltd Chapter, Chichester, vol. **6**, pp. 171–242.

Vial, J., Nocairi, H., Sassiat, P., *et al*. (2009) *J. Chromatogr. A*, **1216**, 2866–2872.

Wang, B., Fang, A., Heim, J., *et al*. (2010) *Anal. Chem.*, **82**, 5069–5081.

Watson, J. T., Schultz, G. A., Tecklenburg, J. R. E., *et al*. (1990) *J. Chromatogr. A*, **518**, 283–295.

www.biotage.com (2013) Isolera™ Dalton mass-directed flash chromatography (accessed January 2014.)

http://www.gerstel.com/en/preparative-gc-fraction-collection.htm (accessed January 2014.)

Sciarrone, D., Pantò, S., Ragonese, C., Tranchida, P.Q., Dugo, P., and Mondello, L. (2012) *Anal. Chem.*, **84**, 7092–7098.

Zougagh, M., Valcarcel, M., and Rios, A. (2004) *TRAC-Trend. Anal. Chem.*, **23**, 399–405.

20

Strategies for Lipid Analysis

Irina A. Guschina and John L. Harwood

School of Biosciences, Cardiff University, Cardiff, UK

1 WHAT LIPIDS ARE WE DEALING WITH?

The major sites of lipids in plants are as constituents of membranes, in lipid stores (when they occur), and as part of the surface coverings of leaves, shoots, and roots. For whole tissues, seeds will be dominated by stored triacylglycerol (when present), whereas leaves, containing large amounts of chloroplasts, will have large amounts of their unique lipids (see later). Algae will also contain chloroplast lipids but, in addition, may have a whole range of rather uncommon lipids whose description is beyond the remit of this chapter (but see Harwood and Jones, 1989; Guschina and Harwood, 2006). Some representative compositions of plant membranes are shown in Table 1. As in other organisms (Gurr, Harwood, and Frayn, 2002), each membrane has its own characteristic lipid composition, which is maintained within tight limits (presumably because of functional requirements). In general, most extra-chloroplastic membranes have phosphoglycerides as their main components. Phosphatidylcholine is the most abundant followed by phosphatidylethanolamine (Table 1). The plasma membrane generally contains significant quantities of plant sterols (e.g., β-sistosterol and stigmasterol), sphingolipids, and phosphorylated derivatives of phosphatidylinositol.

The chloroplast thylakoids have a distinct lipid composition that is shared with the thylakoids of algae and cyanobacteria. The major components are three glycosylglycerides – monogalactosyldiacylglycerol, digalactosyldiacylglycerol, and sulfoquinovosyldiacylglycerol (the plant sulfolipid) (Figure 1 and Table 2). A detailed review on their distribution, metabolism, and analysis (Heinz, 1996) is available. Only one other lipid is found in significant amounts, phosphatidylglycerol

(Table 2). Owing to their different metabolism, plants are sometimes divided into 16:3 and 18:3 plants. These are distinguished by the acyl composition of their chloroplast lipids (Table 3).

Plant tissues (especially leaves) are highly unsaturated with, for example, α-linolenic acid accounting for over 65% of the total fatty acids of leaves. This is due to the dominance of chloroplasts in such photosynthetic tissues and their high linolenate content (Table 3).

For storage tissues, usually triacylglycerol is accumulated. The main exception among commercially important crops is jojoba where wax esters are found (Gunstone, Harwood, and Dijkstra, 2007). Although most of the important crops accumulate triacylglycerols containing a narrow variety of fatty acids (palmitate, oleate, linoleate, and α-linolenate), plant species with a whole variety of unusual fatty acids are possible (see Gunstone, Harwood, and Dijkstra, 2007; Murphy, 1994, 2005). Many of these have important application in industry. The main fatty acids in plant or algal tissues are listed in Table 4, together with their short-hand and systematic names.

The surface covering of plants consist of the cuticle (waxes and cutin) for the aerial parts and suberin for the roots. Surface waxes are exposed to the environment and, hence, are chemically rather stable. Some general futures of surface waxes are shown in Table 5. However, it must be emphasized that both the structures and the composition of surface waxes vary considerably from organism to organism (Kolattukudy, 1975, 1980; Hamilton, 1996; Walton, 1990).

Cutin consists of a hydroxy fatty acid polymer that is embedded in or associated with a complex mixture of lipids (Kolattukudy, 1980, 1987; Walton, 1990). These are two families of acids (Table 6)

Table 1 Comparison of the major acyl lipids of different plant membranes.

Membrane	Composition (% of total acyl lipids)							
	MGDG	DGDG	SQDG	PC	PE	PG	PI	DPG
Plasma[a]	—	—	—	32	46	—	19	3
Chloroplast envelope[b]	22	32	5	27	—	8	1	—
Chloroplast thylakoid[b]	51	26	7	3	—	9	1	—
Mitochondrial (outer)[a]	—	—	—	54	30	5	11	—
Mitochondrial (inner)[a]	—	—	—	41	37	3	5	15
Glyoxysomal[c]	—	—	—	51	27	3	9	2
Endoplasmic reticulum[c]	—	—	—	45	29	4	13	3

Abbreviations: —, <1%; MGDG, monogalactosyldiacylglycerol; DGDG, digalactosyldiacylglycerol; SQDG, sulfoquinovosyldiacylglycerol; PC, phosphatidyl-choline; PE, phosphatidylethanolamine; PG, phosphatidylglycerol; PI, phosphatidylinositol; DPG, diphosphatidylglycerol (cardiolipin).
[a] Potato tuber.
[b] Spinach leaf.
[c] Castor bean seeds.
Source: Data from Harwood (1998) where further information is given.
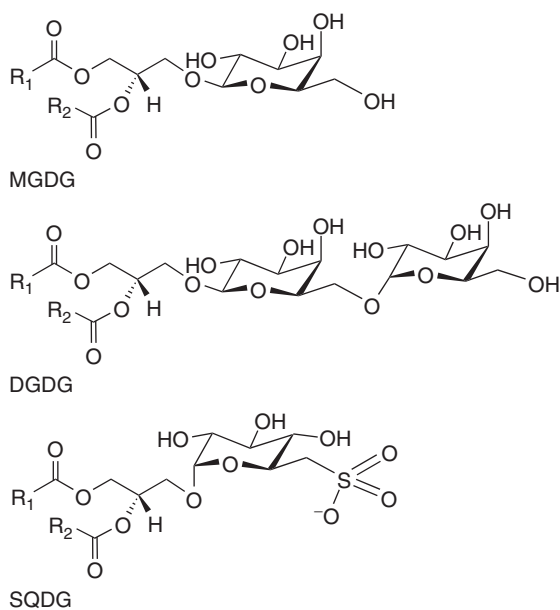
Figure 1 Structures of the major plant glycosylglycerides. (a) Monogalactosyldiacylglycerol (1,2-diacyl[β-D-galactopyran-osyl-(1′→3)]-*sn*-glycerol), MGDG. (b) Digalactosyldiacyl-glycerol (1,2-diacyl[α-D-galactopyranosyl-(1″→6′)-β-D-galacto-pyranosyl-(1′→3)]-*sn*-glycerol, DGDG. (c) Sulfoquinovo-syldiacylglycerol (1,2-diacyl[6-sulfo-α-D-quinovopyranosyl-(1′→3)]-*sn*-glycerol), SQDG. Also known as the plant sulfolipid.

based on palmitic and oleic acids, respectively. In the polyester intramolecular structure, there is much cross-linking, influenced by the availability of secondary hydroxyl groups. The building up of different types of structures is described by Deas and Holloway (1977).

The main lipid components of suberin are ω-hydroxy fatty acids and dicarboxylic acids (Table 7). Phenolics are also major constituents unlike cutin. In both cutin and suberin, environmental conditions and developmental stages can alter the overall composition significantly (Harwood, 1980). Moreover, because of their polymeric nature, special methods are needed for their quantitative and qualitative analyses (Walton, 1990).

2 EXTRACTION AND INITIAL HANDLING OF LIPIDS

Each of the three main types of plant lipids (nonpolar acyl lipids, polar membrane lipids and surface lipids) requires different considerations for extraction. Ideally, the latter should be quantitative and preserve the individual lipids in an intact state. For some lipid classes, this is relatively easy; for others (particularly some surface layer components), it is problematic. Because of the nature of lipids and their interactions with each other (or alternative cellular constituents), a mixture of solvents is always required (Kates, 2010). Nonpolar/polar solvent mixtures (e.g., hexane/methanol) are usually used. Of course, these will not be effective for covalently bound lipids and, in that case, acid or alkaline hydrolysis must be employed.

There are two particular problems often encountered with plant tissues. First, as mentioned in the previous section, polyunsaturated fatty acids are often major constituents (Table 8). These are prone to rapid

Table 2 Lipid composition of plant leaves and chloroplast thylakoids.

	% of total acyl lipids						
Plant leaf	MGDG	DGDG	SQDG	PC	PE	PG	PI
Maize	42	31	5	6	3	7	1
Rye grass	39	29	4	10	5	7	2
Clover	46	28	4	7	5	6	1
Broad bean	38	30	6	7	4	6	2
Thylakoid							
Spinach	51	26	7	3	—	9	1
Wheat	45	36	8	2	—	10	—

Abbreviations: MGDG, monogalactosyldiacylglycerol; DGDG, digalactosyldiacylglycerol; SQDG, sulfoquinovosyldiacylglycerol; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PG, phosphatidylglycerol; PI, phosphatidylinositol.
Source: Data taken from Gunstone, Harwood, and Dijkstra (2007).

Table 3 Comparison of the chloroplast thylakoid acyl lipid composition in a "16 : 3" (spinach) and an "18 : 3"-plant (barley).

Plant	Lipid	Major fatty acids (% total fatty acid)					
Pea		16 : 0	16 : 1[a]	16 : 3	18 : 1	18 : 2	18 : 3
	MGDG	3	1	—	1	3	90
	DGDG	9	2	—	3	7	78
	SQDG	32	3	—	2	5	55
	PG	18	27	—	2	11	38
Spinach							
	MGDG	1	—	25	1	2	72
	DGDG	3	—	5	2	2	87
	SQDG	29	1	—	7	26	36
	PG	11	32	—	2	4	47

16 : 1[a] = *cis*-9-hexadecenoic acid (palmitoleic acid) in all lipids except PG where it is mainly *trans*-3-hexadecenoic acid.
For lipid abbreviations see Table 1.

oxidation and particular precautions [rapid processing, use of a nitrogen headspace, and addition of an antioxidant such as BHT (2,6-di-*tert*-butyl-*p*-cresol)] must be taken to prevent their destruction. Second, many plant tissues contain catabolic enzymes (e.g., acyl hydrolases) that, if not inactivated, will destroy complex lipids. Such enzymes are active at low temperatures – hence, the reason for briefly boiling Brussels sprouts before freezing! Traditionally, isopropanol is used to inactivate these enzymes.

Before giving some details of recommended methods, a few more general remarks should be made. If possible, the tissues to be analyzed should be extracted as rapidly as possible. If this cannot be done, then storage at as low a temperature as possible is advised (−80°C better than −20°C). Because freezing destroys subcellular structures and, hence, allows catabolic enzymes access to lipid substrates, then even if they have not hydrolyzed lipids on storage, they certainly will when samples are unfrozen. So, it is wise to homogenize the samples in isopropanol as they begin to thaw out. Of course, the better overall procedure is to expose samples to hot isopropanol and carry out extraction before storing samples. All lipids should be handled in glassware (which can be acid-washed to clean it thoroughly) – plasticizers easily contaminate lipids and, of course, many plastics are themselves soluble in organic solvents! Addition of BHT (10–50 mg L^{-1}) and storage under nitrogen minimize oxidation.

How effective these precautions have been will be immediately revealed as analysis proceeds. Large amounts of unesterified fatty acids, partial glycerides, and oxidized fatty acids are tell-tale signals. Also remember that in many purification procedures, antioxidants will often be separated from the

Table 4 Major membrane fatty acids of plants and algae.

Plants	
16:0	Hexadecanoic acid (palmitic acid)
16:1	9-Hexadecenoic acid (palmitoleic acid)
16:3	7,10,13-Hexadecatrienoic acid
18:0	Octadecanoic acid (stearic acid)
18:1	9-Octadecenoic acid (oleic acid)
18:2	9,12-Octadecadienoic acid (linoleic acid)
18:3	9,12,15-Octadecatrienoic acid (α-linolenic acid)
Additional algal acids	
14:0	Tetradecanoic acid (myristic acid)
16:4	4,7,10,13-Hexadecatetraenoic acid
18:4	6,9,12,15-Octadecatetraenoic acid (stearidonic acid)
20:4	5,8,11,14-Eicosatetraenoic acid (arachidonic acid)
20:5	5,8,11,14,17-Eicosapentaenoic acid (EPA)
22:6	4,7,10,13,16,19-Docosahexaenoic acid (DHA)

Acids are shown with their short-hand nomenclature on the left. This is used commonly for GLC and is convenient since it does not specify the double bond position which requires additional analysis to be properly defined. All the acids shown have *cis* double bonds but in plants a specific fatty acid is found esterified to the *sn*-2 position of phosphatidylglycerol – *trans*-Δ3-hexadecenoic acid. The double bonds for acids in the Table are numbered from the carboxyl group, that is, they are all Δ.

The 16:3 in higher plants is found in the leaves of certain species (e.g., spinach) termed 16:3 plants (Table 3). It is preferentially esterified to MGDG. For algae, not all species contain the fatty acids shown but they are present in a good variety.

lipids, necessitating further additions to the extracts as manipulations proceed.

Most (all?) solvents are toxic, so appropriate precautions must be taken. Because of this, many new extraction mixtures, which do not contain chloroform, have been developed. However, this solvent has one advantage in being nonflammable. Working in a fume hood alleviates most danger, although one must still consider the different methods of disposal for various solvents and their mixtures. Flammable solvents are an obvious hazard and must always be handled carefully and, preferably, in small volumes. Storage of lipids in such solvents should be done in sealed containers and in a spark-proof fridge.

Because of the extreme sensitivity of modern analytical techniques and the delicate nature of many lipids, it may be necessary to redistill solvents before use. This would, for example, remove peroxides that could destroy fatty acid double bonds.

2.1 Some Extraction Methods

2.1.1 General Extraction Methods

For the efficient extraction of tissues, it is usually necessary to first homogenize (or disrupt) the tissue in the presence of a water-miscible solvent or solvent mixture. Afterward, a two-phase partition needs to be produced in order to separate lipids from other solubilized material. For quantitative purposes, it may be necessary to re-extract the tissue and to rewash the phases. It must be born in mind that there are few cases where all the individual lipids present in a tissue (Table 9) can be adequately extracted in a single procedure (Christie and Han, 2010).

Most methods for extracting lipids from tissues use a technique based on the classic methods of Folch, Lees, and Sloane-Stanley (1957) or Bligh and Dyer (1959). For plant tissues, we routinely use a method that is based on a modification of the latter that gives better recovery of polar glycerolipids (e.g., phosphatidate and polyphosphoinositides) (Garbus *et al*., 1963). Because of the potential problems with endogenous catabolic enzymes, we routinely incorporate an initial treatment with hot isopropanol (Smith, Douce, and Harwood, 1982).

2.1.2 Special Considerations

Some seed tissues contain intractable (and rather polar) lipids such as lysophospholipids. Special methods such as those of Morrison, Tan, and Hargin (1980), Colborne and Laidman (1975), or Osagie and Kates (1984) are needed. Some algae may have tough cell walls and may need special homogenization techniques (Fuschino *et al*., 2011). Fortunately, for photosynthetic tissues, the extraction of pigments is a sign that the method is proving adequate. For surface lipids (waxes, cutin, and suberin), special adaptations are needed and, in the latter case, some hydrolytic steps are required (Walton, 1990).

2.1.3 Extraction Before Mass Spectrometry and Lipidomics

When a lipidomic study is envisaged (or other use of mass spectrometry (MS)), then it may be advisable to

Table 5 The main components of plant waxes.

Class	Comments	Chain length range in plants (in *Arabidopsis*)	% in *Arabidopsis* stems
Fatty acids	Very common	C16–C34	3
	Usually even-chain saturated	(C30, C28)	
Fatty alcohols	Common	C22–C34	12
	Even chains predominate	(C28, C30, C26)	
Wax esters	Common	C32–C64	1
n-Alkanes	Common	C21–C35	38
	Usually C29, C31	(C29, C31, C27)	
Ketones	Not as common	C21–C35	30
	as alkanes	(C29)	
Secondary alcohols	About as common	C21–C35	10
	as ketones	(C29, C31, C27)	

Data taken from Kolattukudy (1980) and Kunst and Samuels (2003). Arabidopsis components are in order of abundance.

Table 6 The cutin acids, major components of plant cutin.

16C family

$H_3C[CH_2]_{14}COOH$

$HOCH_2[CH_2]_{14}COOH$

$HOCH_2[CH_2]_x CHOH[CH_2]_y COOH$

$(x + y = 13; y = 5–8)$

18C family[a]

$H_3C[CH_2]_7 CH = CH[CH_2]_7 COOH$

$HOCH_2[CH_2]_7 CH = CH[CH_2]_7 COOH$

$HOCH_2[CH_2]_7 CH–CH[CH_2]_7 COOH$

$HOCH_2[CH_2]_7 CHOHCHOH[CH_2]_7 COOH$

[a] Δ12 unsaturated analogues also occur.

See Harwood (1980) and Kolattukudy (1975, 1977) for more details.

use methods that minimize the presence of inorganic salts. Two suitable procedures are those of Cheng, Guan, and Han (2006) and Han *et al*. (2005).

2.1.4 Further Information About Extraction

General comments and many references to different extraction methods can be found in Christie and Han (2010) and Kates (2010). Further information can be found in Gunstone, Harwood, and Dijkstra (2007) and on the AOCS Lipid Library site (www.lipidlibrary.aocs.org).

3 RADIOACTIVE LABELING

Metabolic studies on plant lipids can be carried out conveniently using appropriate radiolabeled precursors. Of the radioisotopes used, ^{3}H, ^{14}C, ^{32}P, and ^{35}S are the most common. Because ^{32}P emits β-rays of high energy, its use requires appropriate screening to minimize exposure. Where appropriate

(and available) ^{33}P is an alternative, being a soft β-emitter and with a longer half life. Before starting any experiment with radioactivity, it is essential to be aware of the dangers and to take all necessary measures to minimize these. Most institutions will have prescribed procedures for the safe handling and disposal of materials.

For metabolic experiments on plant lipids, [^{14}C]acetate is particularly useful, as (in most tissues) it labels the fatty acyl chains particularly well. [^{14}C]Glycerol and [^{14}C]mevalonate label glycero-lipids and isoprenoid compounds, respectively. For particular enzyme reactions or to follow individual pathways, there are often more specific precursors available (e.g., acetyl-CoA for acetyl-CoA carboxylase, malonyl-CoA for fatty acid synthase, and CDP-choline for cholinephosphotransferase).

In vitro enzyme reactions can be easily measured using radiolabeled substrates. The three main requirements are to choose a substrate that is only used by the enzyme being measured (often a problem with crude subcellular fractions), to consider how to prevent the product being metabolized further, and to have a convenient method to easily separate

Table 7 Differences in the usual compositions of plant cutin and suberin.

Component	Suberin	Cutin
Phenolics	High	Low
Dicarboxylic acids	Major	Minor
Very long chain (20–26 C) acids	Common, substantial	Rare, minor
Very long chain alcohols	Common, substantial	Rare, minor
In-chain substituted acids	Minor	Major

Source: Data taken from Kolattukudy (1975).

the radioactive product from (a large amount of) the radioactive substrate. The multiple volumes of *Methods in Enzymology* are a good first source of information.

For *in vivo* experiments and the labeling of intact tissues, consideration should be given as to how to get adequate uptake of the radioactive precursor. Clearly, it is no good trying to use, for example, nucleotide derivatives with whole tissues – they will never penetrate the cell wall or plasma membrane and other layers. In some cases, a detergent or a mild solvent can be used. A perusal of the literature will often give clues as to how to facilitate labeling, but we would advise a careful independent check of the specific system to be used because plant species can often have individual characteristics.

Some examples are as follows:

- monocotyledon leaves with [^{14}C]acetate (Walker, Ridley, and Harwood, 1988)
- dicotyledonous leaves with [^{14}C]acetate (Walker, Ridley, and Harwood, 1988)
- [^{14}C]fatty acid labeling of leaves (Murphy *et al.*, 1985)
- potato tuber labeling for suberin formation (Walker and Harwood, 1986)
- callus tissue preparations of oil crops (Ramli *et al.*, 2002)
- developing oilseed rape embryos (Tang *et al.*, 2012)
- tissue suspension culture (soybean) (Guschina *et al.*, 2004)
- algal suspensions (Fuschino *et al.*, 2011).

Once the metabolic experiment has been completed, it is necessary to analyze the data. For enzyme reactions *in vitro*, it may be that a simple extraction followed by liquid scintillation counting is enough. It is recommended that a scintillant with an adequate capacity for dissolving some water is used. Quench corrections should always be made and it must be born in mind that some solvents (e.g., chloroform) are strong quenchers. For mixtures of lipids, separations will be needed such as liquid or thin-layer chromatography (TLC). The latter can be assessed by autoradiography, although for accuracy we would recommend scraping relevant areas and scintillation counting. When scraping plates, great care should be taken to minimize breathing in dust. Not only is silica harmful but the particles may also be radioactive! The procedures must be carried out in a fume hood. Lightly spraying the plates (with water) will minimize dust. For liquid chromatographic separations,

Table 8 Fatty acid composition of different plant tissues.

	Percentage of total fatty acids					
	16:0	18:0	18:1	18:2	18:3	Others
Castor bean seed	1	tr	3	5	tr	91[a]
Oilseed rape seed	4	1	15	14	9	57[b]
Narcissus bulb	17	1	9	69	3	1
Potato tuber	19	3	3	60	15	tr
Turnip root	15	2	8	14	59	2
Barley leaf	13	2	6	6	64	9[c]
Pea leaf	12	1	2	25	53	7[c]

[a] Ricinoleic acid is over 90%.
[b] An erucate-containing line with 45% 22:1.
[c] Leaves contain around 3% *trans*-3-16:1 which is exclusively found esterified in PG.
Source: Data taken from Harwood (1980).

Table 9 Acyl lipid composition of some plant tissues.

	Percentage of weight of total lipids			
	Apple fruit	Soybean seed	Barley leaf	Clover leaf
MGDG	1	tr	43	46
DGDG	5	tr	26	28
SQDG	1	tr	5	4
PG	1	tr	6	6
PC	23	4	11	7
PE	11	2	4	5
PI	6	2	1	1
DPG	1	tr	tr	tr
TAG	5	88	nd	nd
Other	46[a]	4	4	3

[a] Sterols, sterol glycosides total 26%. Lipid abbreviations in Tables 1 and 2.
Source: Data taken from Harwood (1980).

eluted lipids can be collected directly for counting (Guiheneuf *et al*., 2013).

Detection and quantification of labeled fatty acids is conveniently carried out using a gas-flow proportional counter attached via a splitter to a gas chromatograph (GC). [14C] is a better isotope that [3H] because the latter has lower energy and, hence, gives relatively poor efficiencies. Radio-GC has been used for over 50 years, since the pioneering developments by A.T. James (see Hitchcock and Nichols, 1971, for early papers). It is a very useful and sensitive technique provided that one has the instrumentation available. For those without, high performance liquid chromatography (HPLC) with tedious collection of samples as they elute or using an instrument with a gas-flow proportional counter and a light scattering detector in series are possibilities.

For general comments on the use of radiolabeled precursors and their use in metabolic experiments, see the book by Kates (2010).

4   SEPARATION TECHNIQUES (TLC)

4.1   Thin-Layer Chromatography

TLC has long been used as the effective technique for the separation of lipid mixtures into their individual lipid classes. Despite the more modern separation technique of HPLC (discussed later in this section), several advantages still make TLC particularly useful for lipid analysis, for example, simplicity and sensitivity with relatively low capital costs. Analyses are

carried out rapidly and many samples can be analyzed simultaneously. The separated individual lipid classes can be visualized using nondestructive spray reagents or specific sprays to detect functional groups in complex lipids, for example, phosphorus in phospholipids and galactose in galactolipids. Moreover, all organic compounds may be detected by charring. Lipid classes separated by TLC and revealed with nondestructive sprays can be recovered by scraping the band and eluting with a particular solvent or solvent mixture (e.g., chloroform–methanol–water, 5 : 5 : 1, by volume, for phospholipids and glycosyldiacylglycerols; chloroform–methanol, 1 : 1, by volume, for nonpolar lipids).

Silica gel (Kieselgel) is the adsorbent most frequently used for the analysis of lipids but other adsorbents such as Kieselguhr, aluminum oxide, magnesium hydroxide, ion-exchange celluloses, and Sephadex may also be used. Silica gel can be modified with various compounds in order to separate the individual lipid classes that are not well resolved by TLC on standard silica gel (such modifications are discussed in later sections). Silica gel on TLC plates should be activated by heating the plates (before use) at 100–110°C to remove water.

Glass is the most commonly used support for the adsorbent on TLC plates, although aluminum and plastic sheets can also be used with the advantage that they can be cut into small pieces. The plates with nonglass supports are available only as precoated plates.

Precoated commercial plates (20 × 20 cm) are available now in both analytical (0.2–0.5 mm layers) and preparative (1–2 mm) thickness with the

following layers of silica gel: H (no binder), G (organic binder), or 60A, having high stability and uniformity. Precoated TLC plates are much more convenient than laboratory-made as they give more reproducible results. For high performance thin-layer chromatography (HP-TLC), only precoated plates with the silica gel 60 adsorbent of 4.5 and 5 μm particle size are used. The adsorbent layer in such plates is slightly thinner (0.15–0.20 mm) than in standard analytical TLC. In addition to the better resolution of the lipid mixture, HP-TLC has several other advantages over standard TLC, such as less sample volume needed, a smaller volume of development solvent, and a shorter analysis time.

Lipid samples can be applied to the TLC plates as discrete spots or narrow streaks (the smaller the area of the application, the sharper the resolution) at a distance of 1.5–2 cm from the bottom of the plate. Samples are applied using a syringe in a solvent (as nonpolar as possible, frequently chloroform). Methanol should not be used for sample application as it produces large spots and wide streaks. The application of authentic lipid standards, either as a mixture or as individual compounds, alongside the sample lipids helps preliminary identification of the compounds present in the lipid samples (lipid identification after TLC separation is discussed in more detail later). The plate is placed in a development chamber/tank containing the mobile phase and lined with filter paper to help saturate the atmosphere. All chambers should have a heavy lid that acts to seal the chamber and maintain the atmosphere saturated with solvent vapor within the chamber. When the mobile phase nears its top, the plate is removed from the tank and dried in air or a stream of nitrogen to evaporate the solvents.

When a complex lipid mixture cannot be separated by TLC in one direction, two-dimensional separation can be more effective. The solvent systems used in the application of one- and two-dimensional TLC for the analysis of nonpolar phospholipids and glycolipids are described in detail by Henderson and Tocher (1992), Kates (2010), and Christie and Han (2010).

For increasing the resolving power of chromatography, TLC separation on modified silica gel can be applied. Such modifications include reversed-phase thin-layer chromatography (RP-TLC) as well as silver nitrate, urea, boric acid, ethylenediaminetetraacetic acid (EDTA), and oxalic acid impregnated TLC. RP-TLC is used for more efficient separation of lipophilic substances including fatty acid methyl and phenacyl esters, triacylglycerols, and cholesteryl

esters (Henderson and Tocher, 1992). Silver nitrate TLC is employed to separate lipids according to the number of double bonds in their fatty acids. TLC with boric-acid-impregnated silica is useful for separation of partial acylglycerols and phospholipids to get clear resolution of phosphatidylserine and phosphatidylinositol (they are very difficult to separate by TLC on standard silica gel 60). Impregnation with EDTA or oxalic acid improves the separation of anionic phospholipids such as phosphatidylinositol, phosphatidylserine, and phosphatidic acid.

4.1.1 Detection Systems and Quantification of Separated Lipids

After development, the plates must be sprayed or treated with reagents or stains to reveal the location of the separated lipids. In general, the detection systems can be divided into two main categories: (i) specific for certain lipids or their functional groups and (ii) nonspecific reagents to visualize all lipids. Methods using nonspecific reagents are:

- spraying the plate with a strong oxidizing agent (e.g., 5–10% H_2SO_4 in methanol or ethanol and 0.25% or 0.6% potassium dichromate in 15% H_2SO_4) followed by charring;
- spraying the plate with a fluorescent reagent (e.g., 0.1% 8-anilino-1-naphthalene-sulfonic acid as the ammonium salt in methanol or 0.05–0.25% rhodamine B in ethanol) and examining under UV light (this is nondestructive);
- other miscellaneous stains including reaction with iodine vapor.

Lipids appear as brown spots after exposure to iodine vapor, and with brief exposure, it is possible to remove the iodine effectively under vacuum. It should be noted that iodine reacts to some extent with polyunsaturated fatty acids, which cannot then be recovered from the silica gel for the further analysis. After spraying the plate with oxidizing agents and charring, lipids appear as black deposits of carbon. Although this method is destructive for lipids, it is very sensitive and as little as 1 μg of lipid can be detected.

Specific stains generally contain a chemical or chemicals in the reagent reacting with particular groups in the lipids that makes them stained or visible

in some way. Some of the most commonly used stains are listed below as follows:

- "Zinzadze" reagent and molybdenum blue reagent for visualization of lipids containing phosphorus;
- ninhydrin solution to stain amino-lipids;
- Dragendorf reagent for choline-containing lipids (quaternary N group);
- α-naphthol, orcinol, and diphenylamine for detection of the sugar groups in glycolipids, such as cerebrosides, sulfatides, gangliosides, and glycosylglycerides;
- periodate-Schiff's reagent to detect lipids containing vicinal diol groups (e.g., inositol and glycerol);
- acidic ferric chloride solution to stain cholesterol and its esters.

Detailed information on various stains and reagents can be found in Henderson and Tocher (1992), Kates (2010), and Christie and Han (2010).

Gravimetric methods should not be used for quantification of lipids separated by TLC as some impurities, for example, silica gel, may also be eluted from the plates and weighed. Photodensitometry of charred plates is probably the most popular of all destructive methods to quantify separated lipids, and modern instruments are capable of generating accurate data after suitable calibration. A nondestructive method for quantification of lipids is to use fluorometry, in which the fluorescence of a dye produced by the presence of lipids is measured by a laser-scanning fluorometer. A number of chemical methods can be used for hydrolyzed products of lipids, for example, phosphorus for phospholipid quantifications, sugars for glycolipids, as well as cholesterol and glycerol (Christie and Han, 2010). It is possible to determine the amount of individual lipid classes by measuring the amount of fatty acids that they contain. In this case, the fatty acids of each lipid class separated by TLC are converted to their methyl esters in the presence of a known amount of an internal standard (usually, an odd-chain fatty acid that is not present in the sample naturally) (Christie and Han, 2010).

4.2 Gas–Liquid Chromatography (GLC or GC)

Gas-liquid chromatography (GLC) is used widely in the analysis of all the major lipid classes (Evershed, 1992). It is a method of choice for rapid, quantitative analysis of volatile lipid components such as hydrocarbons, fatty acids esters, fatty alcohols, and sterols (Christie and Han, 2010). The GLC analysis of lipids involves (i) introduction of a small volume (1–5 μL) of a dilute solution (nanogram or milligram quantities of lipid per microliter, dissolved in a suitable organic solvent, such as hexane, cyclohexane, or petroleum ether) in the GC column via a gas-tight injection port; (ii) volatilization of the diluting solvent and dissolved lipids into the inert carrier gas, followed by continuous passage through the GC column; (iii) separation of mixtures of volatile and semivolatile lipids according to their different vapor pressures owing to their different boiling points; and (iv) detection of the compounds eluting from the GC column.

Most commercial GC systems will meet the operating specifications required for lipid analyses. Variations between different instruments may relate to variability in the range of injector and detector systems, mode of carrier gas-flow control, oven-temperature operating range, and a range of data recording/collection options.

4.2.1 The Column and Stationary Phase

There are two main types of column used for GLC, capillary columns and packed columns. Wall-coated open tubular (WCOT) capillary columns (flexible fused silica capillary column coated externally with either polyimide or aluminum) are the most common type used in the analysis of lipids nowadays. They are widely available commercially. The polymeric stationary phase is coated as a thin film (0.1- to 10-μm film thickness) on the internal wall of the capillary. Typical column dimensions are 10 to $50\,\text{m} \times 0.1$ to 0.32 mm i.d. Advantages of using capillary columns are (i) high column efficiencies and high resolving power; (ii) enhanced sensitivity; and (iii) compatibility with mass spectrometers owing to low gas flow rates. Disadvantages include long analysis times, low sample capacities, and the fact that columns are expensive and fragile.

Typical packed-column dimensions are 1 to $3\,\text{m} \times 4$ mm i.d. The column (made of deactivated glass) is filled with packing material, usually fine particles of washed and deactivated diatomaceous earth (support) coated with polymeric stationary phase (typical loadings of stationary phase in the range of 1–10% w/w, depending on the compounds being

analyzed). Advantages of using packed columns are (i) robust nature; (ii) large sample capacity; and (iii) short analytical time. Relatively low efficiencies and limited separation power are disadvantages of this type of columns (Evershed, 1992).

The most important factor in the selection of a suitable stationary phase is the nature of the separation required. Over the years, certain types of phases have emerged as favorites for special purposes, and these are discussed by Christie and Han (2010) in relation to each lipid class. A list of commonly used stationary phases and supports as well as their characteristics is given by Kates (2010). The most commonly used stationary phases are the polysiloxanes, particularly the apolar dimethyl polysiloxanes. A polar stationary phase can be employed to separate individual compounds based on the differences in dipole–dipole interactions, and the polyethylene glycol stationary phases are widely used for this purpose.

The stationary phase on a column can deteriorate mainly because of chemical attack. Most polar liquid phases are very sensitive to oxygen and water; thus, it is strongly advised that all traces of these be removed from the carrier gas by introducing traps containing suitable molecular sieves and oxygen scrubbers (available commercially) between the gas cylinder and the column. Some polar solvents and traces of polar impurities may slowly react with the liquid phase and or even displace it from the column, and nonvolatile compounds injected onto a column along with the analyzed samples may gradually build up and change the characteristics of the liquid phase. Special precautions should be taken regarding operating temperatures because an exposure of the column to excessive temperatures will ruin the column. The carrier gas should always flow through the column when heated. Most damages usually occur in the first few coils of the column, so cutting off 10–15 cm of this part of the column may be advised. Moreover, when the damage is minimal, reversing the column may be all that is required.

4.2.2 Injection Systems

With WCOT columns, there are several injection systems available from commercial sources. The properties of the major types are discussed in detail by Evershed (1992) and Christie and Han (2010). During the injection process, it is necessary that the sample should not change in composition (e.g., due to thermal degradation or rearrangement), the solvent peak should not interfere with the analytes, and retention times and relative peak areas should be reproducible. Autoinjection is best when it is available.

In an on-column injection method, the small volume (less than 1 μL) of dilute sample is injected in a solvent directly into the interior of the capillary column. This eliminates problems of discrimination between high and low boiling sample components during injection. Correct injection technique is vital, and in particular, the syringe plunger should be depressed rapidly so that the sample is "sprayed" into the column.

In the split injection mode, the sample solution is introduced using a syringe via a rubber septum inlet into a heated injector zone (200–300°C). Such a high temperature ensures rapid volatilization of the sample solution in the carrier gas that is divided into two streams: one is directed into the column while the second is vented to the atmosphere. The flow through the latter is regulated before injection by a control valve to give the desired split ratio, usually between $1:20$ and $1:200$. This factor, combined with the high flow rate (100–200 mL min^{-1}) of carrier gas through the injector area, ensures that the sample remains there only transiently, before delivery into the column as a narrow plug and it is followed by the pure carrier gas. This is the most common method used for fatty acid analysis, it is easily married to autoinjection and can give excellent results.

As with on-column injection, splitless injection is best suited to the analysis of dilute samples. The sample solution is volatilized in the injector port and the vapor delivered by the low carrier gas flow (2 mL min^{-1}) onto the GC column that is held at low temperature (50°C), so the solvent and sample components condense as a narrow plug at the head of the column. After a lapse of around 1 min, a flow of gas is introduced to purge any remaining sample from the injector area. Splitless injection is a less reliable type of injection for quantitative analysis of lipids due to problems of discrimination between high and low molecular weight components in the analyzed samples (Evershed, 1992).

Programmed temperature vaporizing (PTV) injection has been recently introduced, and this combines split and direct modes in a single module. As direct on-column injection can lead to a relatively rapid deterioration of the column, a specially designed

injection port takes over the function of the top part of the column. The sample is introduced into this inlet at a temperature below the boiling point of the solvent, and then the temperature is raised at a controlled rate so that the sample components are selectively vaporized. PTV injection is best suited to the analysis of lipids of high molecular weight, such as intact triacylglycerols.

4.2.3 Detection of Eluting Compounds

A large number of detectors operating on different principles have been developed for use in GLC, but only a few of them continue to be used to a significant extent. The flame ionization detector (FID) is now almost universally adopted as it can be used with practically all organic compounds and has high sensitivity and stability and a low dead volume. The FID detector has a fast responsive time, which is linear over a very wide dynamic range and is very reliable in long-term operations.

In the electron-capture detector, a radioactive source is used to bombard the carrier gas with β-particles as it passes through an ionization chamber. Each β-particle can generate up to a thousand thermal electrons, which are collected by applying a voltage potential.

MS has become an invaluable tool for the identification and detection of lipids separated by GLC, sometimes even when they are incompletely resolved. The use of mass spectrometry-gas–liquid chromatography (MS-GLC) is discussed in Section 5.5.

4.3 High Performance Liquid Chromatography (HPLC)

HPLC is the analog of GC where the stationary phase is held in a metal column and the liquid mobile phase is forced through under pressure. The sample is injected into the mobile-phase stream and the separated components are detected as they elute from the column. There are several modes of separation in chromatography that are applicable to lipid analysis, namely adsorption, normal-phase liquid partition, liquid–liquid (in which reversed-phase chromatography is a special case), ion-exchange and ion-pairing, and exclusion and chiral-phase chromatography. Separation in all modes arises from molecular interactions between the solute and mobile phase and stationary phases in which different types of forces are involved. With lipids, hydrophobic interactions are particularly important, while dispersive, that is, weak impermanent electrical charges, polar and ionic forces may be involved to a limited extent.

Silica gel has long been used by lipid analysts as an adsorbent for TLC and low pressure column chromatography, and now, it is widely used for the HPLC of lipids, especially for the classes of lipids separated according to the number and nature of the polar functional groups such as ester bonds, phosphate, hydroxyl, and amine groups in molecules. The activity of silica is due to the presence of silanol groups (Si-OH) on the surface and these have a strong affinity for water and, therefore, for reproducible retention, the amount of water in the mobile phase must be controlled. Other factors that will affect the chromatographic properties of the silica are surface area, pore size, and pore volume. A particular form of adsorption chromatography, argentation chromatography, is widely used for lipid separation. Silica gel is impregnated with silver nitrate, and this technique is used to separate the lipid samples according to the number and configuration (*cis-* or *trans-*isomers) of double bonds in the acyl or alkyl moieties.

Two forms of partition chromatography, normal phase and reversed phase, are utilized for lipid analysis and they are recognized according to the relative polarities of the stationary and mobile phases. So, the stationary phase is more polar than the mobile phase for normal-phase mode, and the stationary phase is less polar than the mobile phase for reversed-phase chromatography. The most widely used stationary phase for this type of HPLC is octadecylsilyl-silica (ODS or C_{18}) which is prepared from silica by reacting the surface silanol groups with an organochlorosilane. The majority of lipid separations using this HPLC type are for the separation of molecular species of lipids within a single lipid class when it depends on the fatty acyl or alkyl chain length and on the configuration at any double bond.

Ion-exchange stationary phases are either silica or polymer- based. The cellulose-based anion exchangers diethylaminoethyl (DEAE) and triethylaminoethyl (TEAE) celluloses have been used to separate complex polar lipids using low pressure HPLC. This technique has applications in the separation of phospholipids.

Exclusion chromatography, also called gel-permeation or gel-filtration, is based on the exclusion of molecules from the pore structure of the stationary phase, which may be a porous polymer or a porous silica. This technique is used for the separation of fatty acids and mono-, di-, and triacylglycerol mixtures as well as for the separation of polymerized lipids.

In chiral-phase HPLC, stationary phases have a chiral molecule bonded chemically to a silica-based matrix, which makes possible the separation of enantiomers by HPLC, for example, long-chain monoacyl, monoalkyl, and diacyl glycerols.

For general discussion of instrumentation and protocols, columns and mobile phases, and useful practical aspects, the reader is referred to the studies by Sewell (1992) and Christie and Han (2010).

4.3.1   Detectors for HPLC

It should be noted that routine HPLC analysis of lipids is limited by the lack of a detector that satisfies the following two major requirements: (i) to respond to all lipid molecules and to be sensitive to all molecular types and (ii) not to be affected by changes in mobile phase composition in gradient elution or by temperature.

Ultraviolet spectrophotometric detectors are the most common detectors used for HPLC as they are relatively inexpensive and can give great selectivity and sometimes sensitivity in the analysis of specific compounds. Moreover, they are relatively little influenced by changes in temperature or solvent flow rate. They can also be used in gradient elution mode, although baseline drift can be a problem. UV detectors exhibit the highest sensitivity toward compounds with conjugated double bonds and aromatic rings, neither of which are common in naturally occurring lipids (although in seed oils, fatty acids with conjugated double bonds may present). To overcome this limitation, chemical derivatization to form compounds that absorb in the UV region should be considered. Carotenoids and tocopherols can be conveniently analyzed with UV detection.

Fluorescing compounds can be detected with a high degree of sensitivity and selectivity, and sensitivity can be as much as 1000 times greater than when using UV detector. Also, fluorescent derivatives can be prepared in cases where lipids do not exhibit natural fluorescence.

The refractive index (RI) detector is widely used, and it will respond to any molecule that has a different RI to that of mobile phase. This type of detection has been widely used in lipid analysis, for example, for the analysis of triacylglycerols in a range of vegetable oils and for fatty acid methyl esters using nonaqueous reversed-phase HPLC, and in the size exclusion separation of fatty acids and mono-, di-, and triacylglycerols.

Infrared spectrophotometric detectors have been used in the analysis of nonpolar lipids containing a carbonyl function that absorbs at 5.75 μm (between 1650 and 1860 cm^{-1}).

This type of detector is not sensitive to temperature and flow-rate changes but can show a high level of background noise because most solvents absorb to some extent in the spectral regions of interest. An advantage of this detection is that gradient elution can be used.

With evaporative light-scattering detection (ELCD), or "mass detection," the solvent emerging from the end of the HPLC column is evaporated in a stream of air or nitrogen in a heating chamber, but the solute is not evaporated and is nebulized and passes as minute droplets through a light beam where it is reflected and refracted. The amount of scattered light measured is related to the amount of material in the eluent. There are no special wavelength lamp requirements for the light source: it is simply a projector lamp. This detection can be considered to be universal in its applicability. The sensitivity is high, and it is not affected by temperature changes or by small changes in mobile phase flow rate, but the choice of mobile phases is limited to those that will evaporate in the heated chamber. The detector needs very careful calibration and strict adherence to the calibration conditions. A wide range of lipids, for example, cholesteryl esters, triacylglycerols, fatty acids, and phospholipids, has been reported to be successfully analyzed using ELSD.

5   MASS SPECTROMETRY

The use of mass spectrometry for lipids has expanded rapidly in the last decade. Mass spectrometry has two main uses. First, it is often used to identify (or confirm provisional identities) the structure of particular fatty acids or lipids. Second, it is used for "lipidomics," whereby a detailed analysis of

the molecular species of individual lipid classes is undertaken, often in association with a particular challenge (environmental stress, disease, etc.) to the tissue being examined.

5.1 Types of Mass Spectrometry for Lipid Analysis

These techniques are thoroughly discussed by Christie and Han (2010) and we summarize them briefly here.

Electron impact and chemical ionization techniques are well suited to the analysis of volatile lipids and are often coupled to GCs. For the ionization of relatively polar intact lipids, atmospheric pressure chemical ionization (APCI) is a better method. It is useful for ionization of unesterified fatty acids, phospholipids, sterols, and triacylglycerols (Byrd-well, 2004). Most use has been for triacylglycerols and its low sensitivity for, say, phospholipids mean that its use has been superseded by electrospray ionization-mass spectrometry (ESI-MS).

Fast atom bombardment has been utilized for many polar lipid molecules in the past, both in positive- and negative-ion modes. For a review on its applications to different phospholipids, see the paper by Murphy and Harrison (1994). Again, it has been largely replaced by ESI-MS. Since it was developed in the 1980s, MALDI-TOF/MS (matrix-assisted laser desorption/ionization, time-of-flight MS) has been used to characterize almost all classes of lipids as well as oxidized derivatives (see e.g., Schiller *et al*., 2004). There are a number of problems with MALDI-TOF/MS but efforts to reduce the drawbacks are being made (Christie and Han, 2010).

The most useful technique for lipidomics is electrospray ionization (ESI). It is one of the least destructive methods for molecular analysis. This also makes ESI-MS a very sensitive technique. Furthermore, the reduced in-source fragmentation also helps quantitative aspects of analysis. The method is also highly adaptable allowing it to be linked to HPLC systems with a whole variety of solvents. Virtually all non-volatile lipids can be analyzed by ESI-MS, which therefore complements GLC-MS rather well. Individual molecular species of a given lipid class have very similar ionization efficiencies, thus aiding their quantitation. Finally, specific lipid classes within a mixture can be selectively ionized – termed intra-source separation.

Because soft ionization techniques (such as ESI-MS) cause little in-source fragmentation, analysis usually depends on tandem MS. Thus, the first analyzer is used to select a precursor ion that is accelerated into a collision cell to collide with an inert gas that fragments it into product ions and neutral fragments. These are detected by a second analyzer. There are also various scan modes that can be set up for lipidomics. For reviews of ESI-MS, see those of Murphy, Fielder, and Hevko (2001); Griffiths (2003); Pulfer and Murphy (2003); and Christie and Han (2010).

5.2 Use of Mass Spectrometry to Elucidate Fatty Acid Structures

Many scientists assume that the use of retention times on GLC columns (especially capillary columns) is enough for identification. Although they are an excellent guide, definitive confirmation needs additional methods. Coupling a mass spectrometer to the GLC is one such method. Methyl esters are not particularly good and it is better to use derivatives designed for MS such as pyrrolidines, picolinyl esters, or 4,4′-dimethyloxazolines (DMOX) (Christie, 1997, 1998).

However, methyl esters are adequate for analysis of saturated fatty acids and these derivatives on GLC-MS will usually allow location of methyl branch points or position and type of oxygen groups in a chain. For monounsaturated fatty acids, methyl esters are unsuitable, and for polyunsaturated fatty acids, they need to be used with some caution (Dijkstra, Christie, and Knothe, 2007). For the location of double bond position, any of the three types of derivatives (detailed earlier) should be used. We usually use DMOX derivatives (Fuschino *et al*., 2011). Examples of mass spectra obtained for α-linolenic acid are shown in Figure 2. For discussion of the interpretation of such spectra, see the reviews by Dijkstra, Christie, and Knothe (2007) and Christie and Han (2010).

5.3 Mass Spectrometry Applied to Intact Lipids

There are several basic precautions that need to be born in mind when using ESI-MS (Christie and Han,
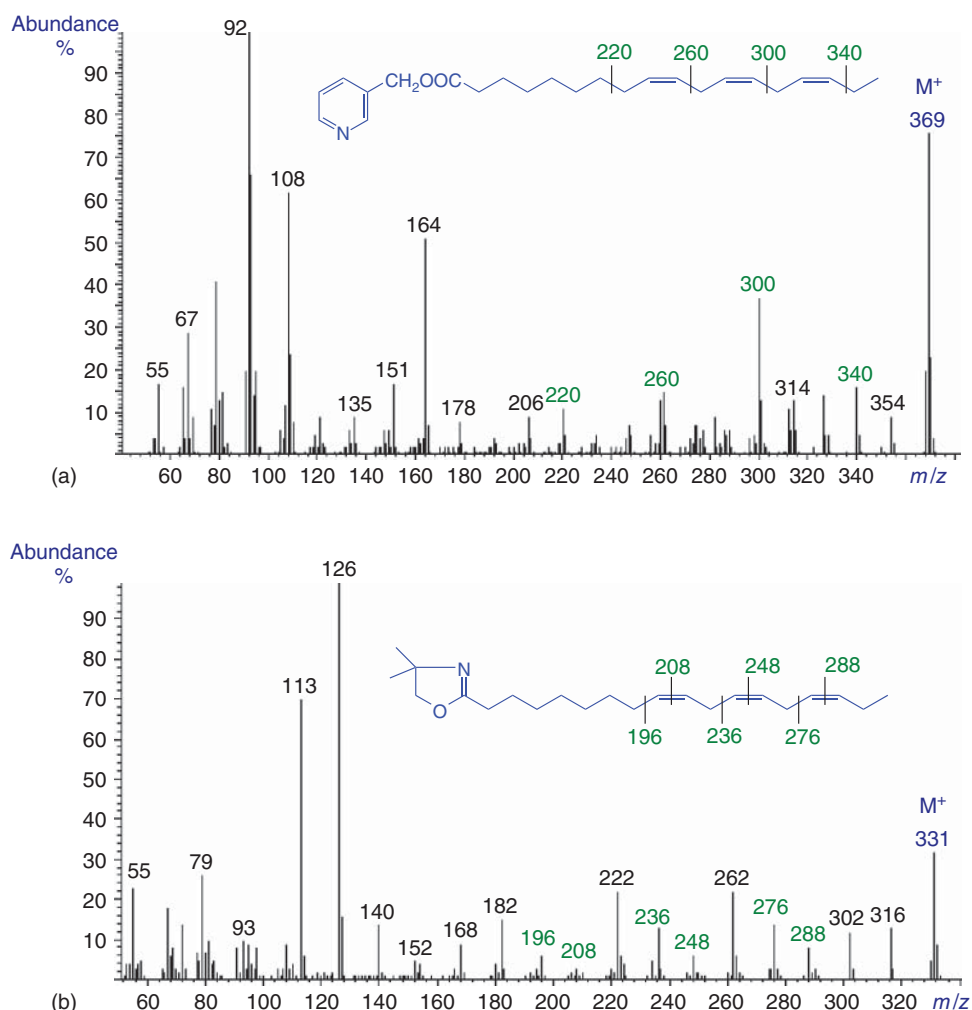
Figure 2   Mass spectra of α-linolenate (9,12,15-octadecatrienoate) in the form of its picolinyl esters (a) and dimethyloxazoline derivative (b). *Source*: Spectra are taken from www.lipidlibrary.aocs.org with permission of the author (W. W. Christie).

2010). For example, the use of collision-induced dissociation (CID) (as is usual) depends very much on the CID conditions and the type of instrument used. The intensities of the product ions also depend on such conditions, making the construction of libraries containing product-ion spectra of individual lipid species not practical. Operation in the positive- or negative-ion modes gives different fragmentation processes and these are used differently for different lipid classes separated by HPLC (online) or previously by other chromatographic techniques. For detailed information about the structural characterization of different lipid classes, refer to the reviews

by Christie and Han (2010) and, especially, Hsu and Turk (2005a,b, 2009).

Examples of mass spectra of plant-derived lipid class species are given in Figure 3. For alternative examples, refer to the Lipid Library (www.lipidlibrary.aocs.org) and "LIPID MAPS" (www.lipidmaps.org).

5.4   Lipidomics

Although separation of lipid classes and their subsequent analysis by ESI-MS gives excellent results,
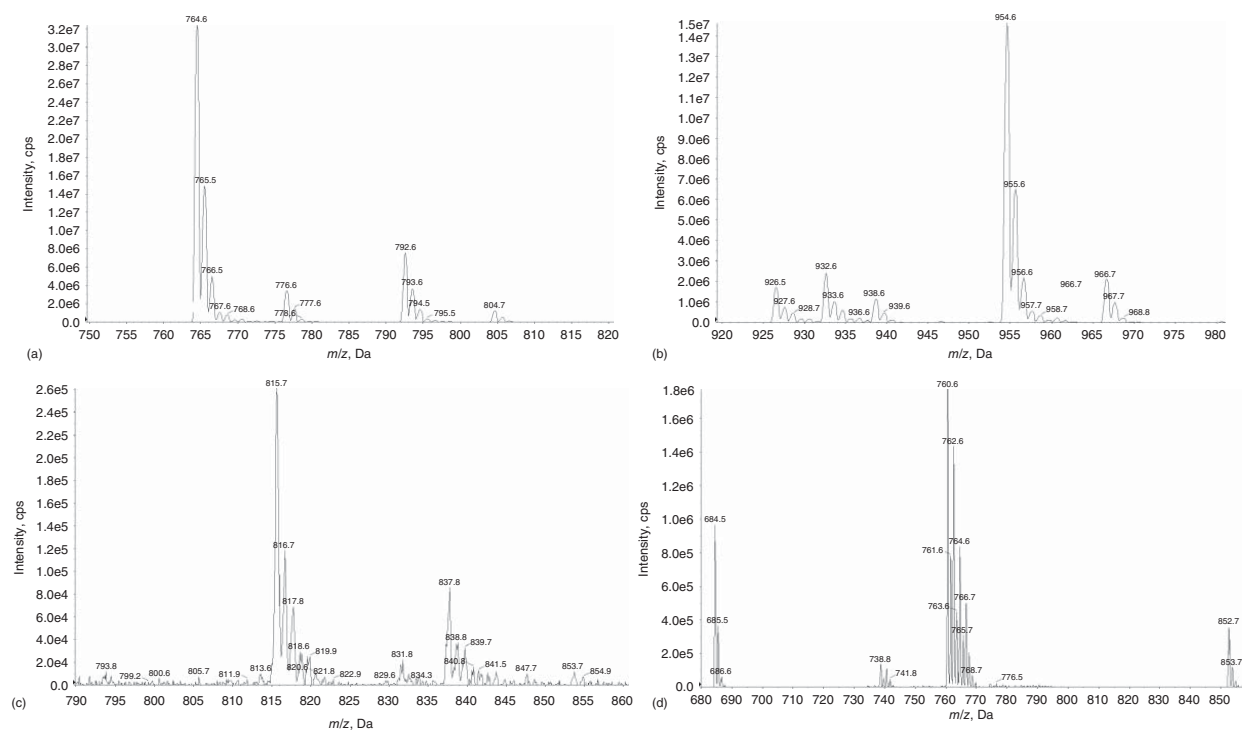
Figure 3 Mass spectrometry of intact plant (*Arabidopsis* leaf) lipids. (a) MGDG. The major peaks correspond to 764.6 (34 : 6) 16 : 3/18 : 3 species; 792.6 (36 : 6) 18 : 3/18 : 3 species. Internal standards at 776.6 (34 : 0) and 804.7 (36 : 0). (b) DGDG. Major peak at 954.6 (36 : 6) 18 : 3/18 : 3 species. Internal standards at 938.6 (34 : 0) and 966.7 (36 : 0). (c) SQDG. Major peaks at 815.7 (34 : 3) 16 : 0/18 : 3 species; 837.8 (36 : 6) 18 : 3/18 : 3 species. No internal standard available. (d) PG. Major peaks at 760.6 (34 : 4) 16 : 1/18 : 3 species; 764.6 (34 : 2) 16 : 0/18:2 species; 766.7 (34 : 1) 16 : 0/18 : 1 species. Internal standards at 684.5 (28 : 0) and 852.7 (40 : 0). *Source*: The spectra were generously supplied by Drs. Sunitka Shive and Ruth Welti of the Kansas Lipidomics Research Centre (Kansas State University) for which the authors are very grateful.

for many purposes, "shotgun lipidomics" using crude lipid extracts is needed. In general, sample sizes of 10 mg wet weight tissues or 200 µg of a protein in a membrane fraction might be needed (Cheng, Jiang, and Han, 2007). If a minor lipid component is of special interest, then more sample will be needed.

A key step is lipid extraction. For shotgun lipidomics, inorganic residues must be minimized. This may involve a "clean up" of the original extract. Detergents should also be avoided. During extraction, a salt is used that matches the adducts preferred for analysis. Lithium chloride is recommended (see Christie and Han, 2010, for discussion).

Comprehensive libraries/databases are available giving important information about structures, masses, isotope patterns, and MS/MS spectra. Two detailed ones are those from LIPID MAPS consortium (www.lipidmaps.org) and METLIN from the Scripps Research Institute (www.metlin.scripps.edu).

In shotgun lipidomics, lipids are sorted into different classes using basic structural blocks. For example, all sphingolipids containing combinations of (i) various sphingoid bases, (ii) fatty acids (attached to the amino moiety of the base), and (iii) the polar carbohydrate moiety (-ies) attached to the terminal hydroxyl of the sphingoid base. These three structures will be revealed by characteristic fragment ions. A characteristic fragment of a class that is associated with the head group is usually visible. This can be used to identify individual species of a class of interest by precursor ion scanning or specific neutral loss (Brugger *et al.*, 1997). All molecular species of an individual lipid class can be analyzed in one analytical run using a crude lipid extract. There are recognized problems (Christie and Han, 2010). Nevertheless, lipidomics has been used successfully in plant biology (Moreau *et al.*, 2008; Welti *et al.*, 2002; Welti *et al.*, 2007). Use of stable isotope labeling with the method allows lipid turnover, biosynthesis, and trafficking to be assessed (e.g., Postle *et al.*, 2007).

One aspect of the use of ESI-MS, in general (including its utilization in lipidomics), that should be born in mind is the issue of quantification. Owing to the chemistry/physics involved in MS, there is not always a clear relationship between ion counts and concentration. The issues of internal standards and normalization are covered by Christie and Han (2010) as are examples of quantification after

HPLC-MS and after direct injection. Particular concerns are ion suppression and the dynamic range. The latter will undoubtedly be improved as new instruments are developed.

Nevertheless, it is clear that ESI-MS/MS has become a method of choice and indispensible for lipidomics.

6 RELATED ARTICLES

Extraction Methodologies: General Introduction; Thin-layer Chromatography, with Chemical and Biological Detection Methods; HPLC and Ultra HPLC: Basic Concepts; New Developments of Laser Desorption Ionization Mass Spectrometry in Plant Analysis; LC and LC-MS: Techniques and Applications.

REFERENCES

Bligh, E. G., and Dyer, W. J. (1959) *Can. J. Biochem. Physiol.*, **37**, 911–917.

Brugger, B., Erben, G., Sandhoff, R., *et al.* (1997) *Proc. Natl. Acad. Sci. U. S. A.*, **94**, 2339–2344.

Byrdwell, W. C. (2004) Review article: APCI-MS in lipid analysis, in *Advances in Lipid Methodology – Five*, ed. R. O. Adlof, Oily Press, Bridgwater, pp. 171–253.

Cheng, H., Guan, S., and Han, X. (2006) *J. Neurochem.*, **97**, 1288–1300.

Cheng, H., Jiang, X., and Han, X. (2007) *J. Neurosci.*, **101**, 57–76.

Christie, W. W. (1997) Structural analysis of fatty acids, in *Advances in Lipid Methodology – Four*, ed. W. W. Christie, Oily Press, Bridgwater, pp. 119–169.

Christie, W. W. (1998) *Lipids*, **32**, 343–353.

Christie, W. W. and Han, X. (2010) *Lipid Analysis*, 4th edn, Bridgwater, Oily Press.

Colborne, A. J. and Laidman, D. L. (1975) *Phytochemistry*, **14**, 2639–2645.

Deas, A. H. B. and Holloway, P. J. (1977) The intracellular structure of some plant cutins, in *Lipids and Lipid Polymers in Higher Plants*, eds. M. Tevini and H. K. Lichtenthaler, Springer-Verlag, Berlin, pp. 293–299.

Dijkstra, A. J., Christie, W. W., and Knothe, G. (2007) Analysis, in *The Lipid Handbook*, 3rd edn, eds. F. D. Gunstone, J. L. Harwood and A. J. Dijkstra, CRC Press, Boca Raton, pp. 415–470.

Evershed, R. P. (1992) Gas chromatography of lipids, in *Lipid Analysis. A Practical Approach*, eds. R. J. Hamilton and S. Hamilton, IRL Press, Oxford, pp. 113–151.

Folch, J., Lees, M., and Sloane-Stanley, G. A. (1957) *J. Biol. Chem.*, **226**, 497–509.

Fuschino, J., Guschina, I. A., Dobson, G., *et al*. (2011) *J. Phycol.*, **47**, 763–774.

Garbus, J., De Luca, H. F., Loomans, H. E., *et al*. (1963) *J. Biol. Chem.*, **238**, 59–63.

Griffiths, W. J. (2003) *Mass Spectrom. Rev.*, **22**, 81–152.

Guiheneuf, F., Ulmann, L., Minouni, V., *et al*. (2013) *Phytochemistry*, **90**, 45–49.

Gunstone, F. D., Harwood, J. L. and Dijkstra, A. J. eds. (2007) *The Lipid Handbook*, 3rd edn, CRC Press, Boca Raton.

Gurr, M. I., Harwood, J. L., and Frayn, K. N. (2002) *Lipid Biochemsitry*, 5th edn, Blackwell Science, Oxford.

Guschina, I. A., and Harwood, J. L. (2006) *Prog. Lipid Res.*, **45**, 160–186.

Guschina, I. A., Kinney, A., Quant, P. A., and Harwood, J. L. 2004. Regulation of lipid biosynthesis in soybean cell cultures. Proceedings of 16th IPLS, Budapest, 1–4 June.

Hamilton, R. J. ed. (1996) *Waxes: Chemistry, Molecular Biology and Functions*, Oily Press, Dundee.

Han, X., Yang, J., Cheng, H., *et al*. (2005) *J. Lipid Res.*, **46**, 1548–1560.

Harwood, J. L. (1980) Plant acyl lipids: structure, distribution and analysis, in *Biochemistry of Plants*, eds. P. K. Stumpf and E. E. Conn, Academic Press, New York, vol. **4**, pp. 1–55.

Harwood, J. L. (1998) What's so special about plant lipids?, in *Plant Lipid Biosynthesis: Fundamentals and Agricultural Applications*, ed. J. L. Harwood, Cambridge University Press, Cambridge, pp. 1–28.

Harwood, J. L., and Jones, A. J. (1989) *Adv. Bot. Res.*, **16**, 1–53.

Heinz, E (1996) Plant glycolipids: structure, isolation and analysis, in *Advances in Lipid Methodology — Three*, ed. W.W. Christie, Oily Press, Dundee, pp. 211–332.

Henderson, R. J., and Tocher, D. R. (1992) Thin-layer chromatography, in *Lipid Analysis. A Practical Approach*, eds. R. J. Hamilton and S. Hamilton, IRL Press, Oxford, pp. 65–111.

Hitchcock, C., and Nichols, B. W. (1971) *Plant Lipid Biochemistry*, Academic Press, London.

Hsu, F.-F., and Turk, J. (2005a) Analysis of sulfatides, in *The Encyclopedia of Mass Spectrometry*, ed. R. M. Caprioli, Elsevier, New York, pp. 473–492.

Hsu, F.-F., and Turk, J. (2005b) Dual parallel liquid chromatography/mass spectrometry for lipid analysis, in *Modern Methods for Lipid Analysis by Liquid Chromatography/Mass Spectrometry and Related Techniques*, ed. W. C. Byrdwell, AOCS Press, Champaign, pp. 510–576.

Hsu, F.-F., and Turk, J. (2009) *J. Chromatogr. B*, **877**, 2673–2695.

Kates, M. (2010) *Techniques in Lipidology*, 3rd edn, Ottawa, Newport Somerville Innovation Ltd..

Kolattukudy, P. E. (1975) Biochemistry of cutin, suberin and waxes on the lipid barriers of plants, in *Recent Advances in the Chemistry and Biochemistry of Plant Lipids*, eds. T. Galliard and E. Mercer, Academic Press, New York, pp. 203–246.

Kolattukudy, P. E. (1977) Biosynthesis and degradation of lipid polymers, in *Lipids and Lipid Polymers in Higher Plants*, eds. M. Tevini and H. K. Lichtenthaler, Springer-Verlag, Berlin, pp. 271–292.

Kolattukudy, P. E. (1980) Cutin, suberin and waxes, in *Biochemistry of Plants*, eds. P. K. Stumpf and E. E. Conn, Academic Press, New York, vol. **4**, pp. 571–645.

Kolattukudy, P. E. (1987) Lipid-derived defensive polymers and waxes and their role in plant-microbial interaction, in *Biochemistry of Plants*, eds. P. K. Stumpf and E. E. Conn, Academic Press, New York, vol. **9**, pp. 291–314.

Kunst, L., and Samuels, A. L. (2003) *Prog. Lipid Res.*, **42**, 51–80.

Moreau, R. A., Doehlert, D. C., Welti, R., *et al*. (2008) *Lipids*, **43**, 533–548.

Morrison, W. R., Tan, S. L., and Hargin, K. D. (1980) *J. Sci. Food Agric.*, **31**, 329–340.

Murphy, D. J. ed. (1994) *Designer Oil Crops*, Weinheim, VCH Press.

Murphy, D. J. ed. (2005) *Plant Lipids: Biology, Utilisation and Manipulation*, Blackwell Publishing, Oxford.

Murphy, R. C., and Harrison, K. A. (1994) *Mass Spectrom. Rev.*, **13**(1), 57–75.

Murphy, D. J., Harwood, J. L., Lee, K. A., *et al*. (1985) *Phytochemistry*, **24**, 1923–1929.

Murphy, R. C., Fielder, J., and Hevko, J. (2001) *Chem. Rev.*, **101**, 479–526.

Osagie, A. U., and Kates, M. (1984) *Lipids*, **19**, 958–965.

Postle, A. D., Wilton, D. C., Hunt, A. N., *et al*. (2007) *Prog. Lipid Res.*, **46**, 200–224.

Pulfer, M., and Murphy, R. C. (2003) *Mass Spectrom. Rev.*, **22**, 332–364.

Ramli, U. S., Baker, D. S., Quant, P. A., *et al*. (2002) *Biochem. J.*, **364**, 385–391.

Schiller, J., Suss, R., Arnhold, J., *et al*. (2004) *Prog. Lipid Res.*, **43**, 449–488.

Sewell, P.A. (1992) High-performance liquid chromatography, in *Lipid Analysis. A practical Approach*, eds. R.J. Hamilton and S. Hamilton, IRL Press, Oxford, pp. 153–203.

Smith, K. L., Douce, R., and Harwood, J. L. (1982) *Phytochemistry*, **21**, 569–573.

Tang, M., Guschina, I. A., O'Hara, P., *et al*. (2012) *New Phytol.*, **196**, 415–426.

Walker, K. A., and Harwood, J. L. (1986) *Biochem. J.*, **237**, 41–46.

Walker, K. A., Ridley, S. M., and Harwood, J. L. (1988) *Biochem. J.*, **254**, 811–817.

Walton, T. J. (1990) Waxes, cutin and suberin, in *Methods in Plant Biochemistry*, eds. J. L. Harwood and J. R. Boyer, Academic Press, London, vol. **4**, pp. 105–158.

Welti, R., Li, W., Li, M., *et al*. (2002) *J. Biol. Chem.*, **277**, 31994–32002.

Welti, R., Shah, J., Li, W., *et al*. (2007) *Front. Biosci.*, **12**, 2494–2506.

HPLC Analysis of Alkaloids

Brás Heleno de Oliveira

Federal University of Paraná, Curitiba, Brazil

1 INTRODUCTION

Alkaloids are very important natural products found in many plants. They are nitrogen-containing compounds, and most are biologically active. Since the description of "morphium" (morphine) in opium by Friedrich Sertürner in 1804 until the more recent discoveries of the antitumor vincristine and vinblastine, many alkaloids with important medicinal use have been described. Some toxic alkaloids for humans and animals have also been described, and others with psychotropic properties constitute important health and legal problem worldwide. Owing to their practical importance, analytical procedures are required for analysis of alkaloids in plant materials.

The analysis of alkaloids by chromatographic methods has been the subject of a few books and book chapters (Flieger, 2011; McCalley, 2008; Popl, Faehnrich, and Tatar, 1990; Ranta, Callaway, and Naaranlahti, 1994; Verpoorte and Baerheim, 1984), in addition to journal reviews (Kabulov, D'Yakonov, and Zalyalieva, 1991; Petruczynik, 2012; Stockigt *et al*., 2002; Verpoorte and Niessen, 1994). They discuss the available literature on various alkaloid classes. Another review contains examples of alkaloids analysis using HPLC-MS (high performance liquid chromatography mass spectrometry) of natural products (Wolfender, Maillard, and Hostettmann, 1994). Other reviews deal with specific alkaloid classes and will be reviewed accordingly.

This chapter shows HPLC analysis of pharmacologically useful alkaloids and of some with toxic and psychotropic properties. It is organized according to alkaloid classes, and examples are described for each class. Most of the information is in tabular form with sufficient data, useful as a starting point for the practical analyst. For basic aspects of HPLC technique (see **HPLC and Ultra HPLC: Basic Concepts**).

2 GENERAL ANALYTICAL PROCEDURES

2.1 Extraction

The simplest extraction procedure involves the use of an appropriate solvent, usually with sonication (Acevska *et al*., 2012b). Considering that most alkaloids are basic compounds, matrix pH has important effect on extraction efficiency. Therefore, in these cases, a preliminary treatment with a base may increase efficiency of extraction with organic solvent. A typical extraction protocol would involve (i) basification of matrix using, for example, NH_4OH; (ii) extraction with suitable organic solvent; and (iii) evaporation and resolubilization in HPLC mobile phase (Garcia *et al*., 2005). This procedure produces alkaloid-rich extracts.

A more selective procedure would involve a multistep acid–base extraction. A typical protocol would involve (i) preparation of a crude extract using a suitable solvent; (ii) liquid–liquid extraction using dilute mineral acid; (iii) neutralization of the acid extract and liquid–liquid extraction with a suitable solvent; and (iv) evaporation and resolubilization in HPLC mobile phase (Lee *et al*., 2007). The use of

an internal standard is advisable as the multistep procedure may compromise recovery. A report compared sample preparation procedures, including acid extraction followed by solid phase extraction (SPE) cleaning, methanol (MeOH) extraction, and Soxhlet extraction (Hosch *et al.*, 1996). More information on extraction techniques can be found on **Extraction Methodologies: General Introduction**; **Supercritical Fluid Extraction**, and **Solid-Phase Microextraction (SPME) and Its Application to Natural Products**.

2.2 Detection and Derivatization

Depending on detection system available, the alkaloid analyte may need derivatization. The compound may not produce a signal with sufficient sensitivity for analysis. This may be due to the lack of a chromophore (UV/vis detector), fluorophore (fluorescence detector), or electrochemical activity (electrochemical detector). Conversion of the alkaloid into a suitable derivative may overcome the limitation. Derivatization may also be useful for the separation of alkaloid enantiomers, which can be previously converted to the corresponding diastereomers (Lee *et al.*, 2008). A few examples of derivatization are shown in Table 1.

2.3 HPLC Analysis

Most alkaloids analyses have been carried out with alkyl silica stationary phases, especially octadecylsilane. Mobile phases used are usually buffers with an organic modifier, MeOH or acetonitrile (ACN). The pH of mobile phase may have significant effect on separation because of the ionizable nature of alkaloids and the possible presence of free silanol groups on silica surface. At acidic pH, the ionization

of silanol groups is avoided, which improves peak shape; however, it may also protonate the alkaloid nitrogen, changing retention times. When basic pH is chosen, the alkaloid molecules are maintained as free bases, but silanol groups may be ionized and the alkylsilane bond may hydrolyze. These problems explain the fact that most analyses are carried out with acid buffers. Nearly all these issues can be addressed by careful choice of mobile phase pH and by modern stationary phase materials (Theodoridis *et al.*, 1995).

Specific examples of analytical conditions for various classes of alkaloids are shown in the following sections. Table 2 facilitates the search for a specific alkaloid class. It shows the alkaloid classes covered in this chapter, their basic structures or representative examples, and the corresponding tables with HPLC analysis data.

3 TRUE ALKALOIDS

True alkaloids are derived from amino acids, and the nitrogen atom participates in a heterocyclic ring. Examples of HPLC analysis of pyridine, piperidine, isoquinoline, aporphine, indole, and imidazole alkaloids are shown in the following sections.
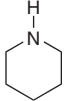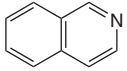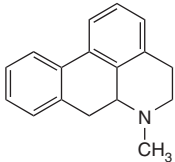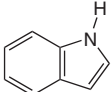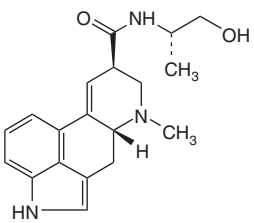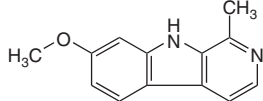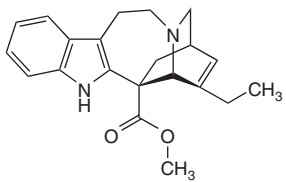
3.1 Pyridine

Alkaloids of this class are found in the plants of the families Leguminosae, Apocynaceae, Rubiaceae, Chenopodiaceae, and Euphorbiaceae. General aspects of pyridine alkaloids have been reviewed previously (O'Hagan, 2000). An HPLC review on the comparison of reversed phase columns for the separation of tobacco alkaloids is also available (McCalley, 1993). Examples of HPLC analysis of pyridine alkaloids are shown in Table 3.

Table 1　Examples of derivatization procedures for HPLC analysis of alkaloids.

Alkaloid class	Analyte	Reagent	Detection	References
Pyrrolizidine	Adonifoline	*o*-Chloranil	223 nm	Xiong *et al.* (2009)
Steroid	α-Solanine, α-chaconine	Tris(2,2′-bipyridine) Ru(III)	Chemiluminescence	Kodamatani *et al.* (2005)
Piperidine	1-Deoxynojirimycin	9-Fluorenylmethyl chloroformate	FLD	Kim *et al.* (2003)
Piperidine	Ammodendrine, anabasine, coniine	*N*-(9-Fluorenylmethoxycarbonyl)-L-alanine	MS	Lee *et al.* (2008)

Table 2 Alkaloid classes and respective HPLC examples data.

Class	Basic structure or example	Table
True alkaloids		
Pyridine	Pyridine	Table 3
Piperidine	Piperidine	Table 4
Isoquinoline	Isoquinoline	Table 5
Aporphine	Aporphine	Table 6
Indole, simple	Indole	Table 7
Indole, ergot	Ergobasine	Table 8
Indole, β-carboline	Harmine	Table 9
Indole, terpenoid	Cantharanthine	Table 10

(Continued Overleaf)

Table 2   (*Continued*)
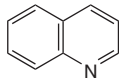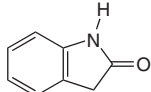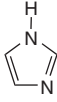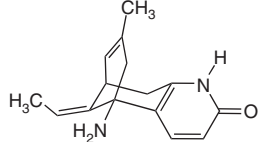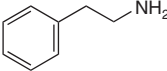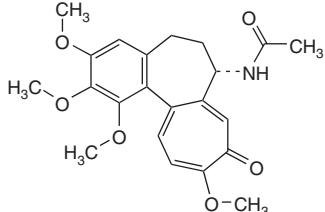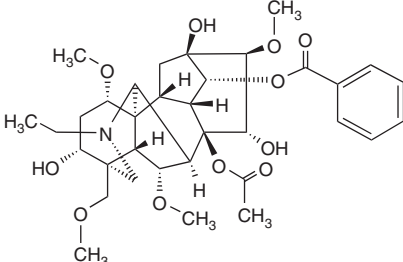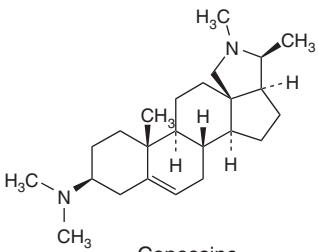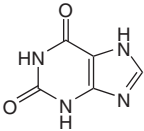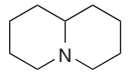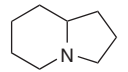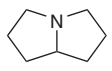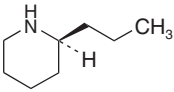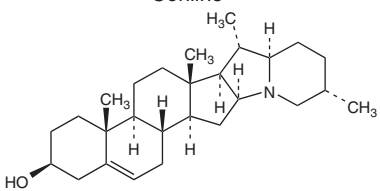
Class	Basic structure or example	Table
Indole, quinoline	 Quinoline	Table 11
Indole, oxindole	 Oxindole	Table 12
Imidazole	 Imidazole	Table 13
Lycodine	 Huperzine A	Table 14
Protoalkaloids		
Phenylethylamine	 Phenylethylamine	Table 15
Colchicine	 Colchicine	Table 16
Pseudoalkaloids		
Terpenoid	 Aconitine	Table 17
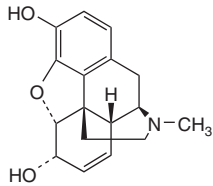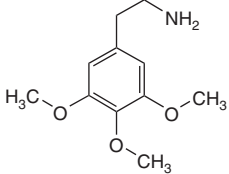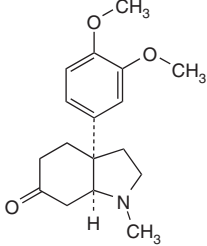
Table 2 (*Continued*)

Class	Basic structure or example	Table
Steroid	Conessine	Table 18
Xanthine	Xanthine	Table 19
Toxic alkaloids		
Quinolizidine	Quinolizidine	Table 20
Indolizidine	Indolizidine	Table 21
Pyrrolizidine	Pyrrolizidine	Table 22
Piperidine	Coniine	Table 23
Steroid	Solanidine	Table 24
Psychotropic alkaloids		
Tropane	Tropane	Table 25

(*Continued Overleaf*)

Table 2 (*Continued*)

Class	Basic structure or example	Table
Opium	 Morphine	Table 26
Pyrrolidine	 Pyrrolidine	Table 27
Phenylethylamine	 Mescaline	Table 28
Mesembrine	 Mesembrine	Table 29

Table 3 HPLC conditions for the analysis of selected pyridine alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Nicotine enantiomers	Standards	Chiral-AGP (100×4.0 mm)	Phosphate buffer: MeOH or ACN (gradient)	254 nm	NA	Demetriou *et al*. (1993)
Nicotine and nornicotine	Standards	Chiralcel OJ (250×4.6 mm)	Hexane: MeOH, EtOH, or iPropOH (80–99:20–0.925%)	254 and 265 nm	NA	Tang, Zielinski, and Bigott (1998)
Trigonelline and nicotinic acid	*Coffea arabica* and *C. canephora* (roast beans)	Spherisorb S5 ODS2 (250×4.6 mm)	0.1 M phosphate buffer (pH 4.0): MeOH	268, 264, and 276 nm	Extraction boiling H_2O	Casal, Oliveira, and Ferreira, (2000)
Ricinine	*Ricinus communis* (seeds)	(Luna, 50×2 mm)	0.05% HCO_2H: 0.05% HCO_2H in MeOH (gradient)	254 nm	Extraction 2% AcOH	Ovenden *et al*. (2010)

Table 4 HPLC conditions for the analysis of selected piperidine alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Synephrine, arecoline, and norisoboldine	Pharmaceutical product	Agilent SCX (250×4.6 mm)	0.2% H_3PO_4:MeOH	215 nm	Filtration	Yi *et al*. (2012)
Lobeline	*Lobelia inflate* (whole plant)	Eurosphere 100-C8 (250×3 mm)	0.1% TFA:ACN (70:30)	250 nm	Extraction 0.1 N HCl:MeOH (1:1), sonication, SPE	Kursinszki, Ludányi, and Szoke (2008)
Piperanine, piperdardine, piperine, etc.	*Piper nigrum* (berries)	Inertsil ODS-3v (250×4.0 mm)	0.5% HCO_2H:ACN (gradient)	280, 340 nm; MS	Extraction 80% EtOH, sonication	Friedman *et al*. (2008)

3.2 Piperidine

Plants of the families Palmae, Lobeliaceae, and Piperaceae are the source of some piperidine alkaloids. Examples of HPLC analysis of these alkaloids are shown in Table 4.

3.3 Isoquinoline

Isoquinoline alkaloids are found in many plant families such as Papaveraceae, Fumariaceae, Berberidaceae, Ranunculaceae, Rubiaceae, Menispermaceae, Leguminosae, and Amaryllidaceae. General aspects of isoquinolines have been reviewed (Bentley, 2006). HPLC analytical methods for this alkaloid class have also been reviewed (László *et al*., 2010). Another review is available on protoberberine alkaloids with a section on analytical methods (Grycova, Dostal, and Marek, 2007). Examples of HPLC analysis of isoquinoline alkaloids are shown in Table 5.

3.4 Aporphine

Aporphine alkaloids are found in Monimiaceae plants. General aspects of these alkaloids have been reviewed (Guinaudeau, Leboeuf, and Cave, 1975, 1979). Examples of HPLC analysis of aporphine alkaloids are shown in Table 6.

3.5 Indole

The structural variation of indole alkaloids is considerable, and the number of compounds of this class found in plants is significant. They may be distributed in subclasses such as simple indole, ergot alkaloids, β-carboline, terpenoid indole, quinoline, and oxindole. HPLC methods for the analysis of indole alkaloids have been reviewed (Anna, 2010).

3.5.1 Simple Indole

A series of reviews on simple indole alkaloids is available (Ishikura, Yamada, and Abe, 2010). Examples of HPLC analysis of simple indole alkaloids are shown in Table 7.

3.5.2 Ergot

Alkaloids of this class are produced by fungi of the genus *Claviceps* that infect some grains. A review on general aspects, including analytical methods, has been published (Flieger, Wurst, and Shelby, 1997). Another review on analysis of this alkaloid class is also available (Scott, 2007). Examples of HPLC analysis of ergot alkaloids are shown in Table 8.

3.5.3 β-Carboline

General aspects of β-carboline alkaloids have been reviewed (Allen and Holmstedt, 1980; Stuart and Wooming, 1975), in addition to their biological properties (Cao *et al*., 2007). A survey of analytical methods, including topics on sample preparation, is also available (Herraiz, 2000). Examples of HPLC analysis of β-carboline alkaloids are shown in Table 9.

Table 5 HPLC conditions for the analysis of selected isoquinoline alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Various	*Neolitsea sericea* (leaves)	LiChroCART (250 × 4.6 mm)	0.1% TFA:ACN (0.1% TFA) (gradient)	280 nm	Acid–base extraction	Lee *et al.* (2007)
Protopine, allocryptopine, sanguinarine, and chelerythrine	*Macleaya cordata* (fruits)	Spherigel C8 (150 × 4.6 mm)	AcONH$_4$:ACN (60:40)	MS	Extraction 0.1 M HCl, microwave	Luo, Chen, and Yao (2006)
Protopine, chelidonine, coptisine, etc.	*Chelidonium majus* (aerial parts)	Luna C18(2) (250 × 4.6 mm)	30 mM HCO$_2$NH$_4$ (pH 2.80): ACN:MeOH (67.3:14.7:18)	280 nm	Extraction HCl in MeOH, sonication, SPE	Kursinszki *et al.* (2006)
Berberine, palmatine, coptisine, etc.	*Coptis chinensis* (whole plant)	Dikma Diamonsil C18 (250 × 4.6 mm)	3.4 mM AcONH$_4$ and 0.2% AcOH, ACN (gradient)	MS, 277 nm	Extraction 75% EtOH	Wu *et al.* (2005)
Morphine, codeine, etc.	*Papaver somniferum* (capsules, stems)	Zorbax Extend C-18 (250 × 4.6 mm)	0.1% TFA (pH 9.6 with TEA):MeOH (gradient)	280 nm	Extraction MeOH, sonication	Acevska *et al.* (2012b)
Emetine and cephaeline	*Psychotria ipecacuanha* (roots)	Shympac CLC-ODS (M), (250 × 4.6 mm)	0.25 M NaOAc (pH 5):ACN (9:5)	288 nm	Basification; extraction EtOEt	Garcia *et al.* (2005)
Galanthamine	*Leucojum aestivum* (bulbs, tissue culture)	Platinium C18, (250 × 4.6 mm)	30 mM TEA acetate:ACN (gradient)	280 nm	Lyophilization, extraction with MeOH (sonication)	Diop *et al.* (2006)
Galanthamine	Various Amaryllidaceae species (bulbs)	Vydac C18, 201SP54 (250 × 4.6 mm)	TFA:ACN:water (0.01:5:95)	210 nm	Extraction with 0.1% TFA (vortex, sonication)	Mustafa, Rhee, and Verpoorte (2003)

Table 6 HPLC conditions for the analysis of selected aporphine alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Boldine	*Lindera aggregata* (roots)	Acquity UPLC BEH C18 (50 × 2.1 mm)	10 mM NH$_4$OAc pH 3:ACN (gradient)	MS	MeOH extr (sonication)	Han *et al.* (2008)
Magnoflorine	*Nigella sativa* (seeds)	Gemini C18 (150 × 4.6 mm)	0.1% AcOH:ACN (gradient)	260 nm; MS	Extraction MeOH, sonication or microwave	Avula *et al.* (2010)
Dicentrine	*Spirospermum penduliflorum* (leaves)	C18 LiChroCART (250 × 4 mm)	20 mM HCO$_2$NH$_4$ pH 3.0:MeOH (gradient)	307 nm, MS	Extraction 1% AcOH in MeOH, etc.	Rafamantanana *et al.* (2012)

3.5.4 Terpenoid Indole

Important examples of this class are the anticancer compounds, vinblastine and vincristine. General aspects of terpenoid indole alkaloids have been reviewed (van der Heijden *et al.*, 2004). Examples of HPLC analysis of terpenoid alkaloids are shown in Table 10.

3.5.5 Quinoline

These alkaloids are found in Rubiaceae and Rutaceae plant families. A review on this alkaloid class is available (Michael, 2008a). HPLC analysis of cinchona alkaloids has been reviewed (McCalley, 2002). Two surveys compared various packing materials for the analysis of *Cinchona* alkaloids

Table 7 HPLC conditions for the analysis of selected simple indole alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Serotonin and melatonin	*Coffea canephora* (various plant tissues)	Atlantis dC18 (150 × 3.9 mm)	0.1 M AcONa, 0.1 M citric acid, 0.5 mM SOS, 0.15 M EDTA, pH 3.7:MeOH (95:5)	MS	Extraction Tris buffer	Ramakrishna *et al.* (2012)
Tryptamine and other biogenic amines	Food samples	C18 ODS Hypersil (250 × 4.5 mm)	ACN:H$_2$O:MeOH (60:25:15)	254 nm	Acid–base extraction , derivatization dansyl chloride	Tameem *et al.* (2010)
Serotonin, tryptamine, and tyramine	Tomatoes, cherry tomatoes, and peppers	Atlantis C18 (150 × 3.9 mm)	0.3% TFA:MeOH (95:5)	280 nm	Extraction MeOH, SPE	Ly *et al.* (2008)

Table 8 HPLC conditions for the analysis of selected ergot alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Ergometrine, ergosine, ergotamine, etc.	Cereals	XBridge MS C18 (150 × 2.1 mm)	NH$_4$HCO$_2$ buffer, MeOH (gradient)	MS/MS	Extraction EtOAc/MeOH/0.2 M NH$_4$HCO$_2$	Di Mavungu *et al.* (2012)
Ergometrine, ergotamine, ergosine, etc.	Cereals, processed foods	Gemini, C18 (150 × 2 mm)	3.03 mM (NH$_4$)$_2$CO$_3$, ACN (gradient)	MS/MS	Dispersive SPE	Krska *et al.* (2008)

Table 9 HPLC conditions for the analysis of selected β-carboline alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Harman and norharman	*Coffea arabica* and *C. canephora* (roast beans)	Tracer Excel ODSA (250 × 4 mm)	0.03 M formate buffer (pH 3.0):ACN (gradient)	Fluorescence	Brew, SPE	Alves, Casal, and Oliveira (2007)
Harmalol, harmaline, harmane, etc.	Foods (beer, coffee, cheese)	Ultrabase C18 (150 × 4.6 mm)	50 mM Na$_2$HPO$_4$ pH 9.0:MeOH:ACN (60:20:20)	ECD	Extraction solvent extr, SPE	Agui *et al.* (2007)
β-Carboline and derivatives	*Passiflora incarnata* and *Tribulus terrestris* (whole plant)	Shim-pack CLC-C8 (M) (250 × 4.6 mm)	ACN:TFA:H$_2$O (18.0:0.2:81.8)	FLD	Multistep extraction	Tsuchiya, Shimizu, and Iinuma (1999)

(Hoffmann, Lammerhofer, and Lindner, 2009; Theodoridis *et al.*, 1995). Examples of HPLC analysis of quinoline alkaloids are shown in Table 11.

3.5.6 *Oxindole*

These alkaloids are found mainly in *Uncaria* sp. (Rubiaceae). The sources and properties of this alkaloids class have been surveyed (Yeoh, Chan, and Morsingh, 1967). Examples of HPLC analysis of quinoline alkaloids are shown in Table 12.

3.6 Imidazole

The literature on this alkaloid class has been reviewed (Jin, 2011). Examples of HPLC analysis of quinoline alkaloids are shown in Table 13.

3.7 Lycodine

Lycodine is an important class of alkaloids found in *Lycopodium* sp. because of their inhibitory property

Table 10   HPLC conditions for the analysis of selected terpenoid alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Vinblastine, vincristine, etc.	*Catharanthus roseus* (hairy roots)	Zorbax Eclipse XDB-C (250 × 4.6 mm)	5 mM Na$_2$ HPO$_4$ (pH 6 with H$_3$PO$_4$):ACN (gradient)	220–330 nm; fluorescence	Extraction MeOH, sonication	Tikhomiroff and Jolicoeur (2002)

Table 11   HPLC conditions for the analysis of selected quinoline alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Cinchonine, cinchonidine, quinidine, and quinine	*Cinchona succirubra* (bark)	Microsorb-MV C8 (250 × 4.6 mm)	50 mM KH$_2$PO$_4$, 30 mM hexylamine pH 2.8, ACN (94:6)	316 nm	Extraction DCM	Fabiano-Tixier *et al.* (2011)
Camptothecin	*Camptotheca acuminata* (leaves)	XTerra C18 (150 × 2.0 mm)	0.05% TFA:ACN (0.05% TFA) (gradient)	254 nm; MS/MS	Extraction 30% EtOH, sonication	Montoro *et al.* (2010)
Cinchonine, cinchonidine, dihydrocinchonine, etc.	*Cinchona succirubra* (bark)	Prodigy ODS (250 × 3.2 mm)	0.05 M TEAP (pH 3.0), ACN (88:12)	FLD	Extraction MeOH	Gatti, Gioia, and Cavrini (2004)
Quinamine, corynantheal, strictosidine, etc.	*Cinchona ledgeriana* and *C. robusta* (shoot cultures)	Bondapak C18 column (300 × 3.9 mm)	100 mM NH$_4$HCO$_2$:ACN (gradient)	UV and MS	Not provided	Giroud *et al.* (1991)

Table 12   HPLC conditions for the analysis of selected oxindole alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Speciophylline, uncarine F, mitraphylline, etc.	*Uncaria tomentosa* (bark)	Zorbax XDB C18 (150 × 4.6 mm)	35 mM TEAA, pH 6.9:ACN (gradient)	245 nm	Extraction 60% EtOH, sonication, SPE	Bertol, Franco, and de Oliveira (2012)
Rhynchophylline, isorhynchophylline, corynoxeine, and isocorynoxeine	*Uncaria rhynchophylla* (hooks and leaves)	Kromasil C18 (200 × 4.6 mm)	0.03% DEA in MeOH, 0.03% DEA in H$_2$O (gradient)	245 nm, MS	Extraction MeOH:H$_2$O (70:30)	Qu *et al.* (2012)

Table 13   HPLC conditions for the analysis of selected imidazole alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Pilocarpine	*Pilocarpus microphyllus* (leaves)	Inertsil ODS-2 (250 × 4.6 mm)	0.05 M (NH$_4$)$_2$OAc pH 4:ACN (gradient)	MS/MS	Acid–base extraction	Sawaya *et al.* (2008)

against acetylcholinesterase, which is a target for the treatment of Alzheimer's disease. A review on *Lycopodium* alkaloids is available (Ma and Gang, 2004). The most studied in this class is huperzine A, and examples of its HPLC analysis are shown in Table 14.

4 PROTOALKALOIDS

The biosynthesis of these compounds also starts from amino acids, but the nitrogen atom does not belong to a heterocyclic ring. Some examples of HPLC analysis of phenylethylamine and colchicine alkaloids are shown below.

4.1 Phenylethylamine

A series of reviews on these compounds is available (Bentley, 2006). The analysis of this alkaloid class in *Citrus aurantium* L. (Rutaceae) by chromatographic and electrophoretic methods has also been reviewed (Pellati and Benvenuti, 2007). Examples of HPLC analysis of phenylethylamine alkaloids are shown in Table 15.

4.2 Colchicine

A survey on general aspects of this alkaloid class has been published (Husek, 1992). Examples of HPLC analysis of colchicine alkaloids are shown in Table 16.

5 PSEUDO ALKALOIDS

These compounds are not derived from amino acids. Some examples of HPLC analysis of terpenoid, steroid, and xanthine alkaloids are shown below.

5.1 Terpenoid

An important example of this class is the anticancer compound paclitaxel (Taxol®). Examples of HPLC analysis of terpenoid alkaloids are shown in Table 17.

5.2 Steroid

A series of reviews on steroid alkaloids is available (Atta ur and Choudhary, 1999). Examples of HPLC analysis of steroid alkaloids are shown in Table 18.

Table 14 HPLC conditions for the analysis of selected lycodine alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Huperzine A	*Huperzia serrata* (whole plant)	Alltima C18 (250 × 4.6 mm)	MeOH:80 mM NH$_4$OAc, pH 6.0 (32:68)	308 nm	Extraction with aqueous tartaric acid, then CHCl$_3$ (after basification)	Ma *et al.* (2005)
Huperzine A	*Huperzia selago* (whole plant)	Symmetry C18 column (250 × 4 mm)	MeOH:10 mM NH$_4$OAc, pH 3.5 (gradient)	308 nm, MS	Basification with conc NH$_4$OH, extraction with CHCl$_3$	Borloz, Marston, and Hostettmann (2006)
Huperzine A	Various *Huperzia* spp. and *Lycopodium* spp. (whole plant)	Inertsil, Japan ODS-3 (250 × 4.6 mm)	MeOH:0.01% TFA (gradient)	308 nm	Extraction MeOH	Sahidan *et al.* (2012)

Table 15 HPLC conditions for the analysis of selected phenylethylamine alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Synephrine	*Citrus aurantium* (leaves, fruits)	Nova-Pak C-18 (150 × 3.9 mm)	ACN:H$_2$O:TFA (5:95:0.01):ACN (gradient)	220 nm	Extraction MeOH	Arbo *et al.* (2008)

Table 16   HPLC conditions for the analysis of selected colchicine alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Colchicine and demecolcine	*Colchicum crocifolium*, *C. ritchii*, and *C. triphyllum* (various plant tissues)	LiChroCART (125 × 4 mm), Purospher1 STAR RP-18.	ACN:3% AcOH (gradient)	245 nm	Basification and extraction DCM	Alai, Tawaha, and El-Elimat (2007)

Table 17   HPLC conditions for the analysis of selected terpenoid alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Aconitine, mesaconitine, and hypaconitine	*Aconitum carmichaelii* (roots)	Zorbax Eclipse XDB-C8 (150 × 4.6 mm)	10 mM $(NH_4)_2OAc$ (pH 8.9):MeOH (gradient)	233 nm, MS	Extraction MeOH-H_2O, MeOH-$CHCl_3$	Csupor *et al.* (2009)
Paclitaxel	*Taxus baccata* (needles)	Symmetry C8 (150 × 4.6 mm)	20 mM AcONH4 (pH 5 with AcOH):MeCN: MeOH:THF (55:39:4:2).	227 nm	Acid–base extraction	Hook *et al.* (1999)
Paclitaxel and other taxanes	*Taxus brevifolia* and *T. baccata* (needles)	Chromegabond pentafluorophenyl (250 × 4.0 mm) or Chromegabond phenyl (150 × 4.6 mm)	2 mM NH_4OAc pH 6.9):ACN (gradient)	230 nm, MS	Extraction MeOH	Kerns *et al.* (1994)

Table 18   HPLC conditions for the analysis of selected steroid alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Pseudojervine, veratrosine, jervine, veratramine, etc.	*Veratrum dahuricum* (whole plant)	Zorbax Extend C18 (250 × 4.6 mm)	Water, 0.1% HCO_2H, and 0.04% ammonia (25%, wt/wt, vol/vol):ACN (gradient)	ELSD and MS	Extraction MeOH and sonication	Tang *et al.* (2008)
Conessine	*Holarrhena antidysenterica* (bark)	Bondapak C18 (250 × 4.6 mm)	MeOH:H_2O (95:5) and ACN:H_2O (both with 0.1% AcOH)	RID and MS	Acid–base extraction	Garg and Bhutani (2008)

5.3   Xanthine

The biosynthesis and metabolism of xanthine alkaloids has been reviewed (Ashihara, Sano, and Crozier, 2008). Reviews on the HPLC analysis of this class of alkaloids have also been published (de Sena, de Assis, and Branco, 2011; Horie and Kohata, 2000). Examples of HPLC analysis of xanthine alkaloids are shown in Table 19.

6   TOXIC ALKALOIDS

Some of the alkaloid classes shown above may also contain toxic compounds. The piperidine alkaloids, lobeline and coniine, are good examples. The former has been indicated for the treatment of tobacco and other drug addictions (Damaj *et al.*, 1997), while the latter is an important teratogen (Green *et al.*, 2012). A review shows examples of plant toxic alkaloids found in the food chain (Koleva *et al.*,

Table 19 HPLC conditions for the analysis of selected xanthine alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Caffeine, theophylline, and theobromine	*Sideritis raeseri* (aerial parts)	PerfectSil Target ODS-3 HD (250×4.6 mm)	MeOH:ACN (95:5):0.01% AcOH (gradient)	270–370 nm	Extraction H_2O and sonication	Samanidou, Tsagiannidis, and Sarakatsianos (2012)
Caffeine and theobromine	*Camellia sinensis* (leaves)	RP-MAX (250×4.6 mm)	ACN:1% HCO_2H (gradient)	250–700 nm and MS	Extraction hot H_2O	Del Rio *et al.* (2004)
Caffeine, theobromine, and theacrine	*Camellia sinensis*, *C. assamica*, *C. ptilophylla*, and *C. assamica* (leaves)	Discovery RP-Amide C16 (150×4.6 mm)	(85%) H_3PO_4 in H_2O (0.05:99.95):ACN (gradient)	210 and 280 nm	Extraction hot H_2O	Peng *et al.* (2008)
Caffeine and theobromine	*Theobroma cacao* (fermented beans)	NOVAPAK C18, (150×3.9 mm)	20% MeOH in water	274 nm	Defatting and extraction hot H_2O	Brunetto *et al.* (2007)

2012). Another review describes some classes of toxic alkaloids including the steroidal, piperidine, indolizidine, pyrrolizidine, and nortropane alkaloids (Molyneux *et al.*, 2007).

6.1 Quinolizidine

The general aspects of quinolizidine and indolizidine alkaloids have been reviewed (Michael, 2008b).

Examples of HPLC analysis of quinolizidine alkaloids are shown in Table 20.

6.2 Indolizidine

A review with general aspects of indolizidine and quinolizidine alkaloids is available (Michael, 2008b). Examples of HPLC analysis of indolizidine alkaloids are shown in Table 21.

Table 20 HPLC conditions for the analysis of selected quinolizidine alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Magnoflorine	*Caulophyllum thalictroides* (dietary supplement)	Acquity UPLC BEH Shield RP18 (50×2.1 mm)	50 mM NH_4OAc:ACN (gradient)	320 nm	Extraction of various solvents and sonication	Avula *et al.* (2011)
9α-Hydroxymatrine, matrine, sophoridine, etc.	*Sophora alopecuroides* (seed, legume, stem, root)	Kromasil C18, 250×4.6 mm)	ACN:0.1% H_3PO_4 + 0.1% TEA (gradient)	205 nm	Extraction MeOH, sonication	Gao, Li, and Wang (2011)
Matrine, oxymatrine, sophocarpine, etc.	*Sophora flavescens* (roots)	Capcell Pak-C18 (150×4.6 mm)	20 mM NH_4OAc, pH 8.0:MeOH (gradient)	MS/MS	Extraction H_2O, sonication	Liu *et al.* (2011)
Baptifoline, anagyrine, and *N*-methylcytisine	*Caulophyllum thalictroides* (roots)	Synergi Max-RP (150×4.6 mm)	10 mM NH_4OAc, pH 8.0:ACN (gradient)	310 nm; ELSD	Extraction MeOH, sonication	Ganzera *et al.* (2003)
N-Methylcytisine, baptifoline, anagyrine, and magnoflorine	*Caulophyllum thalictroides* (roots)	Econosphere ODS (150×4.6 mm)	MeOH:10 mM NaH_2PO_4 pH 7.5 (75:25)	230 nm	Extraction 70% EtOH, SPE	Woldemariam, Betz, and Houghton (1997)

Table 21 HPLC conditions for the analysis of selected indolizidine alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Swainsonine	*Undifilum oxytropis* (mycelia)	XBridge (150 × 4.6 mm)	ACN:5 mM NH$_4$OAc (1:1)	ELSD	Extraction MeOH	Yang *et al.* (2012)
Several indole and indolizidine	*Nauclea officinalis* (twigs)	Diamonsil C18 (4.6 × 250 mm).	0.1% AcOH:ACN (gradient)	MS	Acid–base extraction	Li *et al.* (2011)

6.3 Pyrrolizidine

Pyrrolizidine alkaloids have been found in some plant species, many are toxic to livestock and humans. Food may also be contaminated by these compounds (Koleva *et al.*, 2012), and they are even found in honey (Kempf *et al.*, 2011). General aspects of alkaloids of this class have been reviewed (Roeder and Wiedenfeld, 2009, 2011; Roeder, 1995, 2000).

The analytical methods for alkaloids of this class have been reviewed (Crews, Berthiller, and Krska, 2010; Roeder, 1999). A review of HPLC analysis of this alkaloid class is also available (Cao, Colegate, and Edgar, 2008). Examples of HPLC analysis of pyrrolizidine alkaloids are shown in Table 22.

6.4 Piperidine

Some piperidine alkaloids are also toxic (teratogenic) to livestock and humans (Green *et al.*, 2012; Lee *et al.*, 2008). Examples of HPLC analysis of toxic piperidine alkaloids are shown in Table 23.

6.5 Steroid

General aspects of steroid alkaloids can be found in a series of reviews (Atta ur and Choudhary, 1999). The toxic properties of some *Solanum* and *Veratrum* steroid alkaloids have been evaluated (Crawford and Kocan, 1993; Crawford and Myhr, 1995). Examples

Table 22 HPLC conditions for the analysis of selected pyrrolizidine alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Senecionine, senkirkin, petasin, etc.	*Petasites hybridus* (dietary supplement)	SynergiMAX RP (150 × 4.6 mm)	0.1% HCO$_2$H:0.1% HCO$_2$H in ACN (gradient)	TOF-MS	Extraction MeOH, sonication	Avula *et al.* (2012)
Retrorsine and senecionine	*Packera candidissima*; *P. bellidifolia* (whole plant)	XTerra RP18 (150 × 4.6 mm)	15 mM NH$_4$OH:ACN (gradient)	214 nm	Extraction MeOH	Fragoso-Serrano *et al.* (2012)
Lycopsamine, echimidine, and lasiocarpine	*Symphytum officinale* (whole plant)	Luna C18 (100 × 2.0 mm)	0.1% HCO$_2$H:0.1% HCO$_2$H in ACN (gradient)	MS	Reflux MeOH:H$_2$O (1:1)	Liu *et al.* (2009)
Adonifoline, clivorine, isoline, etc.	*Senecio scandens* (whole plant)	Polaris C18-A (250 × 4.6 mm)	1% HCO$_2$H:1% HCO$_2$H in ACN (35:65)	MS	Acid–base extraction	Zhang *et al.* (2007)
3′-*O*-Acetylintermedine, leptanthine-*N*-oxide, etc.	*Echium plantagineum I* (leaves, flowers)	Aqua C8 (150 × 2.1 mm)	0.1% HCO$_2$H:0.1% HCO$_2$H in ACN (gradient)	MS	Extraction MeOH, SPE	Colegate *et al.* (2005)
Senecionine, Seneciphylline, and spartioidine	*Senecio leucophyllus* (whole plant)	LiChrosphe 60 RP-select B (125 × 4 mm and 250 × 4 mm)	10% ACN (with hexanesulphonic acid):80% ACN	220 nm	Extraction MeOH or HCl, SPE	Hosch *et al.* (1996)

Table 23 HPLC conditions for the analysis of selected toxic piperidine alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Ammodendrine, anabasine, and coniine	*Lupinus sulphureus* (aerial parts)	Betasil C18 (100 × 2.1 mm)	20 m*M* NH$_4$OAc:MeOH (45:55)	MS	Acid–base extraction, derivatization	Lee *et al*. (2008)

of HPLC analysis of toxic steroid alkaloids are shown in Table 24.

7 PSYCHOTROPIC ALKALOIDS

This section contains examples of psychoactive alkaloids of the tropane, opium, pyrrolidine, phenylethylamine, and mesembrine classes.

7.1 Tropane

General aspects of tropane alkaloids have been reviewed (O'Hagan, 2000). Another review discusses the distribution of tropane alkaloids in Solanaceae, Erythroxylaceae, Proteaceae, Euphorbiaceae, Rhizophoraceae, Convolvulaceae, and Brassicaceae plant families (Griffin and Lin, 2000). A survey of tropane alkaloids useful in medicine is also available (Grynkiewicz and Gadzikowska, 2008).

General chromatographic methods for analysis of tropane alkaloids has been reviewed (Drager, 2002). Another review describes chromatographic methods for their analysis on coca leaves (Moore and Casale, 1994). HPLC and CE methods have also been surveyed (Aehle and Drager, 2010; Tomasz, 2010).

Examples of HPLC analysis of tropane alkaloids are shown in Table 25.

7.2 Opium

The latex of opium poppy (*Papaver somniferum* L.) is the most common source of alkaloids of this class. A review describes general aspects of opium alkaloids (Kalant, 1997). HPLC analytical procedures for these alkaloids have also been surveyed (Janicot, Caude, and Rosset, 1986). Examples of HPLC analysis of opium alkaloids are shown in Table 26.

7.3 Pyrrolidine

Pyrrolidine alkaloids have no pharmacological use, but they have been found in toxic plants. Hygrines, for example, have been found in Erythroxylaceae, which also contains tropane alkaloids. Stachydrine, on the other hand, has been found in Lamiaceae and Leguminosae. The general aspects of pyrrolidine alkaloids have been reviewed (O'Hagan, 2000; Ohagan, 1997; Pinder, 1984, 1986, 1987, 1989a, b, 1990; Plunkett, 1994). Examples of HPLC analysis of pyrrolidine alkaloids are shown in Table 27.

Table 24 HPLC conditions for the analysis of selected toxic steroid alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
α-Solanine, α-chaconine, β-chaconine, etc.	*Solanum* sp. (leaves and tubers)	LiChrosph 100 RP-18 (250 × 4 mm)	0.1% HCO$_2$H:0.1% HCO$_2$H in ACN	MS	Lyophilization, extraction 1% AcOH (sonication), SPE.	Distl and Wink (2009)
Verdine, pseudojervine, veratrosine, etc.	*Veratrum nigrum* (roots, rhizomes)	Venusil XBP-C18 (200 × 4.6 mm)	ACN:0.03% TEA (gradient)	ELSD	Extraction H$_2$O, MeOH	Cong *et al*. (2008)
Solanine, solanidine, and solasodine	*Solanum lyratum* (cell culture)	Cosmosil 5 C18-AR-II (250 × 4.6 mm)	ACN:TEA phosphate buffer (gradient)	205 nm	Extraction MeOH	Kuo, Chao, and Lu (2012)

Table 25   HPLC conditions for the analysis of selected tropane alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Hyoscyamine and scopolamine	*Datura innoxia*, *Atropa belladonna* (hairy roots)	Luna C18 (250 × 4.6 mm)	ACN:MeOH:30 mM phosphate buffer pH 6.00 (12:7.9:80.1).	210 nm	Solvent extraction, SPE	Kursinszki *et al.* (2005)
Yosciamine and scopolamine	*Datura* sp. (leaves, fruits)	XTerra C18 (150 × 4.6 mm)	ACN:15 mM NH$_4$OH (gradient)	205 nm	PLE, SPE	Mroczek, Glowniak, and Kowalska (2006)
Cocaine	*Erythroxylum coca* (leaves)	C8 (125 × 4.6 mm)	ACN:1.0% TEA, pH 4 (40:60)	240 nm	Reflux 95% EtOH	Glass and Johnson (1993)
Scopolamine and others	*Datura ferox* (seeds)	Bondapak C18 (250 × 4.6 mm)	0.2% H$_3$PO$_4$ pH 7.3 with di-*m*-propylamine: MeOH (60:40)	254 nm	Acid–base extraction, SPE	Vitale, Acher, and Pomilio (1995)

Table 26   HPLC conditions for the analysis of selected opium alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Morphine, codeine, papaverine, and thebaine	*Papaver somniferum* (capsules, stems)	Zorbax Extend C-18 (250 × 4.6 mm)	0.1% TFA, pH 9.6 using TEA:MeOH (gradient)	280 nm	Extraction MeOH, sonication	Acevska *et al.* (2012a)
Morphine, codeine, thebaine, and papaverine	*Papaver somniferum*, *P. setigerum*, *P. bracteatum*, *P. pseudo-orientale* (capsules)	TSKgel ODS 120A (250 × 4.6 mm)	ACN:10 mM Na 1-heptanesulphonate (pH 3.5 with H$_3$PO$_4$), gradient	284 nm	Extraction 0.1 M Na citrate buffer, SPE	Yoshimatsu *et al.* (2005)

Table 27   HPLC conditions for the analysis of selected pyrrolidine alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Hygrine	*Datura ferox* (seeds)	Bondapak C18 (250 × 4.6 mm)	0.2% H$_3$PO$_4$ pH 7.3 with di-*m*-propylamine: MeOH (60:40)	254 nm	Acid–base extraction	Vitale, Acher, and Pomilio (1995)
Cocaine, cuscohygrine, etc.	*Erythroxylum coca* (seeds)	Phenomenex C18 (150 × 4.6 mm)	0.1% TFA:ACN (gradient).	MS	Acid–base extraction	Casale, Toske, and Colley (2005)
Stachydrine	*Citrus* sp. (juices)	Discovery C8 (150 × 3.0 mm)	0.1% HCO$_2$H in water	MS/MS	SPE	Servillo *et al.* (2011)

7.4 Phenylethylamine

This alkaloid class has been the subject of a series of reviews (Bentley, 2006). Examples of HPLC analysis of phenylethylamine alkaloids are shown in Table 28.

7.5 Mesembrine

A general review containing a section on the chemistry of crinane alkaloids of *Sceletium* spp. has been published (Gericke and Viljoen, 2008). Examples of

Table 28 HPLC conditions for the analysis of selected phenylethylamine alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
Mescaline, N-methylmescaline, and hordenine	*Lophophora williamsii, L. diffusa* (whole plant)	Spherisorb ODS-1 (150 × 4.6 mm)	ACN-H_2O (108:892) with 5.0 mL H3PO4, 0.28 mL hexylamine	205 nm	Defatting, extraction methanol: ammonia 33% (99:l)	Helmlin, Bourquin, and Brenneisen (1992)

Table 29 HPLC conditions for the analysis of selected mesembrine alkaloids.

Alkaloid	Sample (source)	Column	Mobile phase	Detection	Sample preparation	References
14 Crinane and 14 tazettine alkaloids	*Crinum latifolium, C. asiaticum* (various plant tissues)	TSK-GEL ODS-100V (150 × 4.6 mm)	ACN:H_2O with 0.2% AcOH, 40 mM AcONH$_4$	MS	Extraction 95% EtOH	Zhang *et al.* (2009)
Δ^7 Mesembrenone, mesembranol, mesembrenone, mesembrine, and epimesembranol	*Sceletium* sp. (whole plant)	Luna C18 (2) (150 × 4.6 mm); Hypersil C18 (150 × 4.6 mm)	H_2O:ACN:25% NH_4OH (70:30:0.01)	228 nm	NA	Patnala and Kanfer (2010)

HPLC analysis of pyrrolidine alkaloids are shown in Table 29.

8 CONCLUSION

The number of reports on the HPLC of alkaloids is significant, which reflects the importance of the technique for the analysis of these compounds. This review showed that alkaloids of all major classes have been successfully analyzed by HPLC. Therefore, this technique should be considered for the analysis of plant materials containing this important class of natural products.

LIST OF ABBREVIATIONS

AChE	acetylcholinesterase
ACN	acetonitrile
AcOH	acetic acid
DCM	dichloromethane
DEA	diethylamine
ECD	electrochemical detector
EDTA	ethylenediaminetetraacetic acid
ELSD	evaporative light-scattering detector
EtOH	ethanol
FLD	fluorescence detector
HCO_2H	formic acid
iPropOH	isopropanol
MS	mass spectrometry
MeOH	methanol
PLE	pressurized liquid extraction
RID	refractive index detector
SPE	solid phase extraction
SOS	sodium octanesulfonate
TEA	triethylamine
TEAA	triethylammonium acetate
TEAP	triethylammonium phosphate
TFA	trifluoracetic acid
THF	tetrahydrofuran
TOF-MS	time-of-flight mass spectrometry
UV	ultraviolet

REFERENCES

Acevska, J., Stefkov, G., Petkovska, R., *et al.* (2012a) *Anal. Bioanal. Chem.*, **403**, 1117–1129.

Acevska, J., Dimitrovska, A., Stefkov, G., *et al.* (2012b) *J. AOAC Int.*, **95**, 399–405.

Aehle, E., and Drager, B. (2010) *J. Chromatogr. B*, **878**, 1391–1406.

Agui, L., Pena-Farfal, C., Yanez-Sedeno, P., *et al.* (2007) *Anal. Chim. Acta*, **585**, 323–330.

Alai, F. Q., Tawaha, K., and El-Elimat, T. (2007) *Pharmazie*, **62**, 739–742.

Allen, J. R. F., and Holmstedt, B. R. (1980) *Phytochemistry*, **19**, 1573–1582.

Alves, R. C., Casal, S., and Oliveira, B. P. P. (2007) *J. Agr. Food Chem.*, **55**, 1832–1838.

Anna, P. (2010) *HPLC of Indole Alkaloids. High Performance Liquid Chromatography in Phytochemical Analysis*, CRC Press, Boca Raton, FL, pp. 731–767.

Arbo, M. D., Larentis, E. R., Linck, V. M., *et al.* (2008) *Food Chem. Toxicol.*, **46**, 2770–2775.

Ashihara, H., Sano, H., and Crozier, A. (2008) *Phytochemistry*, **69**, 841–856.

Atta ur, R., and Choudhary, M. I. (1999) *Nat. Prod. Rep.*, **16**, 619–635.

Avula, B., Wang, Y. H., Ali, Z., *et al.* (2010) *J. AOAC Int.*, **93**, 1778–1787.

Avula, B., Wang, Y. H., Rumalla, C. S., *et al.* (2011) *J. Pharmaceut. Biomed.*, **56**, 895–903.

Avula, B., Wang, Y. H., Wang, M., *et al.* (2012) *J. Pharmaceut. Biomed. Anal.*, **70**, 53–63.

Bentley, K. W. (2006) *Nat. Prod. Rep.*, **23**, 444–463.

Bertol, G., Franco, L., and de Oliveira, B. H. (2012) *Phytochem. Anal.*, **23**, 143–151.

Borloz, A., Marston, A., and Hostettmann, K. (2006) *Phytochem. Anal.*, **17**, 332–336.

Brunetto, M. R., Gutiérrez, L., Delgado, Y., *et al.* (2007) *Food Chem.*, **100**, 459–467.

Cao, R., Peng, W., Wang, Z., *et al.* (2007) *Curr. Med. Chem.*, **14**, 479–500.

Cao, Y., Colegate, S. M., and Edgar, J. A. (2008) *Phytochem. Anal.*, **19**, 526–533.

Casal, S., Oliveira, M. B., and Ferreira, M. A. (2000) *Food Chem.*, **68**, 481–485.

Casale, J. F., Toske, S. G., and Colley, V. L. (2005) *J. Forensic Sci.*, **50**, 1402–1406.

Colegate, S. M., Edgar, J. A., Knill, A. M., *et al.* (2005) *Phytochem. Anal.*, **16**, 108–119.

Cong, Y., Zhou, Y. B., Chen, J., *et al.* (2008) *J. Pharmaceut. Biomed.*, **48**, 573–578.

Crawford, L., and Kocan, R. M. (1993) *Toxicol. Lett.*, **66**, 175–181.

Crawford, L., and Myhr, B. (1995) *Food Chem. Toxicol.*, **33**, 191–194.

Crews, C., Berthiller, F., and Krska, R. (2010) *Anal. Bioanal. Chem.*, **396**, 327–338.

Csupor, D., Wenzig, E. M., Zupko, I., *et al.* (2009) *J. Chromatogr. A*, **1216**, 2079–2086.

Damaj, M. I., Patrick, G. S., Creasy, K. R., *et al.* (1997) *J. Pharmacol. Exp. Ther.*, **282**, 410–419.

Del Rio, D., Stewart, A. J., Mullen, W., *et al.* (2004) *J. Agr. Food Chem.*, **52**, 2807–2815.

Demetriou, D., Rustemeier, K., Voncken, P., *et al.* (1993) *Chirality*, **5**, 300–302.

Di Mavungu, J. D., Malysheva, S. V., Sanders, M., *et al.* (2012) *Food Chem.*, **135**, 292–303.

Diop, M. F., Ptak, A., Chretien, F., *et al.* (2006) *Nat. Prod. Commun.*, **1**, 475–479.

Distl, M., and Wink, M. (2009) *Potato Res.*, **52**, 79–104.

Drager, B. (2002) *J. Chromatogr. A*, **978**, 1–35.

Fabiano-Tixier, A. S., Elomri, A., Blanckaert, A., *et al.* (2011) *Int. J. Mol. Sci.*, **12**, 7846–7860.

Flieger, J. (2011) *HPLC of alkaloids from the other biosynthetic groups*, CRC Press, Boca Raton, FL.

Flieger, M., Wurst, M., and Shelby, R. (1997) *Folia Microbiol.*, **42**, 3–29.

Fragoso-Serrano, M., Figueroa-González, G., Castro-Carranza, E., *et al.* (2012) *J. Nat. Prod.*, **75**, 890–895.

Friedman, M., Levin, C. E., Lee, S. U., *et al.* (2008) *J. Agr. Food Chem.*, **56**, 3028–3036.

Ganzera, M., Dharmaratne, H. R. W., Nanayakkara, N. P. D., *et al.* (2003) *Phytochem. Anal.*, **14**, 1–7.

Gao, H. Y., Li, G. Y., and Wang, J. H. (2011) *J. Chromatogr. B*, **879**, 1121–1125.

Garcia, R. M. A., de Oliveira, L. O., Moreira, M. A., *et al.* (2005) *Biochem. Syst. Ecol.*, **33**, 233–243.

Garg, S., and Bhutani, K. K. (2008) *Phytochem. Anal.*, **19**, 323–328.

Gatti, R., Gioia, M. G., and Cavrini, V. (2004) *Anal. Chim. Acta*, **512**, 85–91.

Gericke, N., and Viljoen, A. M. (2008) *J. Ethnopharmacol.*, **119**, 653–663.

Giroud, C., Vanderleer, T., Vanderheijden, R., *et al.* (1991) *Planta Med.*, **57**, 142–148.

Glass, R. L., and Johnson, E. L. (1993) *J. Liq. Chromatogr.*, **16**, 3543–3555.

Green, B. T., Lee, S. T., Panter, K. E., *et al.* (2012) *Food Chem. Toxicol.*, **50**, 2049–2055.

Griffin, W. J., and Lin, G. D. (2000) *Phytochemistry*, **53**, 623–637.

Grycova, L., Dostal, J., and Marek, R. (2007) *Phytochemistry*, **68**, 150–175.

Grynkiewicz, G., and Gadzikowska, M. (2008) *Pharmacol. Rep.*, **60**, 439–463.

Guinaudeau, H., Leboeuf, M., and Cave, A. (1975) *Lloydia*, **38**, 275–338.

Guinaudeau, H., Leboeuf, M., and Cave, A. (1979) *J. Nat. Prod.*, **42**, 325–360.

Han, Z., Zheng, Y. L., Chen, N., *et al.* (2008) *J. Chromatogr. A*, **1212**, 76–81.

van der Heijden, R., Jacobs, D. I., Snoeijer, W., *et al.* (2004) *Curr. Med. Chem.*, **11**, 607–628.

Helmlin, H. J., Bourquin, D., and Brenneisen, R. (1992) *J. Chromatogr.*, **623**, 381–385.

Herraiz, T. (2000) *J. Chromatogr. A*, **881**, 483–499.

Hoffmann, C. V., Lammerhofer, M., and Lindner, W. (2009) *Anal. Bioanal. Chem.*, **393**, 1257–1265.

Hook, I., Poupat, C., Ahond, A., *et al.* (1999) *Phytochemistry*, **52**, 1041–1045.

Horie, H., and Kohata, K. (2000) *J. Chromatogr. A*, **881**, 425–438.

Hosch, G., Wiedenfeld, H., Dingermann, T., *et al.* (1996) *Phytochem. Anal.*, **7**, 284–288.

Husek, A. (1992) *Chem. Listy*, **86**, 445–460.

Ishikura, M., Yamada, K., and Abe, T. (2010) *Nat. Prod. Rep.*, **27**, 1630–1680.

Janicot, J. L., Caude, M., and Rosset, R. (1986) *Analusis*, **14**, 441–455.

Jin, Z. (2011) *Nat. Prod. Rep.*, **28**, 1143–1191.

Kabulov, B. D., D'Yakonov, A. L., and Zalyalieva, S. V. (1991) *Chem. Nat. Compd.*, **27** 521–540.

Kalant, H. (1997) *Addiction*, **92**, 267–277.

Kempf, M., Wittig, M., Reinhard, A., *et al.* (2011) *Food Addit. Contam. Part A Chem.*, **28**, 332–347.

Kerns, E. H., Volk, K. J., Hill, S. E., *et al.* (1994) *J. Nat. Prod.*, **57**, 1391–1403.

Kim, J. W., Kim, S. U., Lee, H. S., *et al.* (2003) *J. Chromatogr. A*, **1002**, 93–99.

Kodamatani, H., Saito, K., Niina, N., *et al.* (2005) *J. Chromatogr. A*, **1100**, 26–31.

Koleva, I. I., van Beek, T. A., Soffers, A., *et al.* (2012) *Mol. Nutr. Food Res.*, **56**, 30–52.

Krska, R., Stubbings, G., Macarthur, R., *et al.* (2008) *Anal. Bioanal. Chem.*, **391**, 563–576.

Kuo, C. I., Chao, C. H., and Lu, M. K. (2012) *Phytochem. Anal.*, **23**, 400–404.

Kursinszki, L., Hank, H., Laszlo, I., *et al.* (2005) *J. Chromatogr. A*, **1091**, 32–39.

Kursinszki, L., Sarkozi, A., Kery, A., *et al.* (2006) *Chromatographia*, **63**, S131–S135.

Kursinszki, L., Ludányi, K., and Szoke, É. (2008) *Chromatographia*, **68**, S27–S33.

László, K., Hajnalka, H., Ágnes, K., *et al.* (2010) *HPLC of Isoquinoline Alkaloids. High Performance Liquid Chromatography in Phytochemical Analysis*, CRC Press, Boca Raton, FL, pp. 769–801.

Lee, S. S., Lai, Y. C., Chen, C. K., *et al.* (2007) *J. Nat. Prod.*, **70**, 637–642.

Lee, S. T., Gardner, D. R., Chang, C. W. T., *et al.* (2008) *Phytochem. Anal.*, **19**, 395–402.

Li, Q., Zhang, Y. F., Wu, B., *et al.* (2011) *Eur. J. Mass Spectrom.*, **17**, 277–286.

Liu, F., Wan, S. Y., Jiang, Z., *et al.* (2009) *Talanta*, **80**, 916–923.

Liu, G. Q., Dong, J., Wang, H., *et al.* (2011) *J. Pharmaceut. Biomed.*, **54**, 1065–1072.

Luo, X. B., Chen, B., and Yao, S. Z. (2006) *Phytochem. Anal.*, **17**, 431–438.

Ly, D., Kang, K., Choi, J. Y., *et al.* (2008) *J. Med. Food*, **11**, 385–389.

Ma, X. Q., and Gang, D. R. (2004) *Nat. Prod. Rep.*, **21**, 752–772.

Ma, X. Q., Tan, C. H., Zhu, D. Y., *et al.* (2005) *J. Agr. Food Chem.*, **53**, 1393–1398.

McCalley, D. V. (1993) *J. Chromatogr.*, **636**, 213–220.

McCalley, D. V. (2002) *J. Chromatogr. A*, **967**, 1–19.

McCalley, D. V. (2008) Liquid chromatographic separations of basic compounds, in *Advances in Chromatography*, ed. E. G. N. Grushka, **46**, pp. 305–350.

Michael, J. P. (2008a) *Nat. Prod. Rep.*, **25**, 166–187.

Michael, J. P. (2008b) *Nat. Prod. Rep.*, **25**, 139–165.

Molyneux, R. J., Lee, S. T., Gardner, D. R., *et al.* (2007) *Phytochemistry*, **68**, 2973–2985.

Montoro, P., Maldini, M., Piacente, S., *et al.* (2010) *J. Pharmaceut. Biomed.*, **51**, 405–415.

Moore, J. M., and Casale, J. F. (1994) *J. Chromatogr. A*, **674**, 165–205.

Mroczek, T., Glowniak, K., and Kowalska, J. (2006) *J. Chromatogr. A*, **1107**, 9–18.

Mustafa, N. R., Rhee, I. K., and Verpoorte, R. (2003) *J. Liq. Chromatogr. Relat. Technol.*, **26**, 3217–3233.

Ohagan, D. (1997) *Nat. Prod. Rep.*, **14**, 637–651.

O'Hagan, D. (2000) *Nat. Prod. Rep.*, **17**, 435–446.

Ovenden, S. P. B., Gordon, B. R., Bagas, C. K., *et al.* (2010) *Aust. J. Chem.*, **63**, 8–21.

Patnala, S., and Kanfer, I. (2010) *J. Pharm. Pharm. Sci.*, **13**, 558–570.

Pellati, F., and Benvenuti, S. (2007) *J. Chromatogr. A*, **1161**, 71–88.

Peng, L., Song, X. H., Shi, X. G., *et al.* (2008) *J. Food Compos. Anal.*, **21**, 559–563.

Petruczynik, A. (2012) *Cent. Eur. J. Chem.*, **10**, 802–835.

Pinder, A. R. (1984) *Nat. Prod. Rep.*, **1**, 225–230.

Pinder, A. R. (1986) *Nat. Prod. Rep.*, **3**, 171–180.

Pinder, A. R. (1987) *Nat. Prod. Rep.*, **4**, 527–537.

Pinder, A. R. (1989a) *Nat. Prod. Rep.*, **6**, 515–521.

Pinder, A. R. (1989b) *Nat. Prod. Rep.*, **6**, 67–78.

Pinder, A. R. (1990) *Nat. Prod. Rep.*, **7**, 447–455.

Plunkett, A. O. (1994) *Nat. Prod. Rep.*, **11**, 581–590.

Popl, M., Faehnrich, J., and Tatar, V. (1990) *Chromatographic Analysis of Alkaloids*, Chromatographic Science Series, Dekker, New York, vol. **53**.

Qu, J. L., Gong, T. X., Ma, B., *et al.* (2012) *Chem. Pharm. Bull.*, **60**, 23–30.

Rafamantanana, M. H., Debrus, B., Raoelison, G. E., *et al.* (2012) *J. Pharmaceut. Biomed.*, **62**, 23–32.

Ramakrishna, A., Giridhar, P., Sankar, K. U., *et al.* (2012) *Acta Physiol. Plant*, **34**, 393–396.

Ranta, V. P., Callaway, J. C., and Naaranlahti, T. (1994) *Mod. Methods Plant Anal.*, **15**, 91–114.

Roeder, E. (1995) *Pharmazie*, **50**, 83–98.

Roeder, E. (1999) *Curr. Org. Chem.*, **3**, 557–576.

Roeder, E. (2000) *Pharmazie*, **55**, 711–726.

Roeder, E., and Wiedenfeld, H. (2009) *Pharmazie*, **64**, 699–716.

Roeder, E., and Wiedenfeld, H. (2011) *Pharmazie*, **66**, 637–647.

Sahidan, N., Choo, C. Y., Latiff, A., *et al.* (2012) *Chin. J. Nat. Med.*, **10**, 125–128.

Samanidou, V., Tsagiannidis, A., and Sarakatsianos, I. (2012) *J. Sep. Sci.*, **35**, 608–615.

Sawaya, A., Abreu, I. N., Andreazza, N. L., *et al.* (2008) *Molecules*, **13**, 1518–1529.

Scott, P. (2007) *Mycotoxin Res.*, **23**, 113–121.

de Sena, A. R., de Assis, S. A., and Branco, A. (2011) *Food Technol. Biotechnol.*, **49**, 413–423.

Servillo, L., Giovane, A., Balestrieri, M. L., *et al.* (2011) *J. Agr. Food Chem.*, **59**, 9410–9416.

Stockigt, J., Sheludko, Y., Unger, M., *et al.* (2002) *J. Chromatogr. A*, **967**, 85–113.

Stuart, K., and Wooming, R. (1975) *Heterocycles*, **3**, 223–264.

Tameem, A. A., Saad, B., Makahleh, A., *et al.* (2010) *Talanta*, **82**, 1385–1391.

Tang, Y. B., Zielinski, W. L., and Bigott, H. M. (1998) *Chirality*, **10**, 364–369.

Tang, J., Li, H. L., Shen, Y. H., *et al.* (2008) *Chromatographia*, **67**, 15–21.

Theodoridis, G., Papadoyannis, I., Hermanslokkerbol, A., *et al.* (1995) *Chromatographia*, **41**, 153–160.

Tikhomiroff, C., and Jolicoeur, M. (2002) *J. Chromatogr. A*, **955**, 87–93.

Tomasz, M. (2010) *HPLC of Tropane Alkaloids. High Performance Liquid Chromatography in Phytochemical Analysis*, CRC Press, Boca Raton, FL, pp. 803–821.

Tsuchiya, H., Shimizu, H., and Iinuma, M. (1999) *Chem. Pharm. Bull.*, **47**, 440–443.

Verpoorte, R., and Baerheim, S. A. (1984) *Chromatography of Alkaloids, Pt. B: Gas–Liquid Chromatography and High-Performance Liquid Chromatography*, Journal of Chromatography Library, Elsevier, Amsterdam, vol. **23**.

Verpoorte, R., and Niessen, W. M. A. (1994) *Phytochem. Anal.*, **5**, 217–232.

Vitale, A. A., Acher, A., and Pomilio, A. B. (1995) *J. Ethnopharmacol.*, **49**, 81–89.

Woldemariam, T. Z., Betz, J. M., and Houghton, P. J. (1997) *J. Pharmaceut. Biomed.*, **15**, 839–843.

Wolfender, J. L., Maillard, M., and Hostettmann, K. (1994) *Phytochem. Anal.*, **5**, 153–182.

Wu, W., Song, F. R., Yan, C. Y., *et al.* (2005) *J. Pharmaceut. Biomed.*, **37**, 437–446.

Xiong, A. Z., Yang, L., Zhang, F., *et al.* (2009) *Biomed. Chromatogr.*, **23**, 665–671.

Yang, G. D., Gao, R., Wang, Y., *et al.* (2012) *Toxicon*, **60**, 44–49.

Yeoh, G. B., Chan, K. C., and Morsingh, F. (1967) *Rev. Pure Appl. Chem.*, **17**, 49–66.

Yi, Y. N., Cheng, X. M., Liu, L. A., *et al.* (2012) *Pharm. Biol.*, **50**, 832–838.

Yoshimatsu, K., Kiuchi, F., Shimomura, K., *et al.* (2005) *Chem. Pharm. Bull.*, **53**, 1446–1450.

Zhang, F., Wang, C. H., Xiong, A. Z., *et al.* (2007) *Anal. Chim. Acta*, **605**, 94–101.

Zhang, X., Huang, H., Liang, X., *et al.* (2009) *Rapid Commun. Mass Sp.*, **23**, 2903–2916.

22

Identification and Characterization of Hydroxycinnamates of Six *Galium* Species from the Rubiaceae Family

Rakesh Jaiswal, Marius F. Matei, Sagar Deshpande and Nikolai Kuhnert

Jacobs University Bremen, School of Engineering and Science, Chemistry, Bremen, Germany

1 INTRODUCTION

Hydroxycinnamates form a large class of low molecular weight secondary plant metabolites, which are believed to be involved in UV protection, UV sensing, and defense from herbivores and pathogens in plants. Some compounds have also been reported to provide reproductive advantage as attractants of pollinators and seed dispersers.

Chlorogenic acids (CGAs) and cinnamoyl-hexoses are hydroxycinnamic acid derivatives, which are ubiquitous plant metabolites. Classically, CGAs are a family of esters formed between quinic acid and certain *trans*-cinnamic acids, most commonly caffeic, *p*-coumaric, and ferulic acid (Clifford, 1999, 2000; IUPAC, 1976; Kuhnert, Karakoese, and Jaiswal, 2012); sinapic acid and dimethoxycinnamic acid are also present in certain plant species (Clifford *et al.*, 2006a; Jaiswal and Kuhnert, 2010; Jaiswal *et al.*, 2010a,b; Jaiswal, Deshpande, and Kuhnert, 2011a; Jaiswal, Kiprotich, and Kuhnert, 2011b; Jaiswal and Kuhnert, 2011a,b, c; Jaiswal, Dickman, and Kuhnert, 2012a; Jaiswal, Jayasinghe, and Kuhnert, 2012b). Representative structures are shown in Figure 1. In the IUPAC system (−)-quinic acid is defined as 1L-1(OH),3,4,5-tetrahydroxycyclohexane carboxylic acid, but Eliel and Ramirez (1997) recommend $1\alpha,3R,4\alpha,5R$-tetrahydroxycyclohexane carboxylic acid. The CGAs show a variety of biological activities such as antioxidant, anti-inflammatory, anti-HIV, anti-HBV, and radical scavenging. They also inhibit mutagenesis and carcinogenesis and are considered to be beneficial to human health (Hemmerle *et al.*, 1997; Kwon *et al.*, 2000; Kweon, Hwang, and Sung, 2001; Wang *et al.*, 2009; Rajavelu *et al.*, 2011).

Galium odoratum L. Scop. (woodruff) is a sweet-scented plant owing to the presence of coumarin. Its leaves are widely used in folk medicine for antispasmodic, anti-inflammatory, diaphoretic, diuretic, and sedative (Launert, 1989; Lust, 1979; Triska, 1975) purposes. An infusion of leaves is used in the treatment of insomnia and nervous tension, varicose veins, biliary obstruction, hepatitis, and jaundice (Brown, 1995; Launert, 1989). In Germany, its infusion is also used to prepare some alcoholic drinks such as May wine, beer, and corn schnapps. Woodruff is very popular in Switzerland, and the cross-shaped flower was incorporated in the Swiss national flag.

Galium verum L. (Lady's bedstraw) leaves are edible and used in folk medicine, especially in India and Turkey for the treatment of kidney stones, epilepsy, skin diseases, and wound healing (Brown, 1995; Launert, 1989). They contain a high amount of organic acids, which cause the curdling of milk
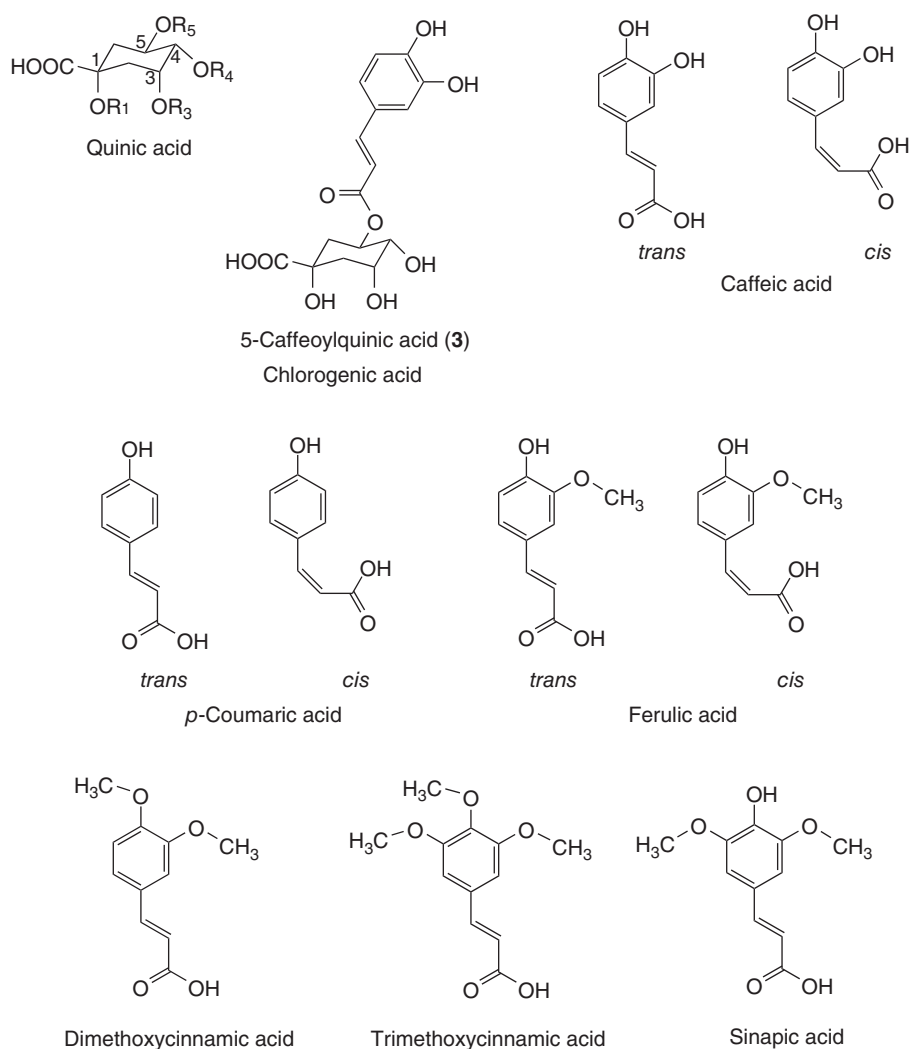
Figure 1 Representative structures of chlorogenic acids, where R_1 and R_3–R_5 are H or a cinnamoyl residue.

(Ahrendt, 1961; Chiej, 1984). The roasted seeds are used as a coffee substitute (Hedrick, 1972).

Galium boreale L. (Northern bedstraw) is a biologically active plant that was demonstrated to have biological properties such as diaphoretic and diuretic (Moerman, 2004). An extract from *G. boreale* was reported to be potentially useful as a contraceptive agent (Moerman, 2004). The raw and cooked leaves are used in human diet, and the flowering stem extract is consumed as tea (Elias and Dykeman, 1990). *Galium mollugo* L., *Galium glaucum* L., and *Galium harcynicum* L. (syn. *Galium saxatile* L.) have been used

in folk medicine for the treatment of epilepsy and hysteria (Brown, 1995).

These plants were previously investigated, and coumarins, iridoids, monoterpenoids, essential oils, and flavonoids were identified as their main secondary metabolites (Brooker, Windorski, and Bluml, 2008; Demirezer *et al*., 2006; Il'ina *et al*., 2009; Raynaud and Mnajed, 1972; Rimpler and Gmelin, 1970; Uesato *et al*., 1984; Zhao *et al*., 2008). Owing to the presence of a large number of natural products and their biological effects, these plants are of great interest to the pharmaceutical industry and the

scientific community. These six plants were selected for research for the following reasons. Firstly, it is well established that plants of the Rubiaceae family are a particularly rich source of hydroxycinnamates that considerably contribute to the dietary intake of CGAs (Clifford, 2000). The hydroxycinnamate profile of herbal remedies is much less explored and prompted in part in this investigation. Secondly, all the six plants under investigation are traditionally used in the treatment of various diseases and as food materials. Antiviral properties, including anti-HIV, antihepatitis, and anti-influenza activities, are studied. Finally, *Galium* spp. are worth investigating because of the established presence of significant quantities of coumarin secondary metabolites. It was previously shown that plant tissue exposed to UV light contains considerable amounts of *cis*-hydroxycinnamates produced by photoisomerization (Clifford *et al*., 2008). Such biochemical intermediates are required in coumarin biosynthesis (Hashidoko, Tanaka, and Tahara, 2001; Kai *et al*., 2006). However, all the plants investigated so far with a high level of *cis*-hydroxycinnamates were known not to produce coumarins, whereas *Galium* spp. in particular are able to do so. Consequently, an interesting biochemical question emerges, whether plants able to biosynthesize coumarins also contain levels of *cis*-hydroxycinnamates.

2 IDENTIFICATION AND CHARACTERIZATION OF HYDROXYCINNAMATES

All the data presented in this chapter were generated by analyzing the aqueous methanolic (70%) extracts of the plant leaves by Iontrap MS (negative ion mode), high resolution TOF-MS (negative ion mode), and RP-HPLC using the diphenyl column. Retention times for the phenolic compounds on a reversed phase diphenyl column are shown in Table 1.

2.1 Identification of Chlorogenic Acids

In 2003, we revolutionized CGA structure elucidation by introducing a tandem mass spectrometry-based method, thus allowing assignment of

regiochemistry based exclusively on fragment spectra (Clifford *et al*., 2003). We observed that all four regioisomeric mono-caffeoyl quinic acids (CQAs), and later all six regioisomeric dicaffeoyl quinic acids (diCQAs), showed dramatically different tandem mass spectra in the negative ion mode, using an ion trap mass spectrometer (Clifford, Knight, and Kuhnert, 2005). Owing to the diagnostic differences in the tandem MS fragment spectra, a consistent and predictive structure diagnostic hierarchical key for CGA structure elucidation has been established, and this allows reliable determination of CGA regiochemistry.

This tandem MS-based method has been successfully adopted by the researchers around the world and can be viewed as the current gold standard method for CGA structure elucidation.

2.1.1 Monoacyl Chlorogenic Acids

Monoacyl CGAs are more polar than the diacyl CGA, so they eluted first and appeared early in the chromatograms (Alonso-Salces, Guillou, and Berrueta, 2009; Clifford *et al*., 2003, 2006a,c, 2007, 2008; Clifford, Knight, and Kuhnert, 2005; Clifford, Wu, and Kuhnert, 2006b; Clifford, Zheng, and Kuhnert, 2006d; Jaiswal *et al*., 2010a,b; Jaiswal and Kuhnert, 2011b,c). Generally, water, water/formic acid, acetic acid, methanol, and acetonitrile have been used as solvents for HPLC with reverse phase stationary phases, such as C18, C18 amide, C8, phenylhexyl, and diphenyl.

All monoacyl CGAs gave the expected parent ion [CGA–H]$^-$. The fragment structures are presented in Figure 2 (Clifford *et al*., 2003, 2006a,c; Clifford, Wu, and Kuhnert, 2006b; Clifford, Zheng, and Kuhnert, 2006d; Jaiswal *et al*., 2010a,b).

In the cases of 3-*p*CoQA **17**, 3-feruloylquinic acid (FQA) **10**, and 3-SiQA **42** (Jaiswal *et al*., 2010a), the MS^2 and MS^3 base peak ions were derived from the cinnamic acid moiety [cinnamic acid–H]$^-$ (Clifford *et al*., 2003). For 3-FQA **10**, 3-*p*CoQA **17**, and 3-SiQA **42**, these ions were [cinnamic acid–H]$^-$ (A$_2$, A$_3$, and A$_4$) ions and the decarboxylation product [cinnamic acid –CO$_2$–H]$^-$ (B$_2$, B$_3$, and B$_4$), respectively. For the remaining monoacyl CGAs, the base peak ions in MS^2 and MS^3 spectra were derived

Table 1   Hydroxycinnamates of *G. odoratum*, *G. verum*, *G. harcynicum*, *G. mollungo*, *G. boreale*, and *G. glaucum* and their retention times on reverse phase diphenyl column.

No.	Hydroxycinnamate	Abbreviation	Retention time (min)	*G. odoratum*	*G. verum*	*G. harcynicum*	*G. mollugo*	*G. boreale*	*G. glaucum*
1	3-Caffeoylquinic acid	3-CQA	12.9	P	P	P	P	—	—
2	4-Caffeoylquinic acid	4-CQA	22.6	P	P	—	P	—	—
3	5-Caffeoylquinic acid	5-CQA	17.8	P	P	P	P	P	P
4	*cis*-3-Caffeoylquinic acid	*cis*-3-CQA	13.9	P	P	P	P	—	—
5	*cis*-4-Caffeoylquinic acid	*cis*-4-CQA	20.4	P	P	P	—	—	—
6	*cis*-5-Caffeoylquinic acid	*cis*-5-CQA	23.5	P	P	P	—	—	—
7	*epi*-Caffeoylquinic acid	*epi*-CQA	11.3	P	—	—	—	—	—
8	*epi*-Caffeoylquinic acid	*epi*-CQA	20.0	P	—	—	—	—	—
9	Methyl 5-caffeoylquinate	Me-5-CQ	28.8	P	—	—	—	—	—
10	3-Feruloylquinic acid	3-FQA	21.3	P	P	—	—	P	—
11	4-Feruloylquinic acid	4-FQA	31.6	P	P	—	—	P	P
12	5-Feruloylquinic acid	5-FQA	27.3	P	P	—	—	P	—
13	*cis*-3-Feruloylquinic acid	*cis*-3-FQA	22.8	—	P	—	—	—	—
14	*cis*-4-Feruloylquinic acid	*cis*-4-FQA	30.2	P	P	—	—	P	P
15	*cis*-5-Feruloylquinic acid	*cis*-5-FQA	31.8	P	P	—	—	P	—
16	*epi*-Feruloylquinic acid	*epi*-FQA	28.2	—	P	—	—	—	P
17	3-*p*-Coumaroylquinic acid	3-*p*CoQA	19.0	P	P	—	P	—	P
18	4-*p*-Coumaroylquinic acid	4-*p*CoQA	28.5	P	P	—	—	—	P
19	5-*p*-Coumaroylquinic acid	5-*p*CoQA	24.8	P	P	—	P	—	P
20	*cis*-4-*p*-Coumaroylquinic acid	*cis*-4-*p*CoQA	26.5	P	P	—	—	—	P
21	*cis*-5-*p*-Coumaroylquinic acid	*cis*-5-*p*CoQA	28.9	P	P	—	—	—	P
22	1,3-Dicaffeoylquinic acid	1,3-diCQA	37.4	—	P	—	—	—	—
23	1,5-Dicaffeoylquinic acid	1,5-diCQA	38.2	P	P	—	—	—	—
24	3,5-Dicaffeoylquinic acid	3,5-diCQA	39.0	P	P	P	—	—	—
25	A *cis*-3,5-Dicaffeoylquinic acid	3,5-diCQA	39.5	—	P	P	—	—	—
26	4,5-Dicaffeoylquinic acid	4,5-diCQA	41.5	P	P	P	—	—	—
27	A *cis*-4,5-Dicaffeoylquinic acid	4,5-diCQA	42.8	—	P	—	—	—	—
28	A *cis*-4,5-Dicaffeoylquinic acid	4,5-diCQA	46.5	—	P	—	—	—	—
29	*cis*-4,5-Dicaffeoylquinic acid	4,5-diCQA	48.5	—	P	—	—	—	—
30	*p*-Coumaroyl-glycoside	*p*Co-Gly	16.1	P	—	—	—	—	—
31	*p*-Coumaroyl-glycoside	*p*Co-Gly	17.0	—	—	—	—	P	—
32	*p*-Coumaroyl-glycoside	*p*Co-Gly	24.7	P	—	—	—	—	—
33	*p*-Coumaroyl-glycoside derivative	*p*Co-Gly-der	12.4	—	—	—	P	—	—
34	Feruloyl-glycoside	F-Gly	22.1	—	—	—	—	P	—
35	Feruloyl-glycoside derivative	F-Gly-der	19.0	—	—	—	—	P	—
36	Feruloyl-glycoside derivative	F-Gly-der	22.1	—	—	—	—	P	—
37	Feruloyl-glycoside derivative	F-Gly-der	19.0	—	—	—	—	P	—
38	Dihydroferuloyl-glycoside	Dihyd-F-Gly	20.2	—	—	—	—	P	—
39	Quercetin-glucoside	Q-Glu	35.0	—	—	—	—	P	—
40	Quercetin-glucoside	Q-Glu	36.0	—	—	—	—	P	—
41	Esculetin hexoside (Esculin)	E-hexoside	14.7	P	—	—	—	—	—

Abbreviations: P, present.
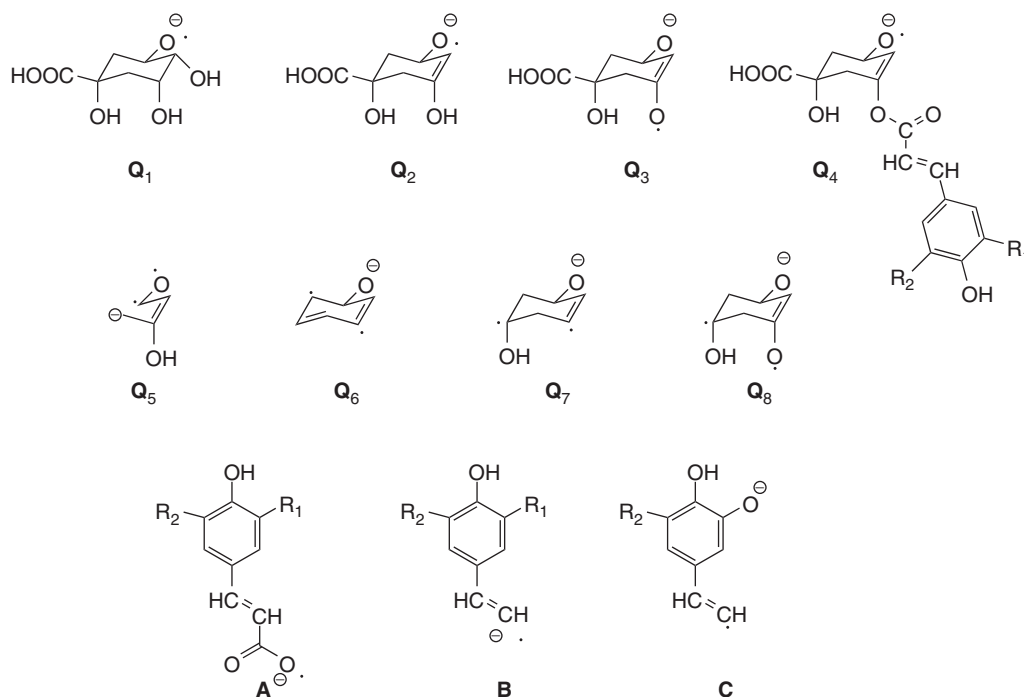
Figure 2 Structures of quinic acid-derived and cinnamic acid-derived fragments of CGAs.

from the quinic acid moiety. Two distinct fragmentation pathways occurred for the quinic acid derived fragments. One pathway gave [quinic acid–H]$^-$ (Q_1) in MS^2 and a fragment (Q_5) at m/z 86 in MS^3; the other gave [quinic acid–H$_2$O–H]$^-$ (Q_2) in MS^2 and a fragment (Q_6) at m/z 93. A 4-acyl CGA could be distinguished by its "dehydrated" MS^2 base peak (Q_2) at m/z 173 (Figure 3), supported by the MS^3 base peak at m/z 93 and Q_7 at m/z 111 (Figure 3). All other monoacyl CGAs and 3-CQA **1** produced the MS^2 base peak at m/z 191 (Q_1), supported by strong MS^3 ions at m/z 86 (Q_5), m/z 127 (Q_7), and m/z 172 (Q_3) (Figure 4). The 3-CQA **1** gave the same base peak as 5-CQA **3** but could be distinguished from 5-CQA **3** by a comparatively intense caffeic acid derived ion A_1 at m/z 179 (Figure 4). For 4-CQA **2** and other 4-acyl CQAs, two characteristic secondary peaks in MS^2 or MS^3 at m/z 255 and 299, respectively, were reported. Similar types of secondary peaks at m/z 269 and 313 were reported for 4-FQA **11** and other 4-acyl FQAs (Clifford *et al.*, 2006a,c).

2.1.2 Diacyl Chlorogenic Acids

The diacyl CGAs behaved similarly, giving the equivalent parent ion [diacyl CGA–H]$^-$ (Clifford, Knight, and Kuhnert, 2005; Clifford *et al.*, 2006a,c; Jaiswal and Kuhnert, 2010; Jaiswal *et al.*, 2010a,b). All diacyl CGAs either produced [diacyl CGA–cinnamoyl–H]$^-$ (Figure 5) or [diacyl CGA–cinnamoyl–H$_2$O–H]$^-$ (Figure 5).

The *vic* diCQA (3,4-diCQA **43** and 4,5-diCQA **26**) gave Q_2 as the MS^3 base peak at m/z 173 (Figure 5), supported by strong MS^4 ions at m/z 93 (Q_6) and Q_7 at m/z 111, which were consistent with the 4-acylated monoacyl CGAs. These ions were absent in 3,5-diCQA **24** (Figure 5), which gave Q_1 as the MS^3 base peak at m/z 191, supported by strong MS^3 ions at m/z 86 (Q_5), m/z 127 (Q_8), and m/z 172 (Q_3); this was consistent with 3-CQA **1** and 5-CQA **3**. The two *vic* diCQA isomers differed (Figure 5) with regard to their intensities of Q_4, the MS^2 "dehydrated" ion [CQA–H$_2$O–H]$^-$. In 3,4-diCQA **43**, Q_4 was more

Fragment	R_1	R_2	Cinnamic acid	Calculated Mass
Q_1				191.06
Q_2				173.04
Q_3				172.04
Q_4	OH	H	Caffeic	335.08
	OCH_3	H	Ferulic	349.08
	H	H	*p*-coumaric	319.08
	OCH_3	OCH_3	Sinapic	379.10
Q_5				85.03
Q_6				93.03
Q_7				111.04
Q_8				127.04
A_1	OH	H	Caffeic	179.04
A_2	OCH_3	H	Ferulic	193.04
A_3	H	H	*p*-coumaric	163.04
A_4	OCH_3	OCH_3	Sinapic	223.06
B_1	OH	H	Caffeic	135.04
B_2	OCH_3	H	Ferulic	149.04
B_3	H	H	*p*-coumaric	119.04
B_4	OCH_3	OCH_3	Sinapic	179.04
C_1			Ferulic	134.04
C_2			Sinapic	164.04

Figure 2   (*Continued*)

intense. In contrast, in 4,5-diCQA **26**, Q_4 was barely detectable. Similarly, the 3,4-isomer **43** produced Q_1 in MS^3 and Q_7 in MS^4 with approximately double the intensities if compared to the 4,5-isomer **26** (Clifford, Knight, and Kuhnert, 2005; Clifford *et al.*, 2006a,c; Jaiswal and Kuhnert, 2010; Jaiswal *et al.*, 2010a,b).

MS^2 base peaks for the diCQA were identical to the parent ions for the CQA (Clifford, Knight, and Kuhnert, 2005). The subsequent degradation of these ions would, therefore, be identical regardless of whether they have derived from CQA or diCQA. By comparison with the CQA MS^2 data (Figures 3–5 and Tables 2 and 3), it was possible to define the precise regiochemistry of the ions responsible for the diCQA MS^2 base peaks.

On the basis of this information, it became possible to specify which of the caffeoyl moieties was removed from the diCQA during MS^2 and MS^3. The MS^2 base peak at *m/z* 173 (Q_2) is characteristic for an isomer substituted at position 4; the MS^2 base
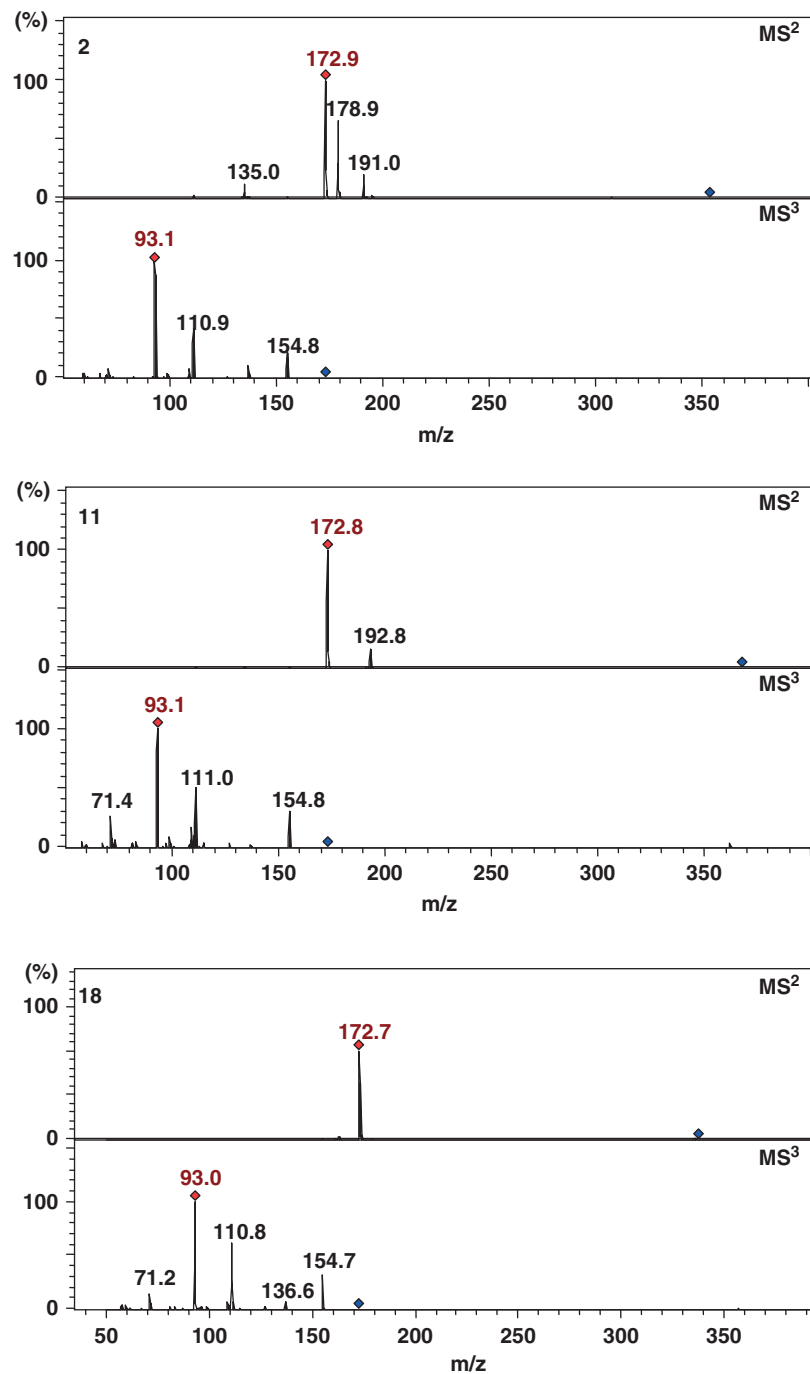
Figure 3   MS^2 and MS^3 spectra of 4-CQA **2** (*m/z* 353), 4-FQA **11** (*m/z* 367) and 4-*p*CoQA **18** (*m/z* 337) in negative ion mode.
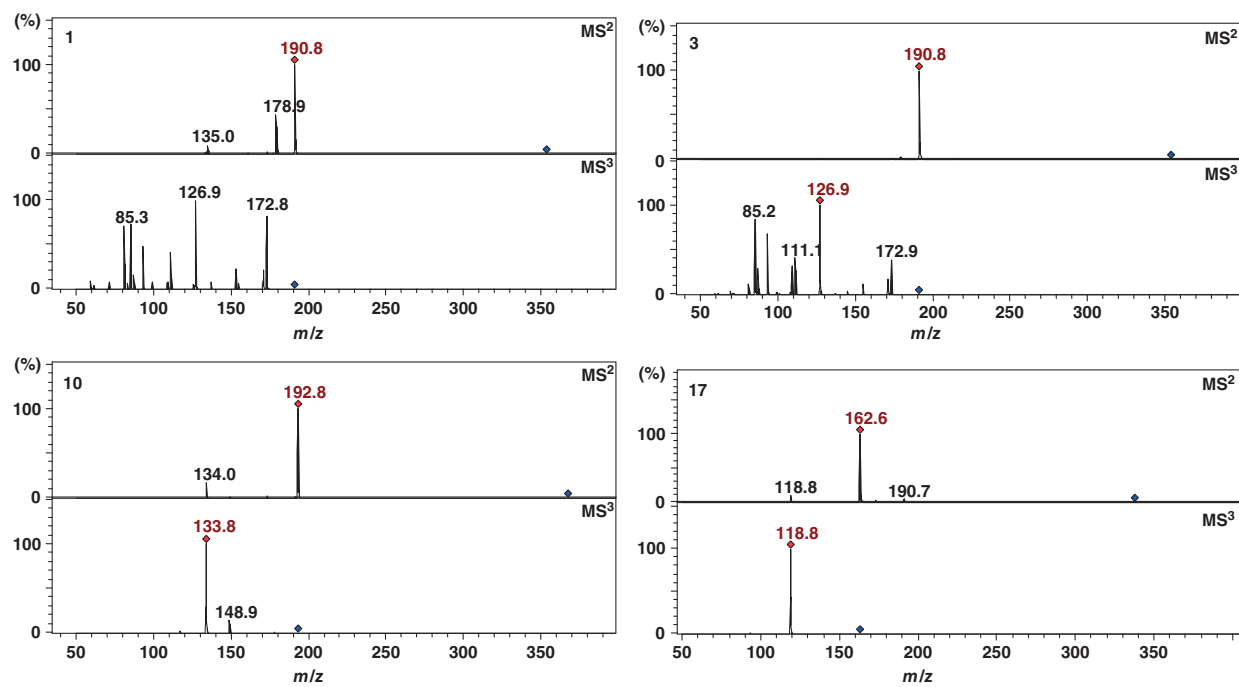
Figure 4   MS³ spectra of 3-CQA **1**, 5-CQA **3** (*m/z* 353), 3-FQA **10** (*m/z* 367) and 3-*p*CoQA **17** (*m/z* 337) in negative ion mode.
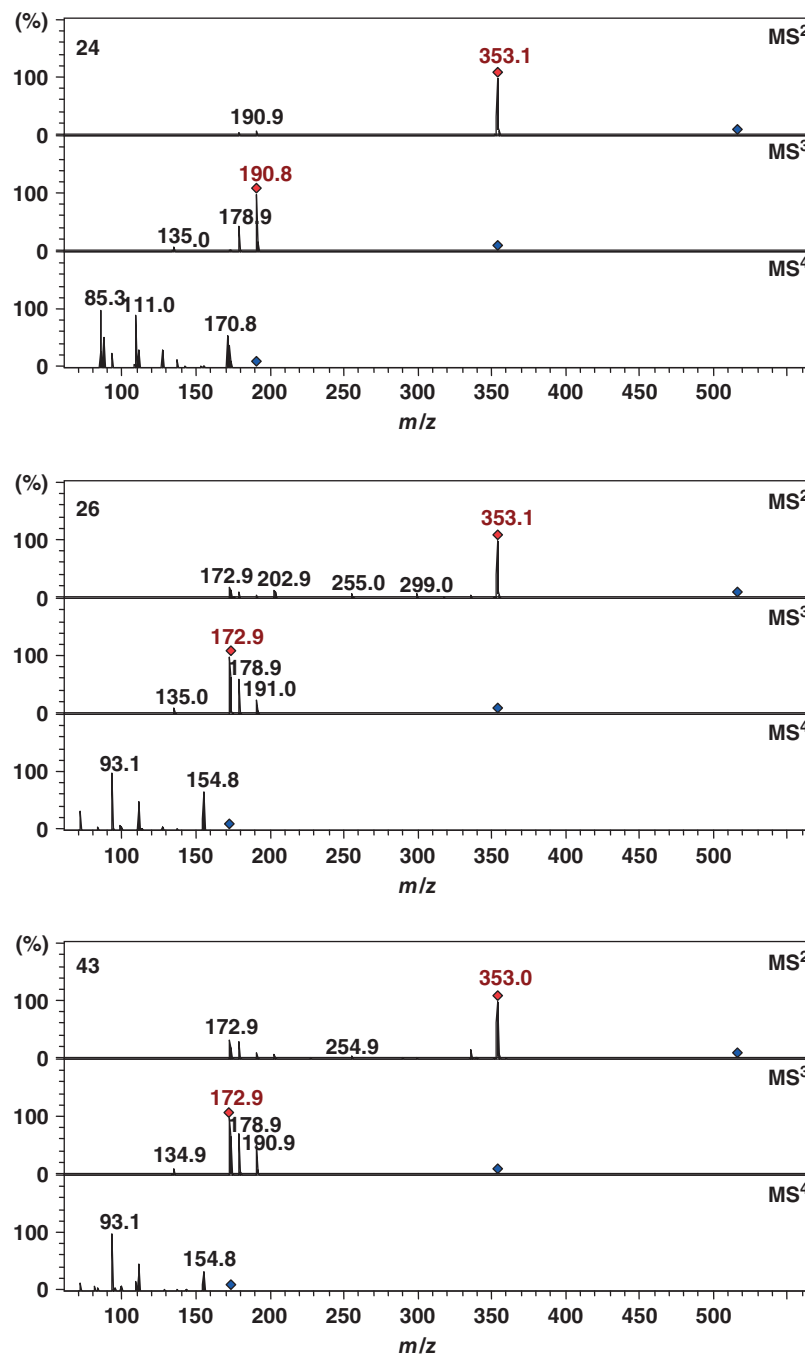
Figure 5 MS[4] spectra of 3,5-diCQA **24**, 4,5-diCQA **26**, and 3,4-diCQA **43**, in negative ion mode (*m/z* 515).

Table 2 MS^3 data of monoacyl CGAs in negative ion mode.

No.	CGA	MS¹ Parent ion	MS² Base peak	MS² Secondary peak						MS³ Base peak	MS³ Secondary peak			
		m/z	m/z	m/z	int	m/z	int	m/z	int	m/z	m/z	int	m/z	int
1	3-CQA	353.1	190.9	178.5	50	134.9	7	—	—	85.3	127.0	71	172.9	67
2	4-CQA	353.1	172.9	178.9	60	190.8	20	135.0	9	93.2	111.0	48	—	—
3	5-CQA	353.2	190.0	178.5	5	135.0	15	—	—	85.2	126.9	66	172.9	27
10	3-FQA	367.2	192.9	191.5	2	173.2	2	—	—	133.9	148.9	23	—	—
11	4-FQA	367.2	172.9	192.9	16	—	—	—	—	93.1	111.5	44	—	—
12	5-FQA	367.2	190.9	172.9	2	—	—	—	—	85.2	126.9	70	—	—
17	3-*p*CoQA	337.1	162.9	190.0	5	—	—	—	—	118.9	—	—	—	—
18	4-*p*CoQA	337.1	172.7	—	—	—	—	—	—	93.0	111.0	61	—	—
19	5-*p*CoQA	337.2	190.9	162.9	5	—	—	—	—	85.2	—	—	—	—

Abbreviations: ND, not detected; int, intensity.

Table 3 MS^4 data of diacyl CGAs in negative ion mode.

No.	CGA	MS¹ Parent ion	MS² Base peak	MS² Secondary peak						MS³ Base peak	MS³ Secondary peak						MS⁴ Base peak	MS⁴ Secondary peak			
		m/z	m/z	m/z	int	m/z	int	m/z	int	m/z	m/z	int	m/z	int	m/z	int	m/z	m/z	int	m/z	int
22	1,3-diCQA	515.1	353.1	335.1	2	173.0	4	178.9	30	190.9	179.0	60	ND		135.1	10	85.1	111.1	35	172.9	60
23	1,5-diCQA	515.2	353.1	335.1	2	173.0	4	ND		190.9	179.0	7	ND		ND		85.1	111.1	55	172.9	75
24	3,5-diCQA	515.2	353.1	335.1	2	173.0	4	ND		190.9	179.0	60	ND		135.1	6	85.1	111.1	86	172.9	60
43	3,4-diCQA	515.2	353.1	335.1	4	172.9	20	178.9	15	172.9	178.9	68	191.0	32	135.1	9	93.2	111.1	30	ND	
26	4,5-diCQA	515.2	353.1	335.1	2	172.9	6	178.9	5	172.9	178.9	76	190.9	9	135.0	19	93.1	111.0	20	ND	

Abbreviations: ND, not detected; int, intensity.

peak of *vic* diCQA must be [4-CQA–H]⁻ rather than [3-CQA–H]⁻ or [5-CQA–H]⁻. The 3,4-diCQA **43** followed the same pattern and initially lost the caffeoyl moiety at position 3, whereas 4,5-diCQA **26** initially lost that at position 5 (Clifford, Knight, and Kuhnert, 2005).

Fragmentation of the MS² base peak for 3,5-diCQA **24** yielded a comparatively intense [caffeoyl–H]⁻ ion (A_1 = 50% of base peak) (Figure 5). This was consistent with [3-CQA–H]⁻ being the MS² base peak rather than [5-CQA-H]⁻ where only a weak [caffeoyl–H]⁻ was detected. On the basis of these arguments, the acylated residue at position 4 is the most difficult to remove, whereas that at position 5 is the easiest. These facts are also true for hetero diacyl CGAs (Clifford *et al*., 2006a,c; Jaiswal *et al*., 2010a,b).

Generally, it was observed that the order of elution for the monoacyl CGAs in RP columns was 1>3>5>4 (in green coffee, there is no 1-acylated CGA) and, similarly, for the diacyl CGAs was 1,3>1,4>1,5>3,4>3,5>4,5 (Clifford *et al*., 2003; Clifford, Knight, and Kuhnert, 2005).

As a rule of thumb, the ease of fragmentation takes place in the following order: 1-acyl substituent < 5-acyl substituent < 3-acyl substituent < 4-acyl substituent, the last of these being always accompanied by loss of water (Clifford *et al*., 2003, 2006a,c; Clifford, Knight, and Kuhnert, 2005; Jaiswal and Kuhnert, 2010; Jaiswal *et al*., 2010a,b).

2.2 Chlorogenic Acids of Galium spp.

2.2.1 *Mono- and Dicaffeoylquinic Acids*

Eight CQAs (**1–8**) were detected in the extracts by virtue of their (M–H)⁻ ions at *m/z* 353 (Tables 1, 4, and 5). Three of these were assigned using the hierarchical keys previously developed (Jaiswal,

Table 4 High resolution mass (MS-TOF) data of hydroxycinnamates present in *G. odoratum* L. Scop., *G. verum* L., and *G. harcynicum* L.

No.	Hydroxycinnamate	Mol. formula	Theoretical m/z (M–H)	*G. odoratum*		*G. verum*		*G. harcynicum*	
				Exp. m/z (M-H)	Error (ppm)	Exp. m/z (M–H)	Error (ppm)	Exp. m/z (M–H)	Error (ppm)
1	3-CQA	$C_{16}H_{18}O_9$	353.0878	353.0873	1.3	353.0885	−2.0	353.0877	0.2
2	4-CQA	$C_{16}H_{18}O_9$	353.0878	353.0866	3.5	353.0884	−1.6	ND	
3	5-CQA	$C_{16}H_{18}O_9$	353.0878	353.0889	−3.0	353.0875	0.9	353.0863	4.2
4	*cis*-3-CQA	$C_{16}H_{18}O_9$	353.0878	353.0869	2.7	353.0895	−4.9	353.0884	−1.6
5	*cis*-4-CQA	$C_{16}H_{18}O_9$	353.0878	353.0865	3.7	353.0869	2.7	353.0866	3.5
6	*cis*-5-CQA	$C_{16}H_{18}O_9$	353.0878	353.0866	3.5	353.0889	−3.0	353.0878	0.0
7	*epi*-CQA	$C_{16}H_{18}O_9$	353.0878	353.0869	2.7	ND		ND	
8	*epi*-CQA	$C_{16}H_{18}O_9$	353.0878	353.0869	2.7	ND		ND	
9	Me-5-CQ	$C_{17}H_{20}O_9$	367.1035	367.1026	2.2	ND		ND	
10	3-FQA	$C_{17}H_{20}O_9$	367.1035	367.1034	0.2	367.1032	0.8	ND	
11	4-FQA	$C_{17}H_{20}O_9$	367.1035	367.1040	−1.5	367.1051	−4.4	ND	
12	5-FQA	$C_{17}H_{20}O_9$	367.1035	367.1035	−2.0	367.1034	1.0	ND	
13	*cis*-3-FQA	$C_{17}H_{20}O_9$	367.1035	ND		367.1020	4.1	ND	
14	*cis*-4-FQA	$C_{17}H_{20}O_9$	367.1035	367.1036	−0.4	367.1041	−1.7	ND	
15	*cis*-5-FQA	$C_{17}H_{20}O_9$	367.1035	367.1030	1.2	367.1040	−1.5	ND	
16	*epi*-FQA	$C_{17}H_{20}O_9$	367.1035	ND		367.1038	−1.0	ND	
17	3-*p*CoQA	$C_{16}H_{18}O_8$	337.0929	337.0934	−1.4	337.0929	−0.1	ND	
18	4-*p*CoQA	$C_{16}H_{18}O_8$	337.0929	337.0913	4.7	337.0912	5.0	ND	
19	5-*p*CoQA	$C_{16}H_{18}O_8$	337.0929	337.0919	2.9	337.0921	2.0	ND	
20	*cis*-4-*p*CoQA	$C_{16}H_{18}O_8$	337.0929	337.0922	2.2	337.0923	1.7	ND	
21	*cis*-5-*p*CoQA	$C_{16}H_{18}O_8$	337.0929	337.0921	2.0	337.0919	3.0	ND	
22	1,3-diCQA	$C_{25}H_{24}O_{12}$	515.1195	ND		515.1198	−0.6	ND	
23	1,5-diCQA	$C_{25}H_{24}O_{12}$	515.1195	515.1197	2.2	515.1213	−3.5	ND	
24	3,5-diCQA	$C_{25}H_{24}O_{12}$	515.1195	515.1195	−0.1	515.1203	−1.5	515.1208	−2.6
25	A *cis*-3,5-diCQA	$C_{25}H_{24}O_{12}$	515.1195	ND		515.1192	0.6	515.1195	−0.1
26	4,5-diCQA	$C_{25}H_{24}O_{12}$	515.1195	515.1180	3.0	515.1176	3.8	515.1215	−3.9
27	A *cis*-4,5-diCQA	$C_{25}H_{24}O_{12}$	515.1195	ND		515.1185	1.9	ND	
28	A *cis*-4,5-diCQA	$C_{25}H_{24}O_{12}$	515.1195	ND		515.1209	−2.8	ND	
29	*cis*-4,5-diCQA	$C_{25}H_{24}O_{12}$	515.1195	ND		515.1215	−4.4	ND	
30	*p*Co-Gly	$C_{15}H_{17}O_8$	325.0929	325.0938	−2.7	ND		ND	
32	*p*Co-Gly	$C_{15}H_{17}O_8$	325.0929	325.0935	−1.9	ND		ND	
41	E-hexoside	$C_{15}H_{17}O_9$	339.0722	339.0739	−5.0	ND		ND	

Abbreviations: ND, not detected.

Deshpande, and Kuhnert, 2011a; Jaiswal, Kiprotich, and Kuhnert, 2011b; Jaiswal and Kuhnert, 2010,b; Jaiswal *et al*., 2010a,b; Clifford *et al*., 2003, 2006a,c; Clifford, Knight, and Kuhnert, 2005) as the well-known 3-CQA (**1**), 4-CQA (**2**), and 5-CQA (**3**) (Tables 1, 4, and 5). Three more peaks present as minor components displayed identical fragmentation patterns to 3-, 4-, and 5-CQAs, respectively, and it was suspected that they might be the *cis* isomers (**4–6**) of the corresponding CQAs (Jaiswal *et al*., 2010a,b; Jaiswal, Deshpande, and Kuhnert, 2011a; Jaiswal, Kiprotich, and Kuhnert, 2011b). For confirmation, extracts of all the plants were irradiated with UV light at 245 nm for 40 min. As expected, the intensity of each of the three putative *cis* isomers in the chromatogram significantly increased compared

to their corresponding *trans* isomers from the original plant extracts.

The presence of *cis* isomers of CGAs was previously reported in coffee leaves, Asteraceae plants, and maté tea (Clifford *et al*., 2008; Karakoese, Jaiswal, and Kuhnert, 2011; Jaiswal, Kiprotich, and Kuhnert, 2011b; Jaiswal *et al*., 2010b; Clifford, Zheng, and Kuhnert, 2006d). It appears that CGAs present in the plant tissue exposed to natural UV light undergo *trans–cis* isomerization, whereas in the tissue unexposed to the UV light, such as the seeds of coffee berries, they remain unchanged. It was proposed that the *cis*-hydroxycinnamate isomers are biosynthetic precursors to coumarins (Hashidoko, Tanaka, and Tahara, 2001; Kai *et al*., 2006). To investigate this, a search was conducted for an MS[2]
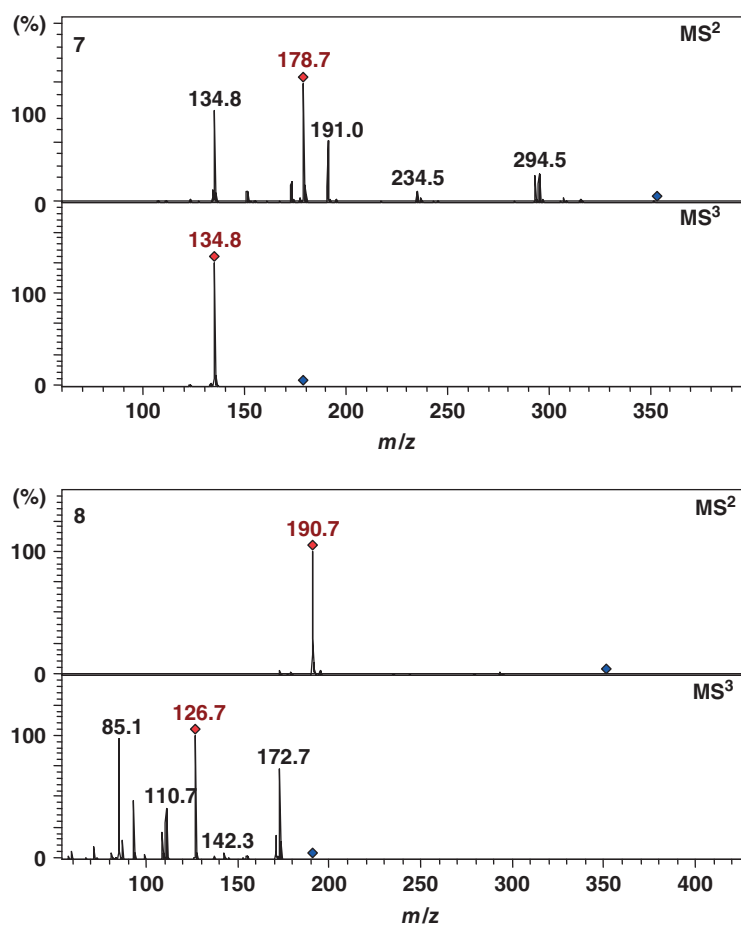
Figure 6 MS^3 spectra of *epi*-caffeoylquinic acids **7** and **8** (parent ion at *m/z* 353) in negative ion mode.
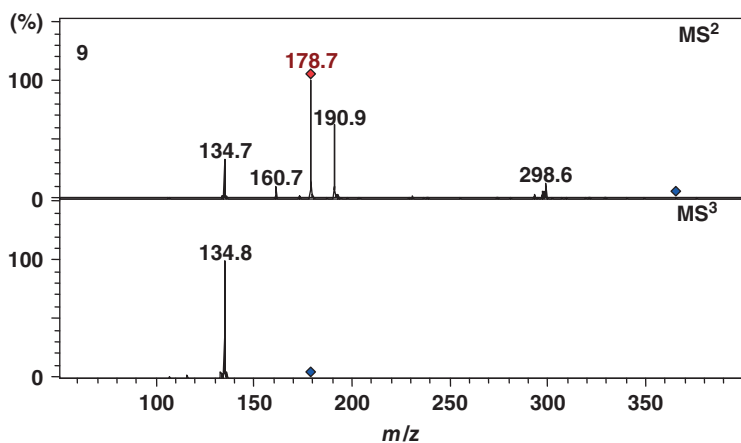
Figure 7 MS^3 spectra of methyl ester of caffeoylquinic acid **9** (parent ion at *m/z* 367) in negative ion mode.

Table 5 High resolution mass (MS-TOF) data of hydroxycinnamates present in *G. mollungo* L., *G. boreale* L., and *G. glaucum* L.

No.	Hydroxycinnamate	Mol. Formula	Theoretical m/z (M–H)	G. mollungo		G. boreale		G. glaucum	
				Exp. m/z (M–H)	Error (ppm)	Exp. m/z (M–H)	Error (ppm)	Exp. m/z (M–H)	Error (ppm)
1	3-CQA	$C_{16}H_{18}O_9$	353.0878	353.0892	−3.9	ND		ND	
2	4-CQA	$C_{16}H_{18}O_9$	353.0878	353.0893	−4.2	ND		ND	
3	5-CQA	$C_{16}H_{18}O_9$	353.0878	353.0878	0.0	353.0867	3.2	353.0887	−2.6
4	*cis*-3-CQA	$C_{16}H_{18}O_9$	353.0878	353.0895	−4.8	ND		ND	
10	3-FQA	$C_{17}H_{20}O_9$	367.1035	367.1034	0.2	367.1038	−1.0	ND	
11	4-FQA	$C_{17}H_{20}O_9$	367.1035	ND		367.1032	0.8	367.1033	0.5
12	5-FQA	$C_{17}H_{20}O_9$	367.1035	ND		367.1032	0.8	ND	
13	*cis*-3-FQA	$C_{17}H_{20}O_9$	367.1035	ND		ND		ND	
14	*cis*-4-FQA	$C_{17}H_{20}O_9$	367.1035	ND		367.1043	−2.3	367.1033	−4.4
15	*cis*-5-FQA	$C_{17}H_{20}O_9$	367.1035	ND		367.1032	0.8	ND	
17	3-*p*CoQA	$C_{16}H_{18}O_8$	337.0929	337.0939	−2.9	ND		337.0932	−0.9
18	4-*p*CoQA	$C_{16}H_{18}O_8$	337.0929	ND		ND		337.0928	0.8
19	5-*p*CoQA	$C_{16}H_{18}O_8$	337.0929	337.0936	−2.1	ND		337.0944	−4.4
20	*cis*-4-*p*CoQA	$C_{16}H_{18}O_8$	337.0929	ND		ND		337.0933	−1.2
21	*cis*-5-*p*CoQA	$C_{16}H_{18}O_8$	337.0929	ND		ND		337.0942	−4.0
31	*p*Co-Gly	$C_{15}H_{17}O_8$	325.0929	ND		367.1032	2.6	ND	
33	*p*Co-Gly-der[a]	$C_{15}H_{21}O_{10}$	—	—	—	ND		ND	
34	F-Gly	$C_{16}H_{19}O_9$	355.1035	ND		—	—	ND	
35	F-Gly-der[a]	$C_{16}H_{23}O_{11}$	391.1246	ND		—	—	ND	
36	F-Gly-der[a]	$C_{16}H_{23}O_{11}$	391.1246	ND		—	—	ND	
37	F-Gly-der[a]	$C_{18}H_{25}O_{10}$	401.1453	ND		—	—	ND	
38	Dihyd-F-Gly	$C_{16}H_{21}O_9$	357.1191	ND		357.1178	3.8	ND	
39	Q-gly	$C_{21}H_{19}O_{12}$	463.0882	ND		463.0881	0.2	ND	
40	Q-gly	$C_{21}H_{19}O_{12}$	463.0882	ND		463.0883	−0.2	ND	

[a] Error was more than 5 ppm.

Abbreviations: ND, not detected.
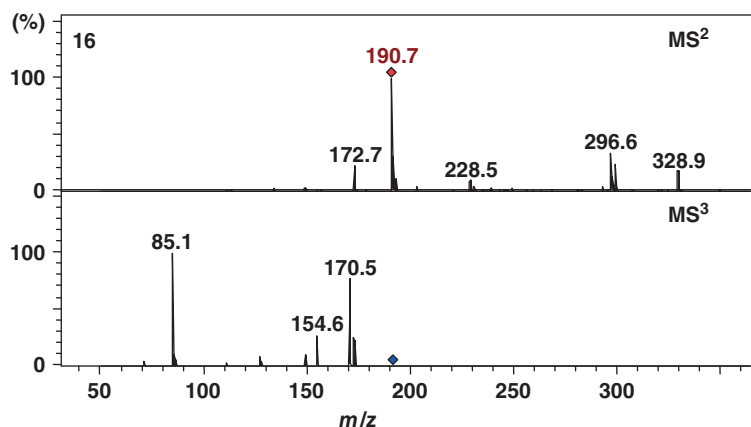
Figure 8 MS[3] spectra of *epi*-feruloylquinic acid **16** (parent ion at *m/z* 367) in negative ion mode.

fragment ion at *m/z* 177.1 and an MS[3] fragment ion at *m/z* 132.8 characteristic of esculetin derivatives (3,4-dihydroxy coumarins) (Sanchez-Rabaneda *et al.*, 2003), the expected products of caffeic acid derivatives, presumably formed in a 6-exo-trig cyclization of a dehydrogenated caffeic acid derivative (Eames, Kuhnert, and Warren, 1999; Kuhnert *et al.*, 2010). Any suspected presence identified can also be supported by the characteristic UV–vis absorption data.
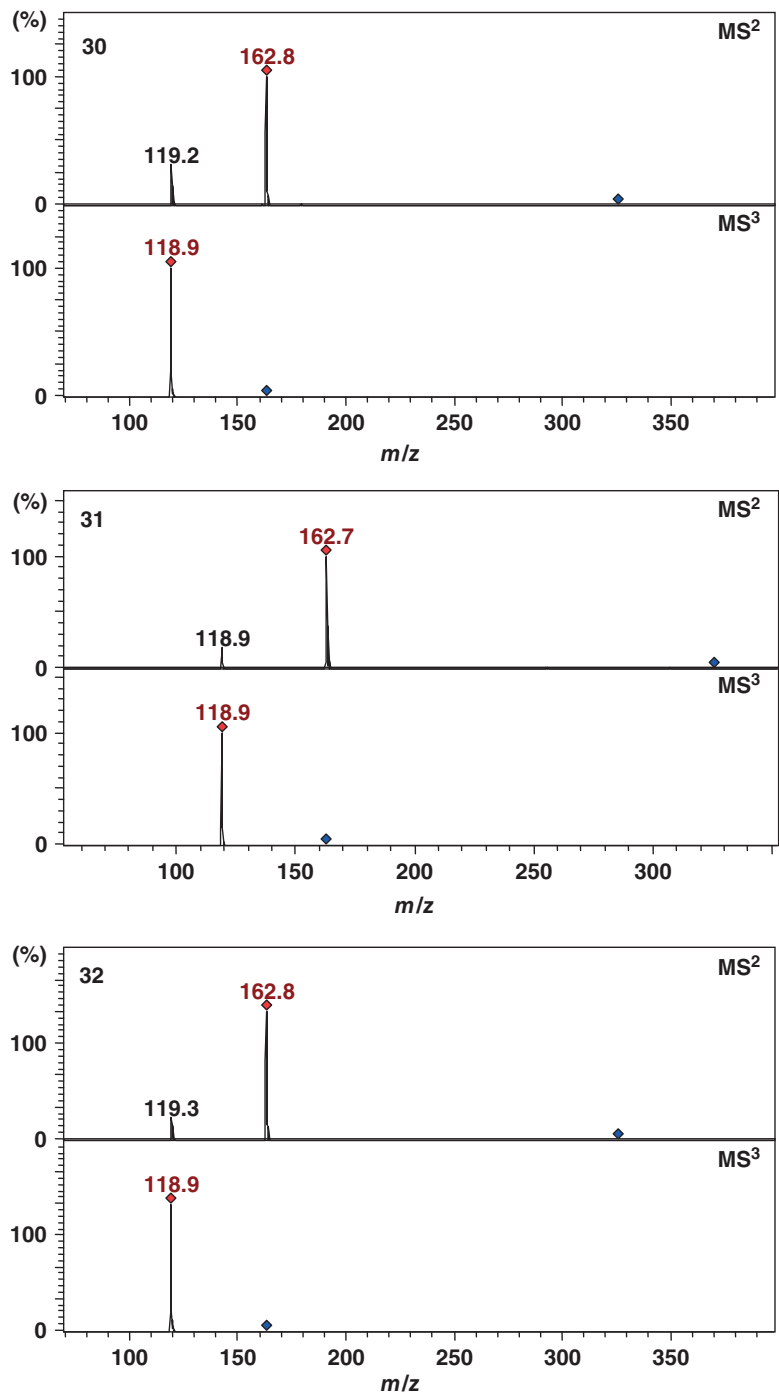
Figure 9 MS³ spectra of *p*-coumaroyl-glucosides **30–32** (parent ion at *m/z* 325) in negative ion mode.
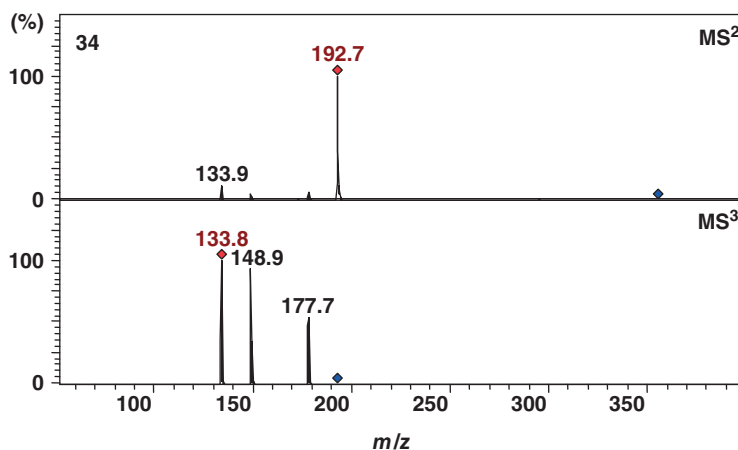
Figure 10 MS[3] spectra of feruloyl-glucoside **34** (parent ion at m/z 355) in negative ion mode.

In this way, one possible esculetin derivative (**41**) was detected in one of the extracts under investigation (Tables 1, 4, and 5).

The remaining two CQAs (**7** and **8**) were tentatively assigned as diastereomers of *epi*-CQAs because they had different retention times and significantly different fragmentation patterns when compared to 1-, 3-, 4-, or 5-CQA (Figure 6) (Clifford *et al.*, 2006a,c; Clifford, Knight, and Kuhnert, 2005; Jaiswal and Kuhnert, 2010, 2011b,c; Clifford *et al.*, 2010; Kuhnert *et al.*, 2010; Jaiswal *et al.*, 2010b).

Similarly, eight diCQA isomers (**22–29**) were tentatively identified by their (M–H)[−] ions at m/z 515 (Tables 1, 4, and 5). Four of these were assigned as 1,3-diCQA (**22**), 1,5-diCQA (**23**), 3,5-diCQA (**24**), and 4,5-diCQA (**26**). The remaining four m/z 515 isomers were suggested to have at least one *cis* caffeic acid group based on their fragmentation pattern, a characteristic shoulder in their UV spectra and the increased intensity after the UV irradiation (Clifford *et al.*, 2008; Karakoese, Jaiswal, and Kuhnert, 2011; Jaiswal, Deshpande, and Kuhnert, 2011a; Jaiswal, Kiprotich, and Kuhnert, 2011b; Jaiswal *et al.*, 2010b). It was not possible to differentiate which of the two caffeoyl groups was *cis* or *trans*. Thus, **25** was suggested to be 3,5-diCQA, with one *cis* caffeoyl moiety, while the last three (**27–29**) seemed to be isomers of a *cis*-4,5-diCQA.

A search for methyl, ethyl, or butyl esters of CQAs, such as previously reported, resulted in the tentative identification of one chromatographic peak as the methyl ester of CQA (**9**) with a parent (M–H)[−] ion at m/z 367 (Figure 7) (Jaiswal and Kuhnert,

2011a). Evidence for this was the fragmentation of m/z 367 to produce the MS[2] base peak at m/z 179 ([caffeic acid–H[+]][−]) and the secondary peaks at m/z 191 ([quinic acid–H[+]][−]) and m/z 135 ([caffeic acid–CO_2–H[+]][−]) confirmed a methyl ester of quinic acid instead of an FQA. The assignment was confirmed by methylating 3-CQA, 4-CQA, and 5-CQA. The retention time and MS fragmentation pattern of the natural product were identical to methyl-5-caffeoylquinate.

2.2.2 Feruloylquinic Acids and p-Coumaroylquinic Acids

Six minor components were detected with (M–H)[−] at m/z 337. Three of these were identified by their MS[n] fragmentation (Jaiswal and Kuhnert, 2010; Clifford *et al.*, 2006a,c) as the well-characterized 3-*p*-coumaroylquinic acid (CoQA) (**17**), 4-*p*CoQA (**18**), and 5-*p*CoQA (**19**). Another two isomers were tentatively identified as the *cis*-4-*p*CoQA (**20**) and *cis*-5-*p*CoQA, whose chromatographic signals again increased after UV irradiation (**21**) (Table 2) (Karakoese, Jaiswal, and Kuhnert, 2011; Jaiswal, Kiprotich, and Kuhnert, 2011b; Jaiswal *et al.*, 2010b).

An analogous experiment monitoring m/z 367 detected eight compounds including 3-FQA (**10**), 4-FQA (**11**), 5-FQA (**12**), *cis*-3-FQA (**13**), *cis*-4-FQA (**14**), and *cis*-5-FQA (**15**) (Clifford *et al.*, 2003, 2006a,c, 2008; Karakoese, Jaiswal, and
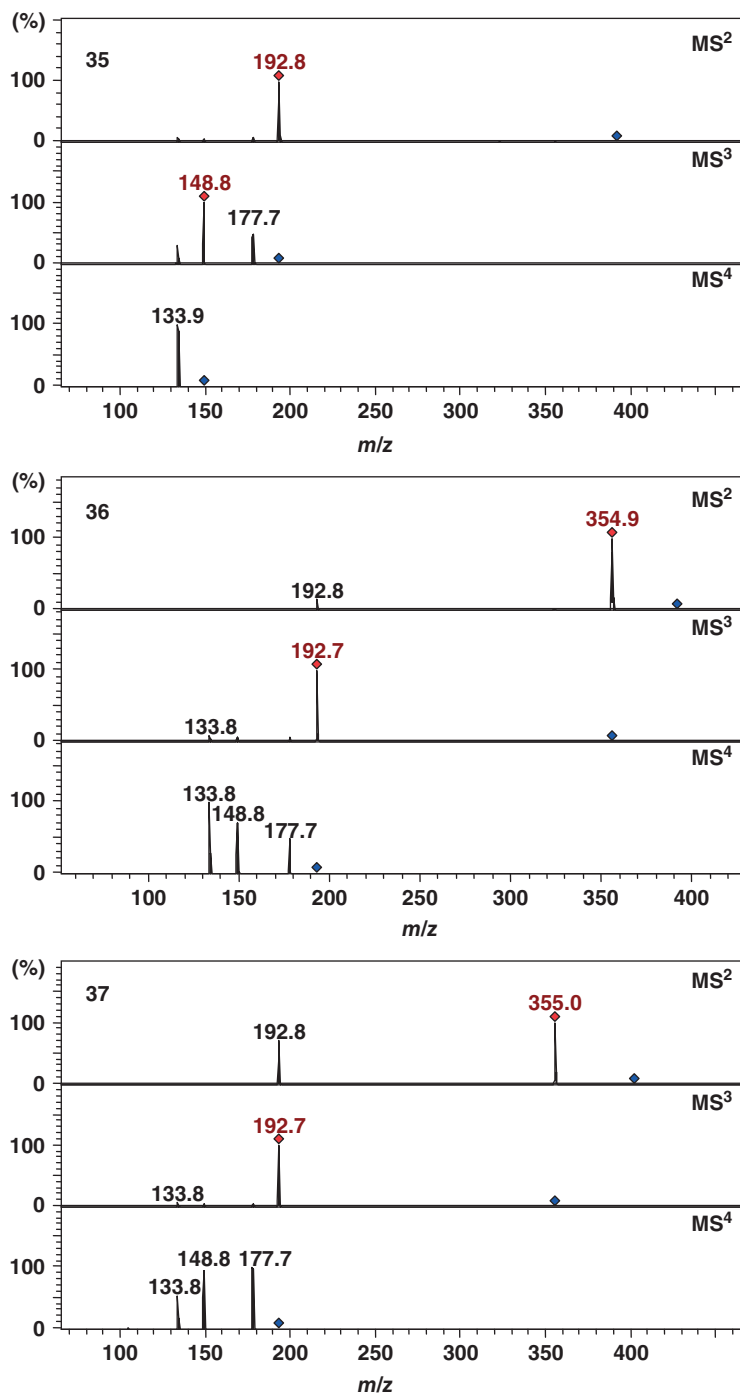
Figure 11   MS³ spectra of feruloyl-glucoside derivatives **35–37** (parent ion at *m/z* 391 and 401) in negative ion mode.
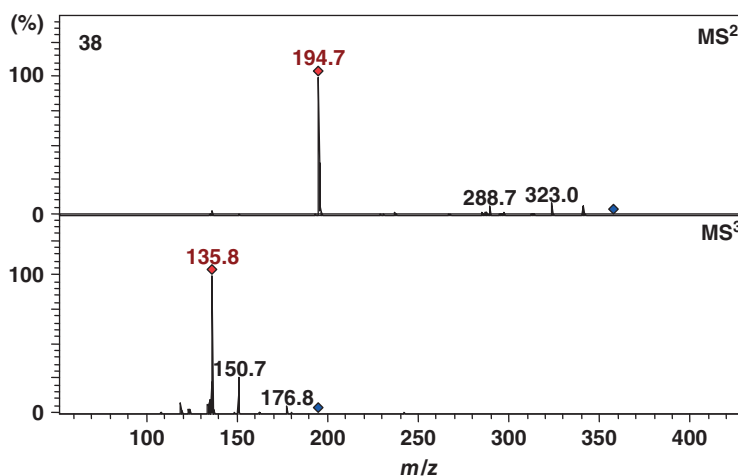
Figure 12 MS^3 spectra of dihydroferuloyl-glucoside derivatives **38** (parent ion at *m/z* 357) in negative ion mode.

Kuhnert, 2011; Jaiswal and Kuhnert, 2011a,b; Jaiswal *et al.*, 2010a,b). The latter *cis* compounds were again enhanced on UV irradiation of the plant extract. The remaining FQA (**16**) was tentatively assigned as an *epi*-FQA comprising an epimer of (−)-quinic acid, because it had different retention time and significantly different fragmentation patterns when compared to 1-, 3-, 4-, or 5-FQA and 1-, 3-, 4-, or 5-iso-FQA (Table 1 and Figure 8) (Kuhnert *et al.*, 2010).

2.3 Identification of Other Hydroxycinnamates

Several minor phenolic compounds were tentatively identified from extracted ion chromatograms displaying selected *m/z* values.

Three compounds were detected with an (M–H)$^-$ at *m/z* 325 and were tentatively assigned as *p*-coumaroyl-glycosides (**30–32**). Each of these compounds had similar MS^n fragmentation patterns with similar intensities of ions. They produced the MS^2 base peak at *m/z* 163 ([*p*-coumaric acid–H$^+$]$^-$) and a secondary peak at *m/z* 119 (≈25% of the base peak) (Figure 9), which were indicative of *p*-coumaric acid derivatives (Jaiswal *et al.*, 2010b; Clifford *et al.*, 2003, 2006c).

One compound (**33**) was detected with an (M–H)$^-$ ion at *m/z* 371 and showed a loss of 46 Da (yielding an *m/z* 325 ion) in the MS^2 experiment. A subsequent MS^3 experiment for *m/z* 371+325 gave MS^3 spectra identical to the *p*-coumaroyl-glycosides (**30–32**), and therefore **33** was tentatively assigned as a *p*-coumaroyl-glycoside derivative (Figure 9).

Another compound was detected with an (M–H)$^-$ parent ion at *m/z* 355 and was tentatively assigned as feruloyl-glycoside (**34**). This feruloyl-glycoside produced the MS^2 base peak at *m/z* 193 ([ferulic acid–H$^+$]$^-$) and a secondary peak at *m/z* 134 (10% of the base peak) (Figure 10). Further fragmentation of the *m/z* 193 ion in an MS^3 experiment resulted in a base peak at *m/z* 134 and secondary peaks at *m/z* 149 ([ferulic acid–CO$_2$–H$^+$]$^-$) and 178 ([ferulic acid–CH$_3$–H$^+$]$^-$) (Figure 10). Ions at *m/z* 134, 149, 178, and 193 have been reported for ferulic acid derivatives.

Two other postulated feruloyl-glycoside derivatives (**35** and **36**) produced an (M–H)$^-$ parent ion at *m/z* 391, while a third postulated feruloyl-glycoside derivative (**37**) produced an (M–H)$^-$ parent ion at *m/z* 401 (Figure 11). In addition to showing fragments indicative of a ferulic acid entity, compound **37** was analogous to the *p*-coumaroyl-glycoside derivative **33** in that a fragmentation loss of 46 Da from the (M–H)$^-$ ion was observed while compound **36** showed a loss of 36 Da from the (M–H)$^-$ ion. The MS^3 and MS^4 spectra of isomer **36** were identical to MS^2 and MS^3 spectra of compound **34** (Figures 10 and 11).

One compound was detected with an (M–H)$^-$ at *m/z* 357 and was tentatively assigned as dihydroferuloyl-glycoside (**38**) based on the
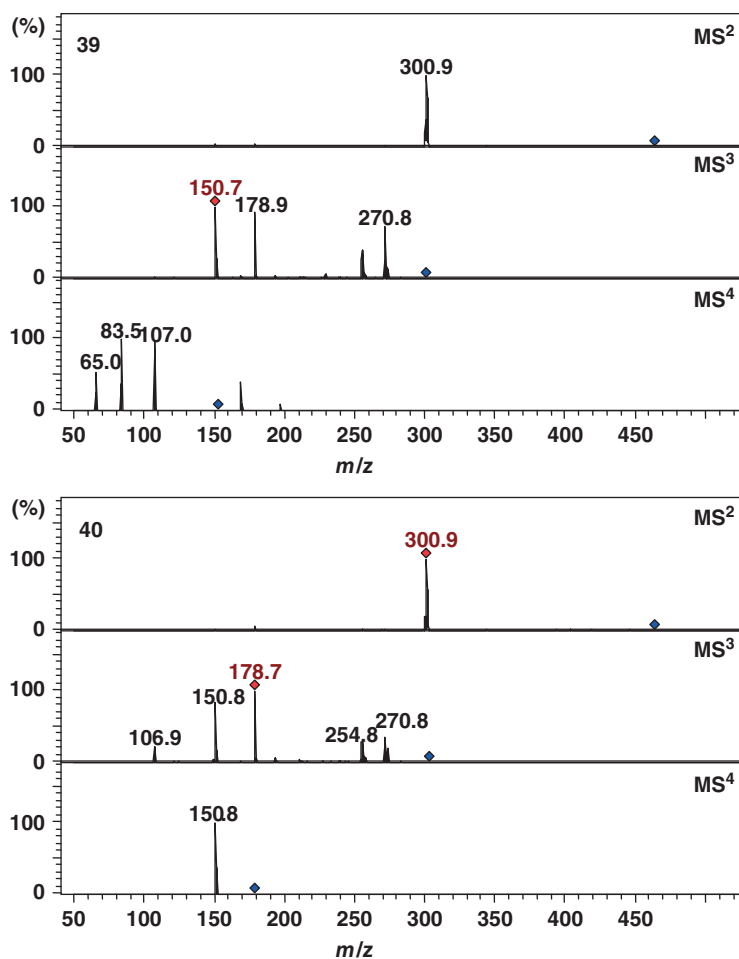
Figure 13 MS[3] spectra of quercetin 3-*O*-galactoside **39** and quercetin 3-*O*-glucoside **40** (parent ion at *m/z* 463) in negative ion mode.

observation of MS fragments, which were two mass units higher than the ones usually observed for a ferulic acid compound, that is, *m/z* 195, *m/z* 151, and *m/z* 136 (Figure 12).

Two compounds detected with an (M–H)⁻ at *m/z* 463 were assigned as quercetin 3-*O*-galactoside and quercetin 3-*O*-glucoside (**39** and **40**, respectively) based on the observation of an MS[2] base peak at *m/z* 301 ([quercetin-H⁺]⁻) and authentic standards (Figure 13).

One compound was detected with an (M–H)⁻ at *m/z* 339 and was assigned as Esculin (**41**) based on the authentic standard and observation of MS[2] fragments observed at *m/z* 163 and *m/z* 119 by the neutral loss of 162 Da (hexoside) and 206 Da (hexoside+CO_2), respectively.

3 CONCLUSIONS

In this study, 29 CGAs, eight cinnamoyl-glycosides, two quercetin-glucosides, and three esculetin-glycosides from six different *Galium* spp. (Rubiaceae) were identified. Twenty-seven of these were not previously reported in the Rubiaceae family, with two of these compounds observed for the first time in nature. Each plant species showed a unique phytochemical profile with only mono-CQAs present in all species investigated. Triacyl CGAs and CGAs derived from *muco*-quinic acid, isoferulic acid (Jaiswal, Kiprotich, and Kuhnert, 2011b; Jaiswal and Kuhnert, 2011c; Jaiswal, Dickman, and Kuhnert, 2012a; Kuhnert *et al.*, 2010; Jaiswal and Kuhnert, 2010), or derivatives carrying unusual side chains, as observed in

green coffee and Gardeniae Fructus (Clifford *et al.*, 2010), the commercially most important member of the Rubiaceae family, were not observed in this study.

REFERENCES

Ahrendt, L. W. A. (1961) *Bot. J. Linn. Soc.*, **57**, 1–410.

Alonso-Salces, R., Guillou, C., and Berrueta, L. A. (2009) *Rapid Commun. Mass Spectrom.*, **23**, 363–383.

Brooker, N., Windorski, J., and Bluml, E. (2008) *Commun. Agric. Appl. Biol. Sci.*, **73**, 81–89.

Brown, D. (1995) *Encyclopedia of Herbs & Their Uses*, Dorling Kindersley Books, London.

Chiej, R. (1984) *The Macdonald Encyclopedia of Medicinal Plants*, Macdonald Co. Ltd., London, p. 447.

Clifford, M. N. (1999) *J. Sci. Food Agric.*, **79**, 362–372.

Clifford, M. N. (2000) *J. Sci. Food Agric.*, **80**, 1033–1043.

Clifford, M. N., Johnston, K. L., Knight, S., *et al.* (2003) *J. Agric. Food Chem.*, **51**, 2900–2911.

Clifford, M. N., Knight, S., and Kuhnert, N. (2005) *J. Agric. Food Chem.*, **53**, 3821–3832.

Clifford, M. N., Knight, S., Surucu, B., *et al.* (2006a) *J. Agric. Food Chem.*, **54**, 1957–1969.

Clifford, M. N., Wu, W. G., and Kuhnert, N. (2006b) *Food Chem.*, **95**, 574–578.

Clifford, M. N., Marks, S., Knight, S., *et al.* (2006c) *J. Agric. Food Chem.*, **54**, 4095–4101.

Clifford, M. N., Zheng, W., and Kuhnert, N. (2006d) *Phytochem. Anal.*, **17**, 384–393.

Clifford, M. N., Wu, W., Kirkpatrick, J., *et al.* (2007) *J. Agric. Food Chem.*, **55**, 929–936.

Clifford, M. N., Kirkpatrick, J., Kuhnert, N., *et al.* (2008) *Food Chem.*, **106**, 379–385.

Clifford, M. N., Wu, W., Kirkpatrick, J., *et al.* (2010) *Rapid Commun. Mass Spectrom.*, **24**, 3109–3120.

Demirezer, L. O., Gurbuz, F., Guvenalp, Z., *et al.* (2006) *Turk. J. Chem.*, **30**, 525–534.

Eames, J., Kuhnert, N., and Warren, S. (1999) *Synlett*, **8**, 1211–1214.

Elias, T., and Dykeman, P. (1990) *Edible Wild Plants*, Sterling Publishing Co., New York.

Eliel, E. L., and Ramirez, M. B. (1997) *Tetrahedron Asymm.*, **8**, 3551–3554.

Hashidoko, Y., Tanaka, T., and Tahara, S. (2001) *Biosci. Biotechnol. Biochem.*, **65**, 2604–2612.

Hedrick, U. P. (1972) *Stutevant's Edible Plants of the World*, Dover Publications, New York, p. 322.

Hemmerle, H., Burger, H. J., Below, P., *et al.* (1997) *J. Med. Chem.*, **40**, 137–145.

Il'ina, T. V., Kovaleva, A. M., Goryachaya, O. V., *et al.* (2009) *Chem. Nat. Compd.*, **45**, 587–588.

IUPAC (1976) *Biochem. J.*, **153**, 23–31.

Jaiswal, R., and Kuhnert, N. (2010) *Rapid Commun. Mass Spectrom.*, **24**, 2283–2294.

Jaiswal, R., and Kuhnert, N. (2011a) *J. Mass Spectrom.*, **46**, 269–281.

Jaiswal, R., and Kuhnert, N. (2011b) *Food Funct.*, **2**, 63–71.

Jaiswal, R., and Kuhnert, N. (2011c) *J. Agric. Food Chem.*, **59**, 4033–4039.

Jaiswal, R., Patras, M. A., Eravuchira, P. J., *et al.* (2010a) *J. Agric. Food Chem.*, **58**, 8722–8737.

Jaiswal, R., Sovdat, T., Vivan, F., *et al.* (2010b) *J. Agric. Food Chem.*, **58**, 5471–5484.

Jaiswal, R., Deshpande, S., and Kuhnert, N. (2011a) *Phytochem. Anal.*, **22**, 432–441.

Jaiswal, R., Kiprotich, J., and Kuhnert, N. (2011b) *Phytochemistry*, **72**, 781–790.

Jaiswal, R., Dickman, M. H., and Kuhnert, N. (2012a) *Org. Biomol. Chem.*, **10**, 5266–5277.

Jaiswal, R., Jayasinghe, L., and Kuhnert, N. (2012b) *J. Mass Spectrom.*, **47**, 502–515.

Kai, K., Shimizu, B., Mizutani, M., *et al.* (2006) *Phytochemistry*, **67**, 379–386.

Karakoese, H., Jaiswal, R., and Kuhnert, N. (2011) *J. Agric. Food Chem.*, **59**, 10143–10150.

Kuhnert, N., Jaiswal, R., Matei, M. F., *et al.* (2010) *Rapid Commun. Mass Spectrom.*, **24**, 1575–1582.

Kuhnert, N., Karakoese, H., and Jaiswal, R. (2012) *Analysis of Chlorogenic Acids and Other Hydroxycinnamates in Food, Plants and Pharmacokinetic Studies. Handbook of Analysis of Active Compounds in Functional Foods*, CRC Press, Boca Raton, pp. 461–510.

Kweon, M. H., Hwang, H. J., and Sung, H. C. (2001) *J. Agric. Food Chem.*, **49**, 4646–4655.

Kwon, H. C., Jung, C. M., Shin, C. G., *et al.* (2000) *Chem. Pharm. Bull.*, **48**, 1796–1798.

Launert, E. (1989) *The Hamlyn Guide to Edible & Medicinal Plants of Britain and Northern Europe*, The Hamlyn Publishing Group Limited, London.

Lust, J. (1979) *The Herb Book*, Bantam Books, New York.

Moerman, D. E. (2004) *Native American Ethnobotany*, 5th edn, Timber Press, Cambridge, pp. 453–459.

Rajavelu, A., Tulyasheva, Z., Jaiswal, R., *et al.* (2011) *BMC Biochem.*, **12**, 1–8.

Raynaud, J., and Mnajed, H. (1972) *Cr. Acad. Sci. D Nat.*, **274**, 1746–1748.

Rimpler, H., and Gmelin, R. (1970) *Phytochemistry*, **9**, 1891–1892.

Sanchez-Rabaneda, F., Jauregui, O., Lamuela-Raventos, R. M., *et al.* (2003) *J. Chrom. A*, **1008**, 57–72.

Triska, J. (1975) *The Hamlyn Encyclopedia of Plants*, The Hamlyn Publishing Group Limited, London.

Uesato, S., Ueda, M., Inouye, H., *et al.* (1984) *Phytochemistry*, **23**, 2535–2537.

Wang, G., Shi, L., Ren, Y., *et al.* (2009) *Antiviral Res.*, **83**, 186–190.

Zhao, C., Shao, J., Li, X., *et al.* (2008) *J. Asian Nat. Prod. Res.*, **10**, 611–615.

Jaiswal, R., Matei, M. F., Golon, A., *et al.* (2012c) *Food Funct.*, **3**, 976–984.

FURTHER READINGS

Jaiswal, R., Matei, M. F., Ullrich, F., *et al.* (2011c) *J. Mass Spectrom.*, **46**, 933–942.

Identification and Characterization of Trimeric Proanthocyanidins of Two Members of the *Rhododendron* Genus (*Ericaceae*) by Liquid Chromatography Multi-Stage Mass Spectrometry

Rakesh Jaiswal, Mohamed G.E. Karar and Nikolai Kuhnert

Jacobs University Bremen, School of Engineering and Science, Chemistry, Bremen, Germany

1   INTRODUCTION

Proanthocyanidins (PAs) are oligomers and polymers of flavan-3-ol units such as afzelechin, epiafzelechin, catechin, epicatechin, gallocatechin, and epigallocatechin (Figure 1). These PAs are the second most abundant natural phenolics after lignin (Hamingway and Karchesy, 1989). PAs are present in many vegetables, fruits (Hong, Barrett, and Mitchell, 2004; Li *et al.*, 2001; Preuss, Bagchi, and Bagchi, 2002), beverages (Kalili and de Villiers, 2010; Omar, Mullen, and Crozier, 2011), and grains (Awika *et al.*, 2004) and are claimed to have potential antioxidant (Bagchi *et al.*, 1997; Serafini, Maiani, and Ferro-Luzzi, 1998), anticancer (Joshi *et al.*, 2000), antimutagenic (Bomser *et al.*, 1999; Ye *et al.*, 1999), antidiabetic (Preuss *et al.*, 2001), anti-inflammatory (Bayeta and Lau, 2000; Erdemoglu *et al.*, 2008), and anti-HIV(He *et al.*, 1997) properties. PAs are also beneficial for healing wounds (Witte *et al.*, 2000), reducing the risks of cardiovascular diseases (Das *et al.*, 1999; Delacroix, 1981; Miyagi, Miwa, and Inoue, 1997) and skin diseases (Ni, Mu, and Gulati, 2002), and

protecting from drug toxicity (Bagchi *et al.*, 2001), UV radiations (Fisher *et al.*, 1996; Saliou *et al.*, 2001), and asthma (Lau *et al.*, 2004).

Rhododendrons are best known for their use in gardens for ornamental purposes. The phytochemical profile of only few species of *Rhododendron* has been investigated so far. This genus is a source of phenolic compounds, especially flavonoids (Harborne and Williams, 1971; Harborne, 1986; Liu *et al.*, 2010) and their glycosides (Yang *et al.*, 2010; Yang and Kong, 2008; Sharma *et al.*, 2010), essential oils (Doss, Hatheway, and Hrutfiord, 1986; Zhao *et al.*, 2006), chromones (Chen *et al.*, 2008), chromanes and chromenes (Kashiwada *et al.*, 2001), β-diketones (Evans *et al.*, 1975), iridoids (Fan *et al.*, 2001), terpenoids (Wang *et al.*, 2010; Zhang *et al.*, 2008; Sakakibara and Kaiya, 1983; Tantry *et al.*, 2011), and steroids (Block and Constant, 1972). Medicinal use of *Rhododendron* plant derived formulations is restricted because of the occurrence of phytotoxins called *grayanotoxins*, a class of terpenoids present in many plants of this genus and family (Zhang *et al.*, 2008; Chen *et al.*, 2004). However,
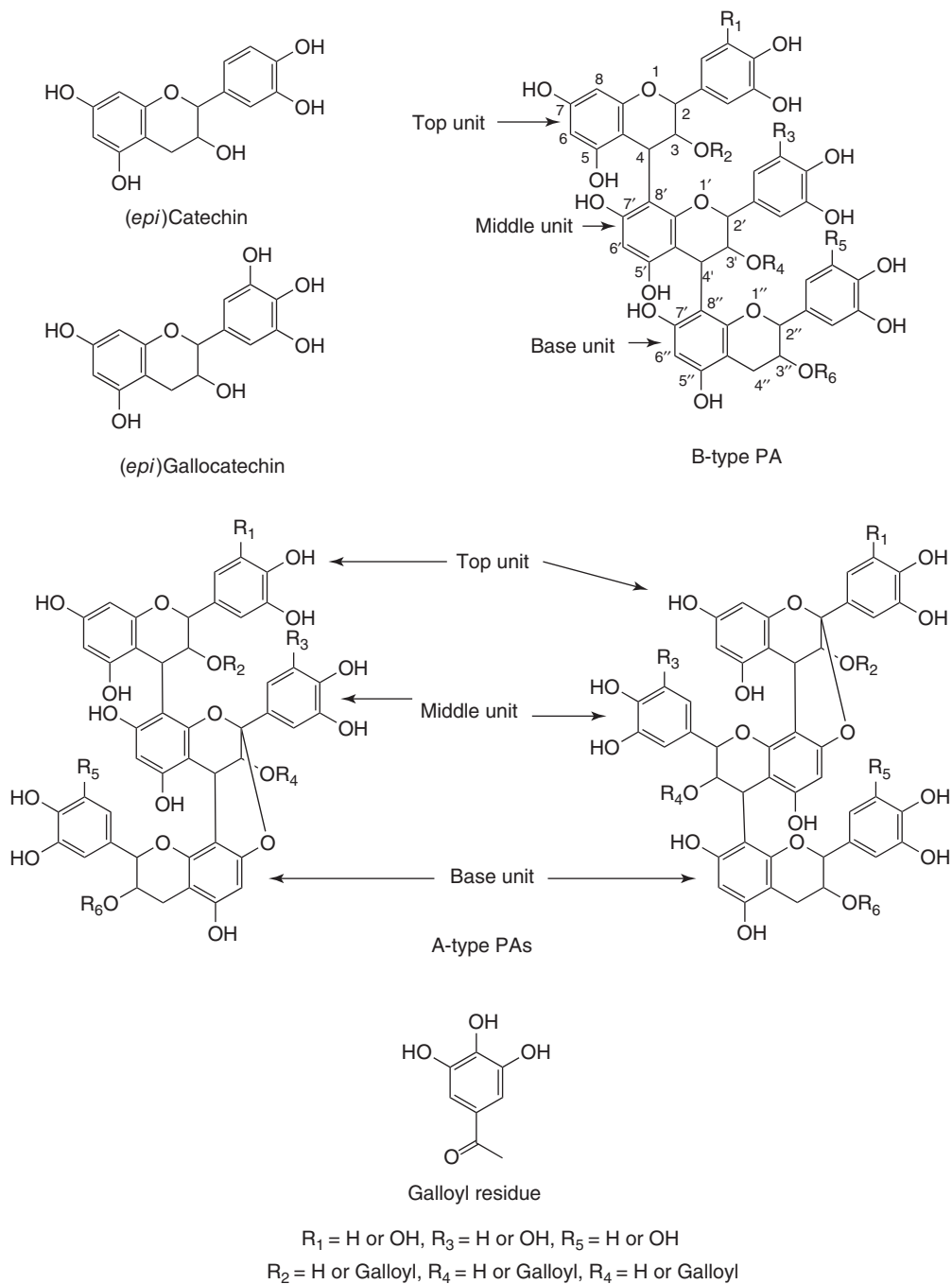
Figure 1　Representative structures of flavan-3-ol units and trimeric proanthocyanidins of *R*. 'Catawbiense Grandiflorum' and *R*. 'Cunningham's White.'

some plants of this genus low in grayanotoxins are commonly used for medicinal purposes in Turkish (Erdemoglu *et al*., 2008), Chinese (Zhang *et al*., 2008; Klocke *et al*., 1991), Homeopathic (Gibson, 1971), and Ayurvedic (Prakash *et al*., 2008) medicinal systems to treat chronic diseases. *Rhododendron* plants are also reported to have antioxidant (Lee *et al*., 2011; Silici, Sagdic, and Ekici, 2010), anti-inflammatory (Erdemoglu *et al*., 2008), antiviral (Kashiwada *et al*., 2001; Gescher *et al*., 2009), and hepatoprotective properties (Prakash *et al*., 2008) because of their phenolic constituents.

In this study, all the PAs with trimeric units analyzed were extracted from the leaves of two *Rhododendron* species, *R*. 'Catawbiense Grandiflorum' and *R*. 'Cunningham's White.' In our previous study, we have reported a series of dimeric PAs in the *Rhododendron* genus and developed a tandem mass spectrometric hierarchical scheme for the identification and characterization of PAs (Jaiswal, Jayasinghe, and Kuhnert, 2012). In this study, we are extending our tandem mass spectrometric hierarchical scheme for the identification and characterization of trimeric PAs without any purification or isolation.

2 IDENTIFICATION AND CHARACTERIZATION OF PROANTHOCYANIDINS

All the data presented in this contribution was generated by analyzing the aqueous methanolic (70%) extracts of the plants leaves by Iontrap MS (negative ion mode), TOF-MS (negative ion mode), and RP-HPLC using a C18 amide column.

In the following section, we discuss in detail the identity of the PAs identified in the two plants under investigation. In this contribution, we focus exclusively on trimeric PAs. In terms of PA nomenclature, we use the system suggested by Porter (1988). Here, for trimeric structures, the (*epi*)catechin unit bearing a (*epi*)catechin substituent at the C-4 position of its C-ring is designated as the top unit, whereas the C-4‴ unsubstituted catechin unit is designated as the base unit. Numbering of the (*epi*)catechin rings follows regular IUPAC system with the base unit carbons designated by a dash. The detailed PAs profile of the plants under investigation is given in Table 1.

The PAs were positively identified by their typical UV-absorptions at 280 nm. All the

PAs showed the loss of the galloyl/gallic acid part, retro Diels–Alder fragmentation, and heterocyclic ring fission (HRF) in the negative ion mode.

For positive identification and characterization of PAs, the following points were considered:

1. UV spectrum at 280 nm (λ_{max}).
2. Molecular ion peaks $(M - H)^-$ in negative ion mode of MS, for example, at *m/z* 865, 881, 897, and 1017 for B-type PAs and at 863, 879, 895, and 1015 for A-type PAs.
3. The fragmentation pathway described by Gu *et al*., for example, HRF and retro-Diels–Alder (RDA) fragmentation give information about the hydroxylation of the B-rings and bonds between two monomeric units, and quinone methide (QM) fragmentation defines the three monomeric units and especially the base unit (Figure 2). (Gu *et al*., 2003) The loss of H_2O molecule from the C-rings indicates the epicatechin unit.
4. The regiochemistry of the galloyl residue was assigned based on the loss of a gallic acid (170 Da, neutral loss) molecule followed by a dehydrated RDA fragment. (Jaiswal, Jayasinghe, and Kuhnert, 2012).

2.1 Characterization of (*epi*)Catechin-(4,8′)-(*epi*)catechin-(4′,8″)-(*epi*)catechin (M_r 866)

Seven peaks were detected at *m/z* 865 in the extracted ion chromatogram (EIC). These seven compounds were tentatively assigned as trimeric B-type PAs (**1–7**) with (*epi*)catechin monomeric units. All these compounds have similar MS^n fragmentation patterns with similar ion intensities (Figures 2 and 3). They produced the MS^2 base peak at *m/z* 695 $([M - H^+ - 170\,Da]^-)$ by the loss of an RDA fragment (152 Da) followed by the loss of a water molecule (18 Da); secondary peaks at *m/z* 577 $([(epi)\text{catechin-}(epi)\text{catechin-}H^+]^-)$ originate from a QM fragment, at *m/z* 739 $([M - H^+ - 126\,Da]^-)$ from an HRF fragment, and at *m/z* 847 $([M - H^+ - H_2O]^-)$ from the loss of a water molecule (Figures 2 and 3). Loss of water was observed for all the seven compounds. Considering that authentic reference compound (PA C1) studies composed of epicatechin moieties show this loss of water, we propose that

Table 1 Proanthocyanidins of *R.* '*Catawbiense Grandiflorum*' and *R.* '*Cunningham*'*s White*' and their retention times (P = presence and − = absence).

No.	Proanthocyanidin	MW	*m/z* [M − H⁺]	RT (min)	*R.* '*Cunningham*'*s White*'	*R.* '*Catawbiense Grandiflorum*'
1	(*epi*)Catechin-(4,8′)-(*epi*)catechin-(4′,8″)-(*epi*)catechin	866	865	6.9	P	P
2	(*epi*)Catechin-(4,8′)-(*epi*)catechin-(4′,8″)-(*epi*)catechin	866	865	14.9	P	−
3	(*epi*)Catechin-(4,8′)-(*epi*)catechin-(4′,8″)-(*epi*)catechin	866	865	15.6	P	−
4	(*epi*)Catechin-(4,8′)-(*epi*)catechin-(4′,8″)-(*epi*)catechin	866	865	16.5	P	−
5	(*epi*)Catechin-(4,8′)-(*epi*)catechin-(4′,8″)-(*epi*)catechin	866	865	18.2	P	−
6	(*epi*)Catechin-(4,8′)-(*epi*)catechin-(4′,8″)-(*epi*)catechin	866	865	19.0	P	−
7	(*epi*)Catechin-(4,8′)-(*epi*)catechin-(4′,8″)-(*epi*)catechin	866	865	24.9	P	−
8	(*epi*)Catechin-(4,8′/2,7′)-(*epi*)catechin-(4′,8″)-(*epi*)catechin	864	863	20.8	P	P
9	(*epi*)Catechin-(4,8′/2,7′)-(*epi*)catechin-(4′,8″)-(*epi*)catechin	864	863	21.3	P	−
10	(*epi*)Catechin-(4,8′/2,7′)-(*epi*)catechin-(4′,8″)-(*epi*)catechin	864	863	25.2	P	P
11	(*epi*)Catechin-(4,8′)-(*epi*)catechin-(4′,8″/2′,7″)-(*epi*)catechin	864	863	31.0	P	−
12	(*epi*)Gallocatechin-(4,8′)-(*epi*)catechin-(4′,8″)-(*epi*)catechin	882	881	11.9	P	P
13	(*epi*)Catechin-(4,8′)-(*epi*)gallocatechin-(4′,8″)-(*epi*)catechin	882	881	12.2	P	P
14	(*epi*)Gallocatechin-(4,8′/2,7′)-(*epi*)catechin-(4′,8″)-(*epi*)catechin	880	879	16.7	P	−
15	(*epi*)Catechin-(4,8′/2,7′)-(*epi*)gallocatechin-(4′,8″)-(*epi*)catechin	880	879	17.4	P	P
16	(*epi*)Gallocatechin-(4,8′/2,7′)-(*epi*)catechin-(4′,8″)-(*epi*)catechin	880	879	18.8	P	P
17	(*epi*)Catechin-(4,8′/2,7′)-(*epi*)gallocatechin-(4′,8″)-(*epi*)catechin	880	879	19.6	P	−
18	(*epi*)Gallocatechin-(4,8′)-(*epi*)gallocatechin-(4′,8″)-(*epi*)catechin	898	897	9.6	−	P
19	(*epi*)Gallocatechin-(4,8′/2,7′)-(*epi*)gallocatechin-(4′,8″)-(*epi*)catechin	896	895	14.1	−	P
20	(*epi*)Catechin-(4,8′/2,7′)-(*epi*)gallocatechin-(4′,8″)-(*epi*)gallocatechin	896	895	14.6	−	P
21	(*epi*)Gallocatechin-(4,8′/2,7′)-(*epi*)catechin-(4′,8″)-(*epi*)gallocatechin	896	895	15.4	−	P
22	(*epi*)Catechin-(4,8′)-3′-*O*-galloyl-(*epi*)catechin-(4′,8″)-(*epi*)catechin	1018	1017	25.1	P	−
23	(*epi*)Catechin-(4,8′)-3′-*O*-galloyl-(*epi*)catechin-(4′,8″)-(*epi*)catechin	1018	1017	27.5	P	−
24	(*epi*)Catechin-(4,8′)-3′-*O*-galloyl-(*epi*)catechin-(4′,8″)-(*epi*)catechin	1018	1017	29.1	P	−
25	3-*O*-Galloyl-(*epi*)catechin-(4,8′/2,7′)-(*epi*)catechin-(4′,8″)-(*epi*)catechin	1016	1015	26.6	−	P
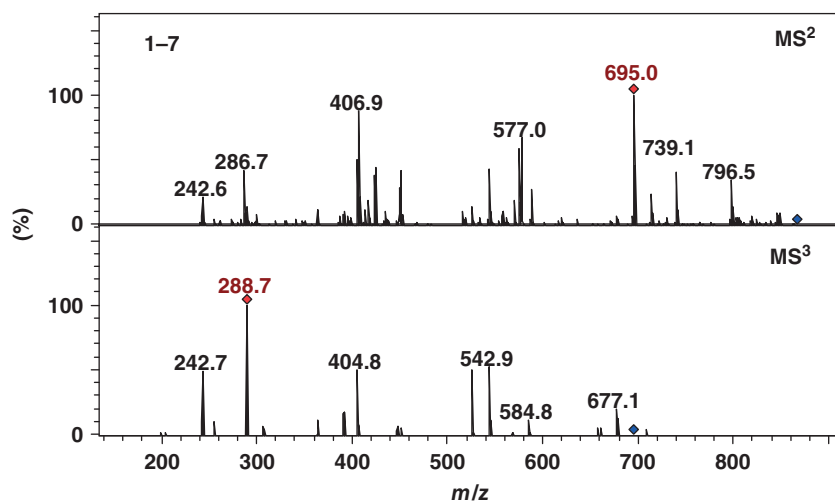
Figure 2 MS³ of trimeric B-type proanthocyanidins (**1–7**) at *m/z* 865 in negative ion mode.
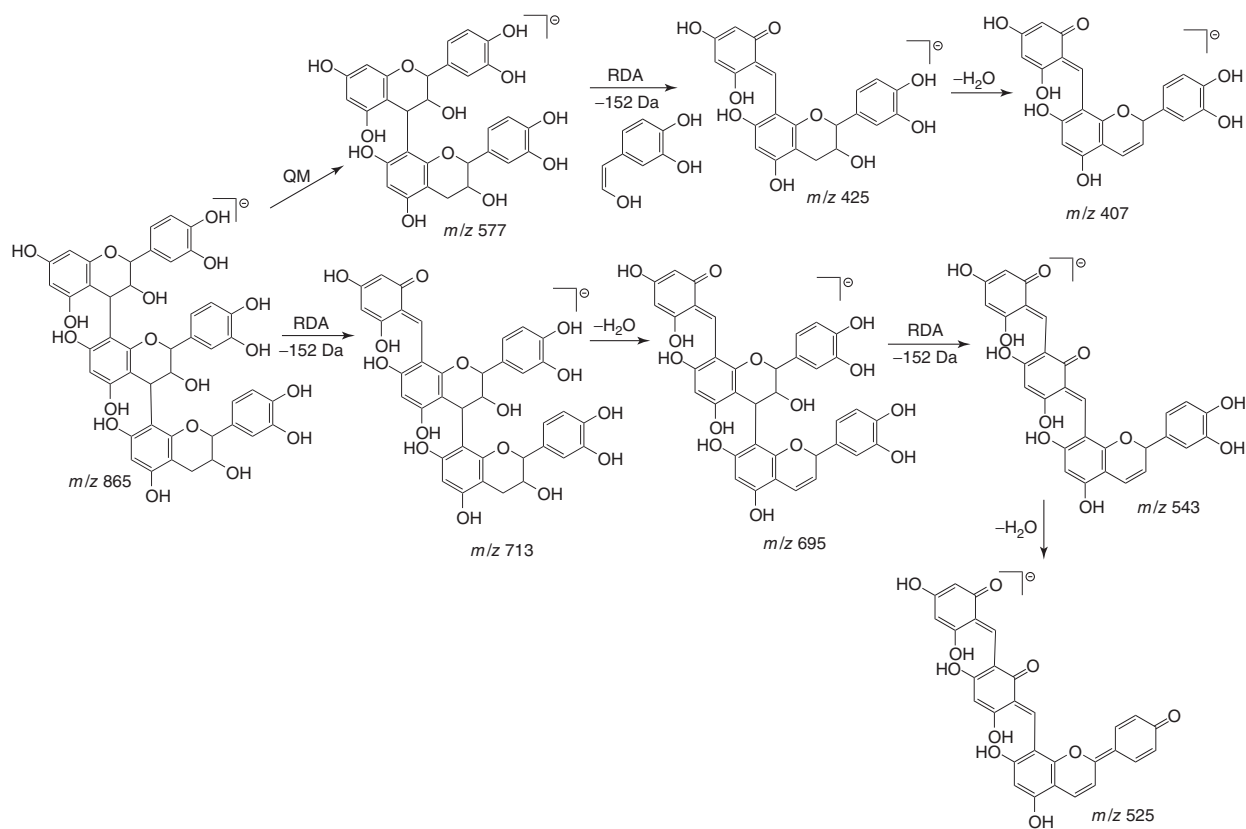
Figure 3 Fragmentation pathways of trimeric B-type proanthocyanidins (*m/z* 865).
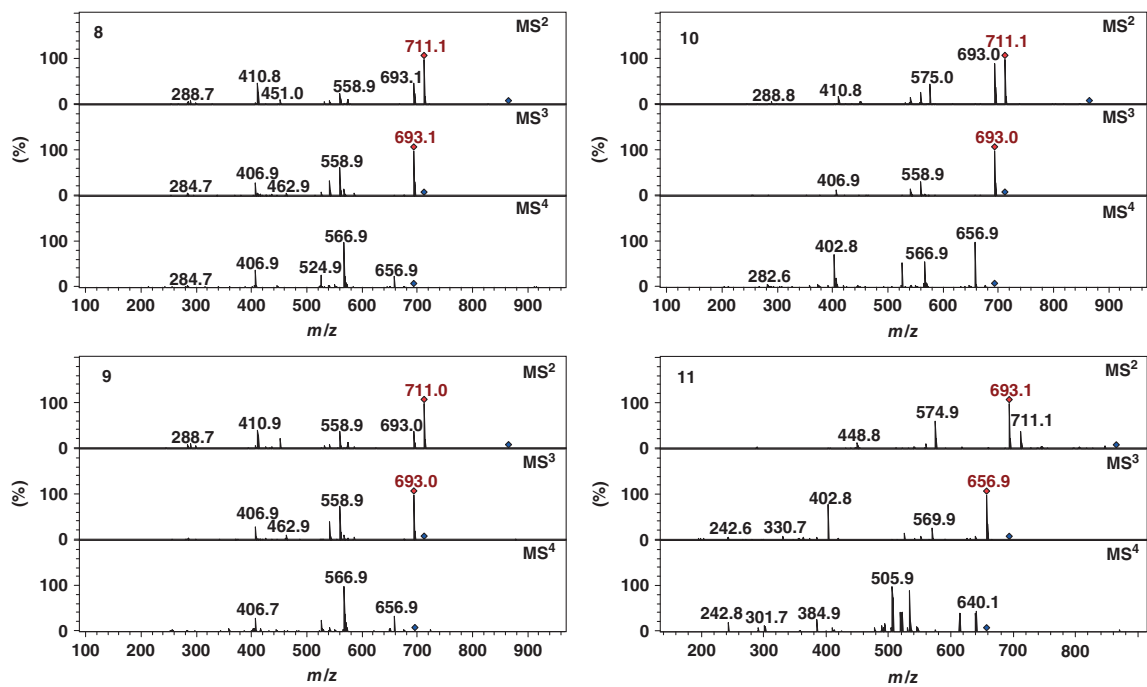
Figure 4 MS^4 of trimeric A-type proanthocyanidins (**8–11**) at *m/z* 863 in negative ion mode.

compounds having two neighboring catechin units do not show this fragmentation pathway. Using this hypothesis, it is possible to distinguish two PAs with 4,6′ and 4,8′ connectivity based on two catechin moieties. On the basis of the above fragmentation mechanism, these PAs were assigned as isomers of (*epi*)catechin-(4,8′)-(*epi*)catechin-(4′,8″)-(*epi*)catechin trimer. For further evidence, PA C1 was used as an authentic standard and showed retention time and fragmentation identical to PA **2**. On the basis of these arguments, isomer **2** was assigned as PA C1 and isomers **1** and **3–7** were assigned as B-type trimers. It was not possible to assign the stereochemistry of monomeric units of a PA by tandem mass spectrometry.

2.2 Characterization of (*epi*)Catechin-(4,8′/2, 7′)-(*epi*)catechin-(4′,8″)-(*epi*)catechin and (*epi*)Catechin-(4,8′)-(*epi*)catechin-(4′,8″/ 2′,7″)-(*epi*)catechin (M_r 864)

Four peaks were detected at *m/z* 863 in the EIC. These four compounds were tentatively assigned as trimeric A-type PAs (**8–11**) with an A-type linkage and (*epi*)catechin monomeric units. Isomers **8–10** produced the MS² base peak at *m/z* 711 ([M − H⁺ − 152 Da]⁻) by the loss of an RDA

fragment (152 Da); they produced secondary peaks at *m/z* 693 ([M − H⁺ − 170 Da]⁻) by the loss of an RDA fragment (152 Da) followed by the loss of a water molecule (18 Da), at *m/z* 289 ([(*epi*)catechin-H⁺]⁻) from a QM fragment, and at *m/z* 411 and 451 from the HRF fragments (Figures 4 and 5). The presence of a QM fragment at *m/z* 289 and of the HRF fragments at *m/z* 411 and 451 showed that the top and middle units were connected by an A-type linkage. On the basis of the above points, isomers **8–10** were assigned as (*epi*)catechin-(4,8′/2,7′)-(*epi*)catechin-(4′,8″)-(*epi*)catechin.

Isomer **11** produced the MS² base peak at *m/z* 693 ([M − H⁺ − 170 Da]⁻) by the loss of an RDA fragment (152 Da) followed by the loss of a water molecule (18 Da); it produced secondary peaks at *m/z* 711 ([M − H⁺ − 152 Da]⁻) by the loss of an RDA fragment (152 Da), *m/z* 449 from an HRF fragment, and at *m/z* 575 ([(*epi*)catechin-(4′,8″/2′7″)-(*epi*)catechin-H⁺]⁻) from a QM fragment (Figures 4 and 5). The presence of a QM fragment at *m/z* 575 showed that the middle and the bottom units were connected by an A-type linkage that was also confirmed by the HRF at *m/z* 449 (Figure 5). On the basis of the above points, isomer **11** was assigned as (*epi*)catechin-(4,8′)-(*epi*)catechin-(4′,8″/2′,7″)-(*epi*)catechin.
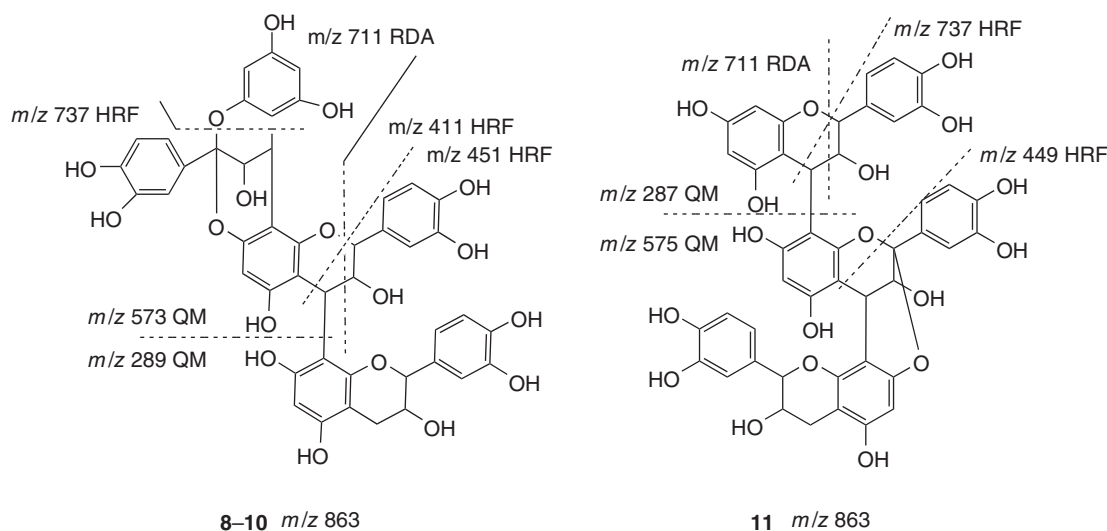
Figure 5 Fragmentation pathways of trimeric A-type proanthocyanidins (*m/z* 863).

2.3 Characterization of (*epi*)Gallocatechin-(4,8′)-(*epi*)catechin-(4′,8″)-(*epi*)catechin and (*epi*)catechin-(4,8′)-(*epi*)gallocatechin-(4′,8″)-(*epi*)catechin (M_r 882)

Two peaks were detected at *m/z* 881 in the EIC and were tentatively assigned as trimeric B-type PAs (**12** and **13**) with two (*epi*)catechin and one (*epi*)gallocatechin monomeric units. Isomer **12** produced the MS^2 base peak at *m/z* 695 ([M − H$^+$ − 186 Da]$^-$) by the loss of an RDA fragment (168 Da) from the top unit of the trimer followed by the loss of a water molecule (18 Da); secondary peaks at *m/z* 577

([(*epi*)catechin-(*epi*)catechin-H$^+$]$^-$), and *m/z* 303 ([(*epi*)gallocatechin-H$^+$]$^-$) originated from a QM fragment (Figures 6 and 7). The presence of a QM fragment at *m/z* 577 showed that the base and middle units were (*epi*)catechin and a QM fragment at *m/z* 303 showed that the top unit was (*epi*)gallocatechin. On the basis of the above arguments, isomer **12** was assigned as (*epi*)gallocatechin-(4,8′)-(*epi*)catechin-(4′,8″)-(*epi*) catechin. Isomer **13** produced the MS^2 base peak at *m/z* 711 ([M − H$^+$ − 170 Da]$^-$) by the loss of an RDA fragment (152 Da) followed by the loss of a water molecule (18 Da); it produced secondary peaks at *m/z* 593 ([(*epi*)gallocatechin-(*epi*)catechin-H$^+$]$^-$)
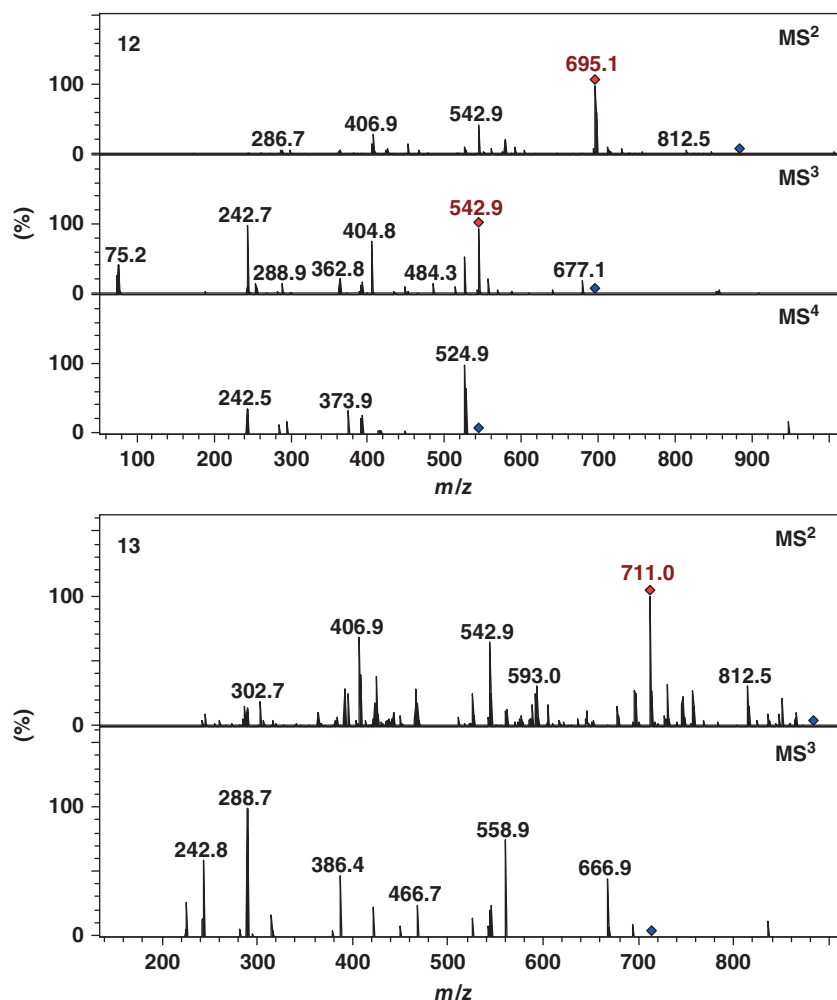
Figure 6 MS^4 of trimeric B-type proanthocyanidins (**12** and **13**) at *m/z* 881 in negative ion mode.
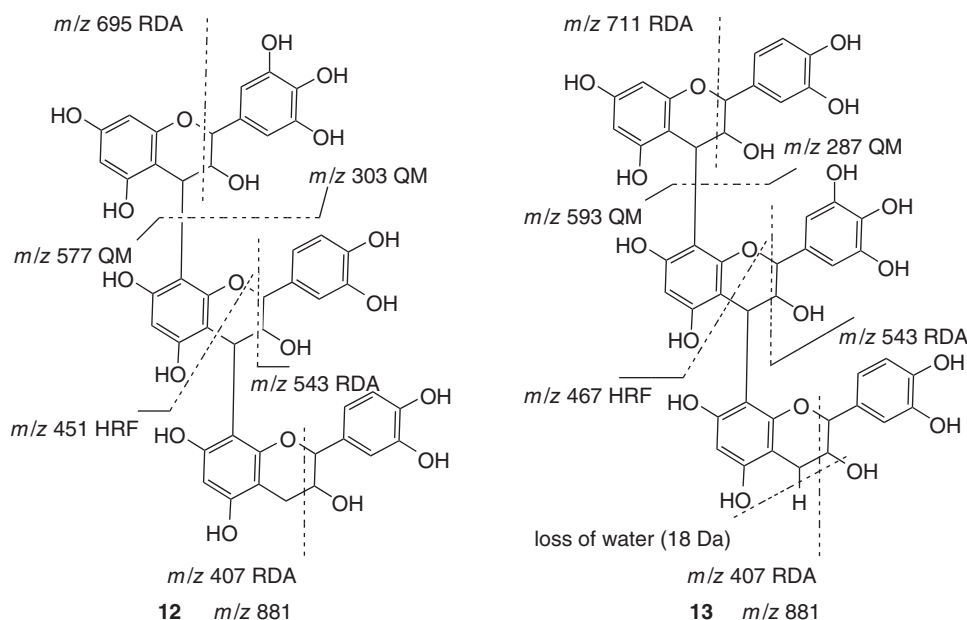
Figure 7   Fragmentation pathways of trimeric B-type proanthocyanidins **12** and **13** at m/z 881.

and at m/z 289 ([(*epi*)catechin-H$^+$]$^-$) from a QM fragment (Figures 6 and 7). The presence of a QM fragment at m/z 289 showed that the base ring was a (*epi*)catechin unit. Further confirmation came from a QM fragment at m/z 593 ([M − H$^+$ − 170 Da]$^-$) ([(*epi*)gallocatechin-(*epi*)catechin-H$^+$]$^-$)   that showed that the middle unit was (*epi*)gallocatechin and the base unit was (*epi*)catechin. On the basis of the above arguments, isomer **13** was assigned as (*epi*)catechin-(4,8′)-(*epi*)gallocatechin-(4′,8″)-(*epi*) catechin.

2.4   Characterization of (*epi*)Gallocatechin-(4, 8′/2,7′)-(*epi*)catechin-(4′,8″)-(*epi*) catechin and (*epi*)Catechin-(4,8′/2,7′)- (*epi*)gallocatechin-(4′,8″)-(*epi*)catechin (M_r 879)

Four peaks were detected at m/z 879 in the EIC and were tentatively assigned as trimeric A-type PAs (**14–17**) with two (*epi*)catechin and one (*epi*)gallocatechin monomeric units. Isomers **14** and **16** produced the MS^2 base peak at m/z 727 ([M − H$^+$ − 152 Da]$^-$) by the loss of an RDA fragment (152 Da) from the middle unit; they produced secondary peaks at m/z 709 ([M − H$^+$ − 170 Da]$^-$) by

the loss of an RDA fragment followed by the loss of a water molecule (18 Da) from the middle unit, and at m/z 575 ([M − H$^+$ − 304 Da]$^-$) by the loss of an RDA fragment (152 Da) from the middle unit followed by another loss of an RDA fragment (152 Da) from the base unit (Figures 8 and 9). They produced MS^2 secondary peaks at m/z 289 ([(*epi*)catechin-H$^+$]$^-$) from a QM fragment and at m/z 427 and 451 from the HRF fragments (Figures 8 and 9). The presence of a QM fragment at m/z 289 showed that the base unit was an (*epi*)catechin, whereas the HRF fragments at m/z 427 and 451 showed that the upper unit was (*epi*)gallocatechin and the middle and bottom units were (*epi*)catechin. On the basis of the above points, isomers **14** and **16** were assigned as (*epi*)gallocatechin-(4,8′/2,7′)-(*epi*)catechin-(4′,8″)- (*epi*)catechin.

It was interesting to notice that a secondary peak at m/z 575 originated from the RDA fragments but in the case of A-type trimeric PAs of (*epi*)catechin units, this ion originated from the QM fragment.

Isomers **15** and **17** produced the MS^2 base peak at m/z 711 ([M − H$^+$ − 168 Da]$^-$) by the loss of an RDA fragment from the middle unit; they produced secondary peaks at m/z 693 ([M − H$^+$ − 186 Da]$^-$) by the loss of an RDA fragment (168 Da) from the middle unit of the trimer followed by the loss of a water
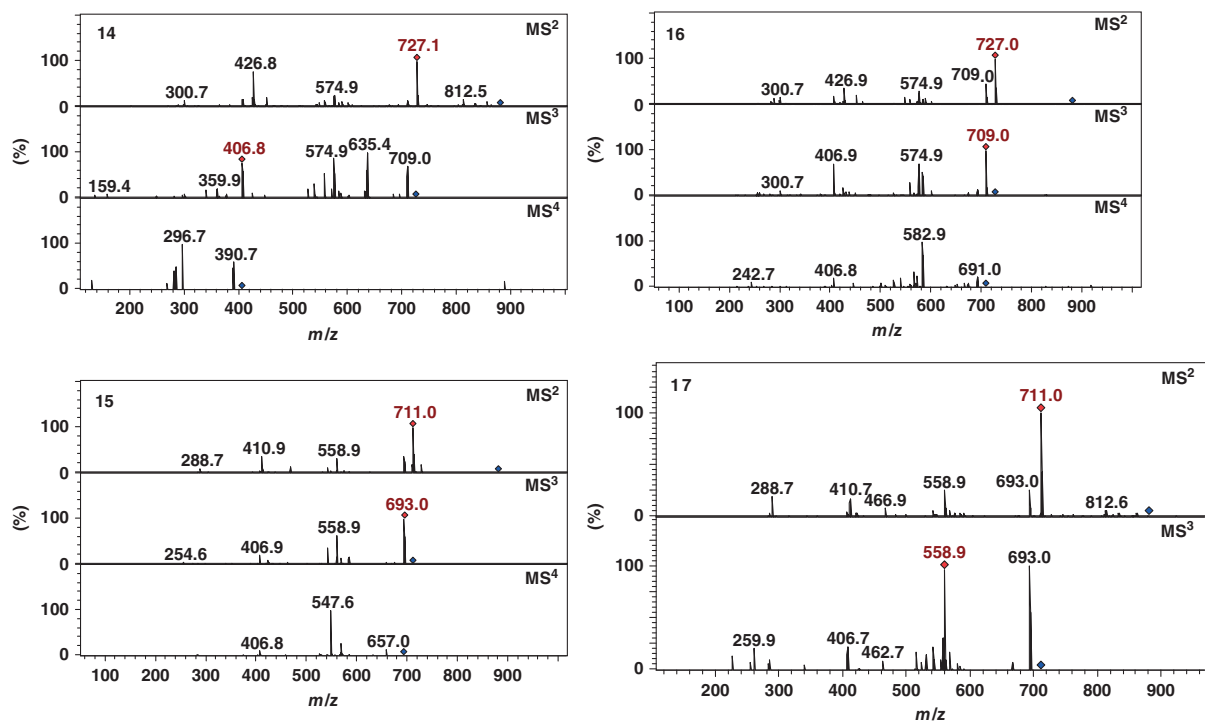
Figure 8 MS[4] of trimeric A-type proanthocyanidins (**14–17**) at m/z 879 in negative ion mode.
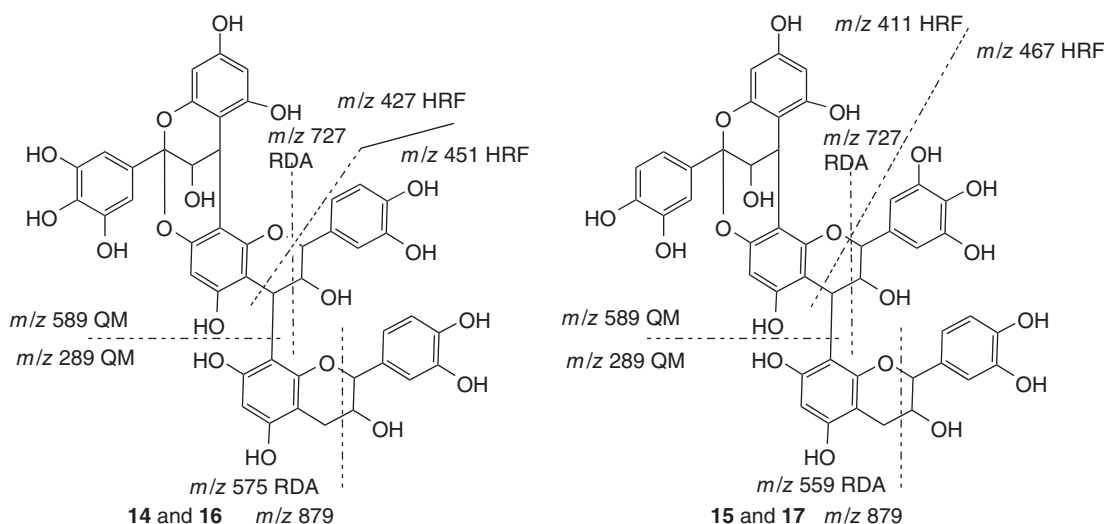
Figure 9 Fragmentation pathways of trimeric A-type proanthocyanidins at *m/z* 879.

molecule (18 Da), at *m/z* 559 ([M − H⁺ − 320 Da]⁻) by the loss of an RDA fragment (168 Da) from the middle unit followed by another loss of an RDA fragment (152 Da) from the base unit of the trimer; the peak at *m/z* 289 ([(*epi*)catechin-H⁺]⁻) originated from a QM fragment (Figures 8 and 9). The presence of a QM fragment at *m/z* 289 showed that the base unit was an (*epi*)catechin unit and the HRF fragments at *m/z* 411 and 467 showed that the upper unit was (*epi*)gallocatechin, whereas the middle unit was (*epi*)gallocatechin. The RDA fragmentation on the middle unit gave a fragment ion (*m/z* 711) with a larger π–π hyperconjugated system that is energetically more favorable than the RDA on the base unit (Jaiswal, Jayasinghe, and Kuhnert, 2012). On the basis of the above arguments, isomers **15** and **17** were assigned as (*epi*)catechin-(4,8′/2,7′)-(*epi*)gallocatechin-(4′,8″)-(*epi*)catechin (*M*r 879).

2.5 Characterization of (*epi*)Gallocatechin-(4,8′)-(*epi*)gallocatechin-(4′,8″)-(*epi*)catechin (*M*r 897)

One peak was detected at *m/z* 897 in the EIC and was tentatively assigned as a trimeric B-type PA **18** with two (*epi*)gallocatechin and one (*epi*)catechin monomeric units. Compound **18** produced the MS² base peak at *m/z* 711 ([M − H⁺ − 186 Da]⁻)

by the loss of an RDA fragment (168 Da) from the top unit of the trimer followed by the loss of a water molecule (18 Da); it produced secondary peaks at *m/z* 543 by the loss of two RDA fragments (168 + 168 Da) from the top and middle units of the trimer, at *m/z* 467 from an HRF fragment, and at *m/z* 289 ([(*epi*)catechin-H⁺]⁻) from a QM fragment (Figures 10 and 11). The presence of a QM fragment at *m/z* 289 showed that the bottom unit was (*epi*)catechin and an HRF fragment at *m/z* 467 showed that the middle unit was (*epi*)gallocatechin. On the basis of the above arguments, isomer **18** was assigned as (*epi*)gallocatechin-(4,8′)-(*epi*)gallocatechin-(4′,8″)-(*epi*)catechin.

2.6 Characterization of (*epi*)Gallocatechin-(4,8′/2,7′)-(*epi*)gallocatechin-(4′,8″)-(*epi*)catechin, (*epi*)catechin-(4,8′/2,7′)-(*epi*)gallocatechin-(4′,8″)-(*epi*)gallocatechin and (*epi*)Gallocatechin-(4,8′/2,7′)-(*epi*)catechin-(4′,8″)-(*epi*)gallocatechin (*M*r 896)

Three peaks were detected at *m/z* 895 in the EIC and were tentatively assigned as trimeric A-type PAs (**19–21**) with one (*epi*)catechin monomeric and two (*epi*)gallocatechin units. Isomer **19** produced the MS² base peak at *m/z* 727 ([M − H⁺ − 168 Da]⁻)
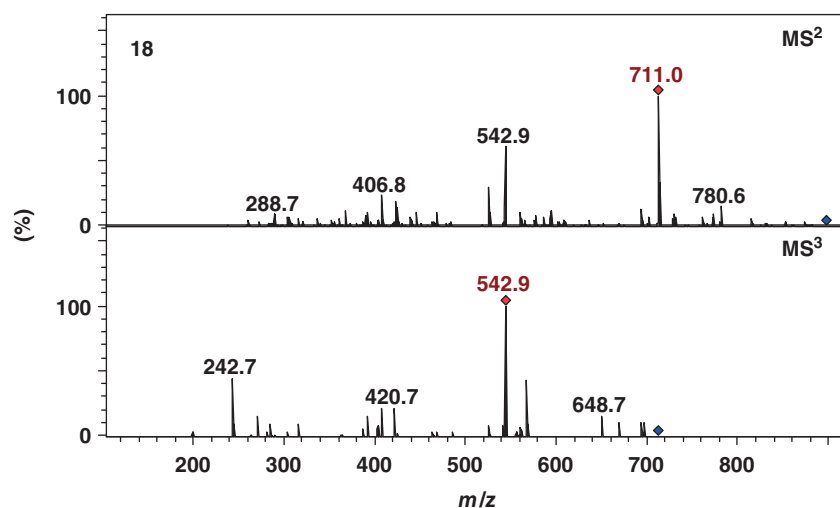
Figure 10   MS^3 of trimeric B-type proanthocyanidin (**18**) at *m/z* 897 in negative ion mode.
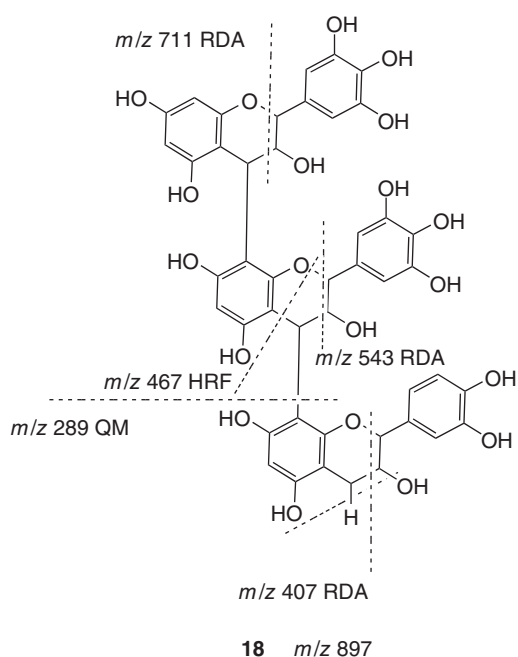
Figure 11   Fragmentation pathways of trimeric B-type proanthocyanidin at *m/z* 897.

by the loss of an RDA fragment from the middle unit; it produced secondary peaks at *m/z* 709 ([M − H$^+$ − 186 Da]$^-$) by the loss of an RDA fragment (168 Da) from the middle unit of the trimer followed by the loss of a water molecule (18 Da), at

m/z 575 ([M − H$^+$ − 320 Da]$^-$) by the loss of an RDA fragment (168 Da) from the middle unit followed by another loss of an RDA fragment (152 Da) from the base unit of the trimer, at *m/z* 427 and 467 from the HRF fragments, and at *m/z* 289 ([(*epi*)catechin-H$^+$]$^-$) from a QM fragment (Figures 12 and 13). The presence of a QM fragment at *m/z* 289 showed that the base unit was −(*epi*)catechin unit and the HRF fragments at *m/z* 427 and 467 showed that the middle and top units were (*epi*)gallocatechin connected by an A-type linkage. On the basis of the above points, isomer **19** was assigned as (*epi*)gallocatechin-(4,8′/2,7′)-(*epi*)gallocatechin-(4′, 8″)-(*epi*)catechin (Figure 14).

Isomer **20** produced the MS^2 base peak at *m/z* 727 ([M − H$^+$ − 168 Da]$^-$) by the loss of an RDA fragment from the middle unit; it produced secondary peaks at *m/z* 709 ([M − H$^+$ − 186 Da]$^-$) by the loss of an RDA fragment (168 Da) from the middle unit of the trimer followed by the loss of a water molecule (18 Da), at *m/z* 559 ([M − H$^+$ − 336 Da]$^-$) by the loss of an RDA fragment (168 Da) from the middle unit followed by another loss of an RDA fragment (168 Da) from the base unit of the trimer, at *m/z* 411 from the HRF fragment and at *m/z* 305 ([(*epi*)gallocatechin-H$^+$]$^-$) from a QM fragment (Figures 12 and 13). The presence of a QM fragment at *m/z* 305 showed that the bottom unit was (*epi*)gallocatechin and the HRF fragment at *m/z* 411 showed that the top unit was
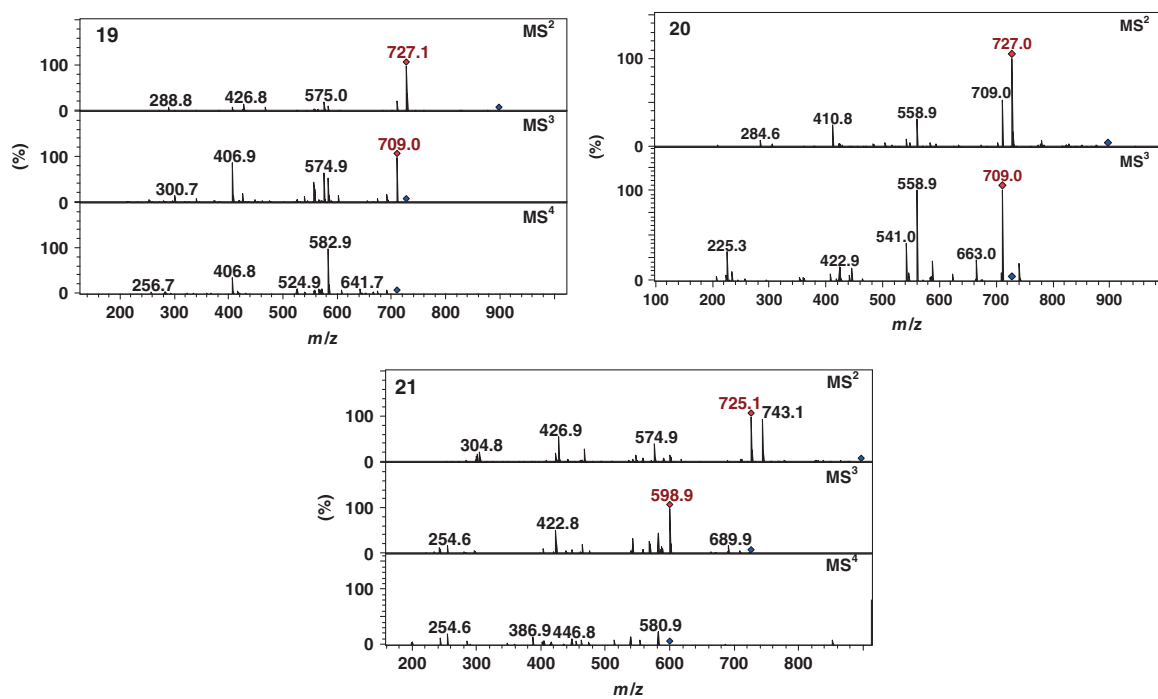
Figure 12 MS^4 of trimeric A-type proanthocyanidins (**19–21**) at *m/z* 895 in negative ion mode.
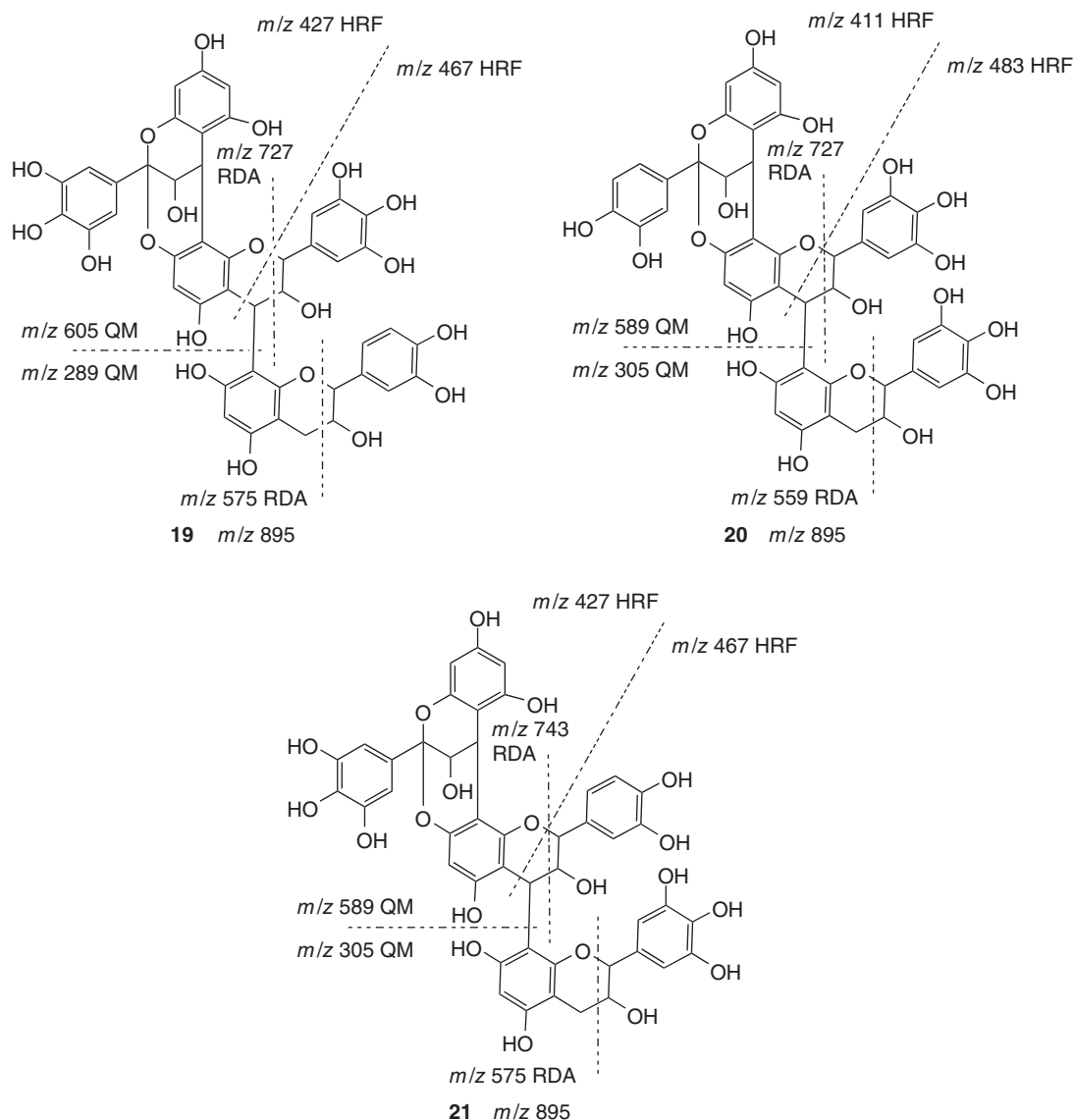
Figure 13 Fragmentation pathways of trimeric A-type proanthocyanidins (**19–21**) at *m/z* 895.

(*epi*)catechin. On the basis of the above points, isomer **20** was assigned as (*epi*)catechin-(4,8′/2,7′)-(*epi*)gallocatechin-(4′,8″)-(*epi*)gallocatechin.

Isomer **21** produced the MS^2 base peak at *m/z* 725 ([M − H$^+$ − 170 Da]$^-$) by the loss of an RDA fragment (152 Da) from the middle unit followed by the loss of a water molecule (18 Da); it produced secondary peaks at *m/z* 743 ([M − H$^+$ − 152 Da]$^-$) by the loss of an RDA fragment (152 Da) from the middle unit of the

trimer, at *m/z* 575 ([M − H$^+$ − 320 Da]$^-$) by the loss of an RDA fragment (152 Da) from the middle unit followed by another loss of an RDA fragment (168 Da) from the base unit of the trimer, at *m/z* 427 and 467 from the HRF fragments, and at *m/z* 305 ([(*epi*)gallocatechin-H$^+$]$^-$) from a QM fragment (Figures 12 and 13). The presence of a QM fragment at *m/z* 305 showed that the base unit was (*epi*)gallocatechin and the HRF fragment at *m/z* 427 showed that the
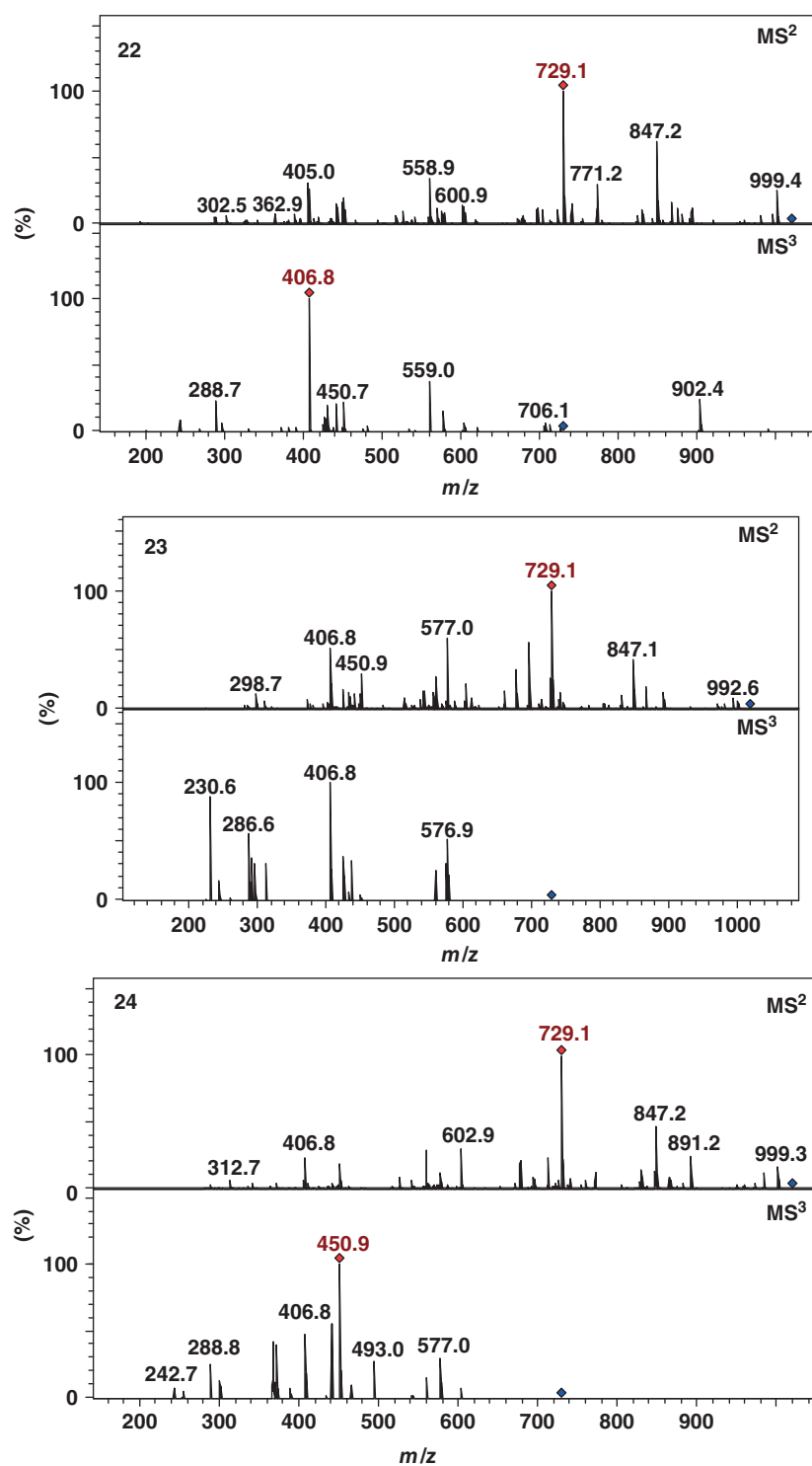
Figure 14 MS^3 of the gallates of trimeric B-type proanthocyanidins (**22–24**) at *m/z* 1017 in negative ion mode.

top unit was (*epi*)gallocatechin. On the basis of the above points, isomer **21** was assigned as (*epi*)gallocatechin-(4,8′/2,7′)-(*epi*)catechin-(4′,8″)-(*epi*)gallocatechin.

2.7 Characterization of (*e*pi)catechin-(4,8′)-3′-*O*-galloyl-(*epi*)catechin-(4′,8″)-(*epi*) catechin (M_r 1018)

Three peaks were detected at *m/z* 1017 in the EIC and were tentatively assigned as gallates of the trimeric PAs with (*epi*)catechin monomeric units. Isomers (**22** and **23**) produced the MS^2 base peak at *m/z* 729 ([M − H$^+$ − 288 Da]$^-$) by the loss of a (*epi*)catechin unit from the top; the secondary peaks were as following: the peak at *m/z* 847 ([M − H$^+$-gallic acid]$^-$) resulted from the loss of a gallic acid molecule, the peak at *m/z* 451 ([M − H$^+$ − 288 Da − 152 Da − 126 Da]$^-$) originated from the HRF, and the one at *m/z* 407 ([M − H$^+$ − 288 Da-gallic acid − 152 Da]$^-$) originated from the RDA fragment. They produced the MS^3 secondary peak at *m/z* 289 ([(*epi*)catechin-H$^+$]$^-$) from a QM fragment that showed that the bottom unit was an (*epi*) catechin unit. The MS^2 fragments of these isomers were similar to the MS^2 fragments of 3-*O*-galloyl-(*epi*)catechin-(4,8′)-(*epi*)catechin that was recently reported in the plants of *Rhododendron* genus (Jaiswal, Jayasinghe, and Kuhnert,

2012). On the basis of the above points, isomers **22** and **23** were assigned as (*epi*)catechin-(4,8′)-3′-*O*-galloyl-(*epi*)catechin-(4′,8″)-(*epi*) catechin (Jaiswal, Jayasinghe, and Kuhnert, 2012). Fragmentation of isomer **24** also produced ions similarly to isomers **22** and **23** except for the MS^3 base peak at *m/z* 451 instead of *m/z* 407. On the basis of these points, isomer **24** was tentatively assigned as (*epi*)catechin-(4,8′)-3′-*O*-galloyl-(*epi*)catechin-(4′, 8″)-(*epi*)catechin.

2.8 Characterization of 3-*O*-Galloyl-(*epi*) catechin-(4,8′/2,7′)-(*epi*)catechin-(4′,8″)- (*epi*)catechin (M_r 1016)

One peak was detected at *m/z* 1015 and was tentatively assigned as a gallate of a trimeric A-type PA with (*epi*)catechin monomeric units. This compound (**25**) produced the MS^2 base peak at *m/z* 693 ([M − H$^+$ − 170 Da − 152 Da]$^-$) by the loss of a gallic acid moiety (170 Da) and an RDA fragment (152 Da); it produced secondary peaks at *m/z* 845 ([M − H$^+$ − 170 Da]$^-$) by the loss of a gallic acid moiety, at *m/z* 451 and 411 from the HRF fragments, at *m/z* 289 from a QM fragment, and at *m/z* 407 from the RDA fragments. It produced the MS^3 base peak at *m/z* 567 from the HRF fragment; it also produced secondary peaks at *m/z* 541 by the loss of an RDA fragment (152 Da), and at *m/z* 407 from an RDA fragment (Figures 15 and 16). The presence of a QM
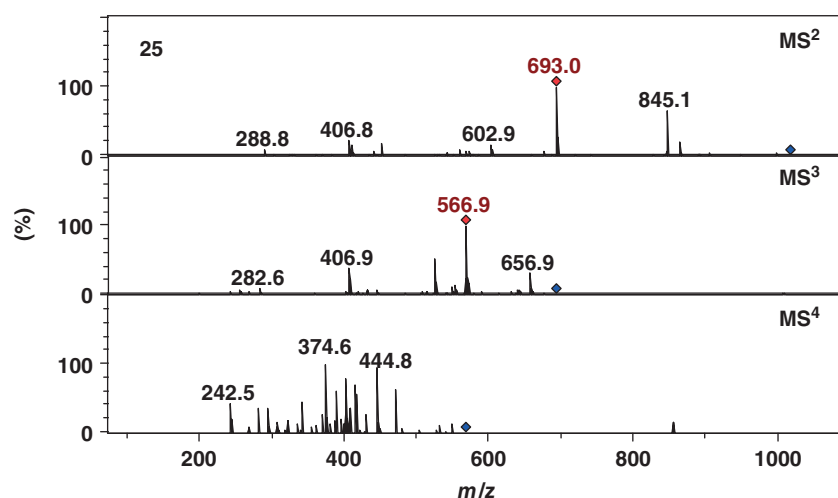
Figure 15   MS^3 of the gallates of trimeric A-type proanthocyanidin (**25**) at *m/z* 1015 in negative ion mode.
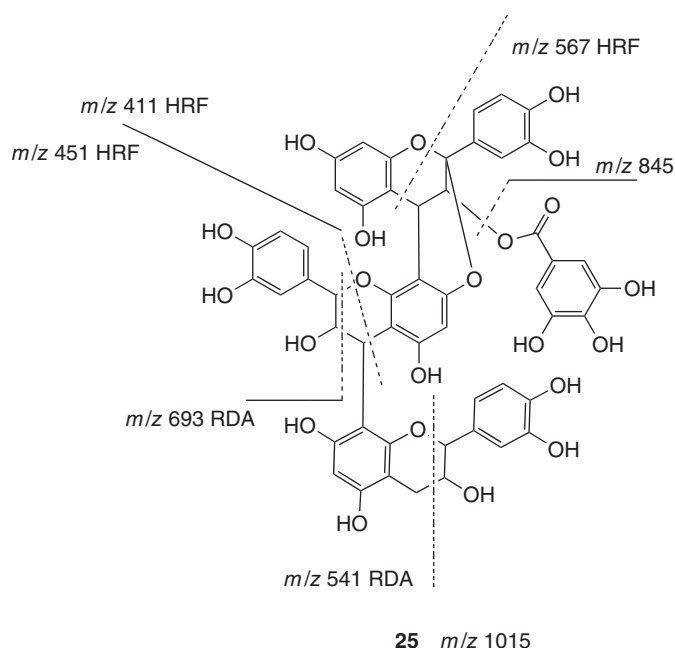
Figure 16 Fragmentation pathways of trimeric A-type proanthocyanidin at m/z 1015.

fragment at m/z 289 confirmed that the base unit was (*epi*)catechin and the HRF fragments at m/z 411 and 451 confirmed that the top and middle (*epi*)catechin units were connected by an A-type linkage. The presence of the MS^3 base peak at m/z 567 showed that HRF was preferred over the second RDA (m/z 541). From this, it was clear that the galloyl group had to be on C-3 of the top unit and it generated a double bond on the C-ring after the loss of a gallic acid moiety. The presence of the double bond on the C-ring favored the loss of an HRF fragment (126 Da) over an RDA fragment (152 Da). On the basis of the above points, this isomer was assigned as 3-*O*-galloyl-(*epi*)catechin-(4,8′/2,7′)-(*epi*)catechin-(4′,8″)-(*epi*)catechin.

3 CONCLUSIONS

We have shown that *Rhododendron* leaves are one of the richest sources of PAs reported in nature considering the number of individual derivatives present. Twenty-five trimeric PAs based on (*epi*)catechin and (*epi*)gallocatechin were detected and characterized on the basis of their unique fragmentation pattern in the negative ion mode tandem MS spectra; all of them for the first time from these sources, with 10 of them previously not reported in nature. The position of the galloyl residue was assigned based on the RDA fragmentation and on the dehydrated RDA fragmentation; it resulted from the loss of gallic acid as a neutral loss in the negative ion mode. For the assignment of PA regiochemistry, we have extended our previously reported hierarchical key of negative ion mode data.

REFERENCES

Awika J. M., Dykes L., Gu L. W., *et al.* (2004) *Abstr. Papers Am. Chem. Soc.*, **228**, U248–U248.

Bagchi D., Garg A., Krohn R. L., *et al*. (1997) *Res. Commun. Mol. Pathol. Pharmacol.*, **95**, 179–189.

Bagchi D., Ray S. D., Patel D., *et al*. (2001) *Drugs Exp. Clin. Res.*, **27**, 3–15.

Bayeta E. and Lau B. H. S. (2000) *Nutr. Res.*, **20**, 249–259.

Block J. H. and Constant G. H. (1972) *Phytochemistry*, **11**, 3279–3282.

Bomser J. A., Singletary K. W., Wallig M. A., *et al*. (1999) *Cancer Lett.*, **135**, 151–157.

Chen S., Zhang H., Wang L., *et al*. (2004) *J. Nat. Prod.*, **67**, 1903–1906.

Chen G., Jin H. Z., Li X. F., *et al*. (2008) *Arch. Pharm. Res.*, **31**, 970–972.

Das D. K., Sato M., Ray P. S., *et al*. (1999) *Drugs Exp. Clin. Res.*, **25**, 115–120.

Delacroix P. (1981) *Rev. Med. Paris*, **22**, 1793–1802.

Doss R. P., Hatheway W. H. and Hrutfiord B. F. (1986) *Phytochemistry*, **25**, 1637–1640.

Erdemoglu N., Akkol E. K., Yesilada E., *et al*. (2008) *J. Ethnopharmacol.*, **119**, 172–178.

Evans D., Knights B. A., Math V. B., *et al*. (1975) *Phytochemistry*, **14**, 2447–2451.

Fan C. Q., Zhao W. M., Ding B. Y., *et al*. (2001) *Fitoterapia*, **72**, 449–452.

Fisher G. J., Datta S. C., Talwar H. S., *et al*. (1996) *Nature*, **379**, 335–339.

Gescher K., Hafezi W., Louis A., *et al*. (2009) *Planta Med.*, **75**, 989–989.

Gibson D. M. (1971) *Br. Homoeopath. J.*, **60**, 294.

Gu L., Kelm M. A., Hammerstone J. F., *et al*. (2003) *J. Mass Spectrom.*, **38**, 1272–1280.

Hamingway R. W. and Karchesy J. J. (1989) Chemistry and significance of condensed tannins, in *Proceedings of the First North American Tannin Conference*, p. 553.

Harborne J. B. (1986) *Phytochemistry*, **25**, 1641–1643.

Harborne J. B. and Williams C. A. (1971) *Phytochemistry*, **10**, 2727–2744.

He J. L., Chen Y. Z., Farzan M., *et al*. (1997) *Nature*, **385**, 645–649.

Hong Y. J., Barrett D. M. and Mitchell A. E. (2004) *J. Agric. Food Chem.*, **52**, 2366–2371.

Jaiswal R., Jayasinghe L. and Kuhnert N. (2012) *J. Mass Spectrom.*, **47**, 502–515.

Joshi S. S., Kuszynski C. A., Bagchi M., *et al*. (2000) *Toxicology*, **155**, 83–90.

Kalili K. M. and de Villiers A. (2010) *J. Sep. Sci.*, **33**, 853–863.

Kashiwada Y., Yamazaki K., Ikeshiro Y., *et al*. (2001) *Tetrahedron*, **57**, 1559–1563.

Klocke J. A., Hu M. Y., Chiu S. F., *et al*. (1991) *Phytochemistry*, **30**, 1797–1800.

Lau B. H. S., Riesen S. K., Truong K. P., *et al*. (2004) *J. Asthma*, **41**, 825–832.

Lee S., Sancheti S. A., Bafna M. R., *et al*. (2011) *J. Med. Plant. Res.*, **5**, 248–254.

Li W. G., Zhang X. Y., Wu Y. J., *et al*. (2001) *Acta Pharmacol. Sin.*, **22**, 1117–1120.

Liu Y. Z., Cao Y. G., Ye J. Q., *et al*. (2010) *Fitoterapia*, **81**, 108–114.

Miyagi Y., Miwa K. and Inoue H. (1997) *Am. J. Cardiol.*, **80**, 1627–1631.

Ni Z., Mu Y. and Gulati O. (2002) *Phytother. Res.*, **16**, 567–571.

Omar M. H., Mullen W. and Crozier A. (2011) *J. Agric. Food Chem.*, **59**, 1363–1369.

Porter L. J. (1988) Flavans and Proanthocyanidins*,* in the Flavonoids*:* Advances in Research Since 1980, Chapman & Hall, London, pp. 21–62.

Prakash T., Fadadu S. D., Sharma U. R., *et al*. (2008) *J. Med. Plant Res.*, **2**, 315–320.

Preuss H. G., Montamarry S., Echard B., *et al*. (2001) *Mol. Cell. Biochem.*, **223**, 95–102.

Preuss H. G., Bagchi D. and Bagchi M. (2002) *Ann. N. Y. Acad. Sci.*, **957**, 250–259.

Sakakibara J. and Kaiya T. (1983) *Phytochemistry*, **22**, 2547–2552.

Saliou C., Rimbach G., Moini H., *et al*. (2001) *Free Radic. Biol. Med.*, **30**, 154–160.

Serafini M., Maiani G. and Ferro-Luzzi A. (1998) *J. Nutr.*, **128**, 1003–1007.

Sharma N., Sharma U. K., Gupta A. P., *et al*. (2010) *J. Food Comp. Anal.*, **23**, 214–219.

Silici S., Sagdic O. and Ekici L. (2010) *Food Chem.*, **121**, 238–243.

Tantry M. A., Khan R., Akbar S., *et al*. (2011) *Chin. Chem. Lett.*, **22**, 575–579.

Wang S., Lin S., Zhu C., *et al*. (2010) *Org. Lett.*, **12**, 1560–1563.

Witte M. B., Thornton F. J., Efron D. T., *et al*. (2000) *Nitric Oxide-Biol. Chem.*, **4**, 572–582.

Yang M. and Kong L. (2008) *Chem. Nat. Compd.*, **44**, 98–99.

Yang M. H., Luo J. G., Huang X. F., *et al*. (2010) *Nat. Prod. Res.*, **24**, 920–925.

Ye X., Krohn R. L., Liu W., *et al*. (1999) *Mol. Cell. Biochem.*, **196**, 99–108.

Zhang W. D., Jin H. Z., Chen G., *et al*. (2008) *Fitoterapia*, **79**, 602–604.

Zhao C., Li X., Liang Y., *et al*. (2006) *Chemometrics Intell. Lab. Syst.*, **82**, 218–228.

Strategies in the Analysis of Flavonoids

Celestino Santos-Buelga and Ana M. González-Paramás

Facultad de Farmacia, Universidad de Salamanca, Salamanca, Spain

1 INTRODUCTION

Flavonoids are one of the largest groups of plant secondary metabolites, which play relevant roles in plant ecology and plant physiology. They comprise several thousand compounds that possess a common C6-C3-C6 skeleton where two aromatic rings (A and B) are linked through a three-carbon bridge that usually, but not always, forms a heterocyclic ring (C) giving rise to a phenylchromane (Figure 1). On the basis of the substitution pattern of ring C, different flavonoids classes are distinguished: flavones, flavonols, flavanones, flavanols, anthocyanins, dihydroflavonols, and isoflavones, in addition to the opened chalcone and dihydrochalcone forms (Table 1). Flavonoids are usually hydroxylated at positions 3, 5, 7, 3′, 4′, and/or 5′ and can be further methylated, acetylated, prenylated, or sulfated. In their natural sources, they may occur in free forms (aglycones), as glycosylated or acylated derivatives, and as oligomeric and polymerized structures, such as the flavan-3-ol–derived condensed tannins [also called as *proanthocyanidin s* (*PAs*)]. But, for flavan-3-ols, most flavonoids occur in plants as *O*-glycosides or, less frequently, *C*-glycosides. The sugar residues are usually linked to 3, 7, or 4′ hydroxyl groups in the case of *O*-glycosides and directly to C-6 or C-8 in the case of *C*-glycosides. The most common sugar substituents are glucose, rhamnose, galactose, and arabinose, and the disaccharides rutinose

(α-L-rhamnopyranosyl-(1→6)-β-D-glucopyranose) and neohesperidose (α-l-rhamnopyranosyl-(1→2)-β-D-glucopyranose). The sugars can be further acylated, either by aliphatic (e.g., malonyl or acetyl residues) or aromatic acids (e.g., *p*-coumaroyl, caffeoyl, or feruloyl residues).

The structural diversity of flavonoids affects their physicochemical behavior, and different flavonoid classes and compounds may have different requisites for their extraction and analysis, so that there is not a unique analytical strategy that applies in all the situations. The usual flow in flavonoid analysis involves sample preparation, extraction and cleanup, chromatographic separation, and detection by spectroscopic techniques. At each of these steps, different approaches can be considered depending on the nature of the plant matrix, type of flavonoids, and intended outcome (e.g., compound isolation, structural characterization, metabolite profiling, or compound quantification). Solvent extraction is commonly used for the preparation of the crude extracts. Solid-phase extraction (SPE) and column chromatography (CC) are the more usual techniques applied for cleanup and/or compound fractionation, whereas high performance liquid chromatography (HPLC) is the technique of choice for flavonoid separation, usually using reversed phases (RPs) and namely C_{18} alkyl-bonded resins. Flavonoid identification and structural characterization involve the use of spectroscopic techniques, such as spectrophotometry, mass spectrometry (MS), and nuclear magnetic

resonance (NMR). The coupling (hyphenation) of these techniques to HPLC has become a fundamental approach for the identification and/or confirmation of the identity of target and unknown compounds, without requiring previous compound isolation or even high purity of sample.

It is not our intention to review all possible techniques and methodological approaches to the analysis of flavonoids, which have already been the object of systematic treatment in different reference publications (Harborne, Mabry, and Mabry, 1975; Harborne and Mabry, 1982; Markham, 1982; Harborne, 1988, 1994; Santos-Buelga and Williamson, 2003; Andersen and Markham, 2006). In this chapter, particular attention will be paid to
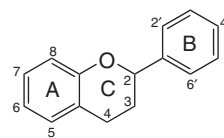
Figure 1 Basic structure of flavonoids.

Table 1 Main classes of flavonoids.

Flavonoid class	Core structure	Examples
Flavones		Apigenin, luteolin, chrysin, scutellarein, diosmetin, chrysoeriol
Flavonols		Quercetin, kaempferol, myricetin, galangin, fisetin, morin
Isoflavones		Genistein, daidzein, glycitein, formononetin, biochanin A, puerarin
Flavanones		Hesperidin, naringenin, taxifolin, eriodictyol, isosakuranetin
Chalcones		Phloretin, arbutin, butein, naringenin chalcone
Flavan-3-ols		(Epi)catechin, (epi)gallocatechin, condensed proanthocyanidins
Anthocyanins		Cyanidin, delphinidin, malvidin, pelargonidin, petunidin, peonidin

more recent extraction techniques and HPLC-based methodologies that currently dominate the field of flavonoid analysis.

2 SAMPLE PREPARATION

Sample collection and preparation are critical steps within the analytical process and have to be carefully performed in order to ensure that quality and representative results are obtained. It has been estimated that sample preparation, including cleanup and extraction, can represent up to 60% of a laboratory technicians timetable and forms one of the principal sources of error (Luthria, 2006). In flavonoid analysis, storage conditions have to be considered carefully, so as to prevent compound degradation and/or enzymatic reactions that may lead to structural modifications. Fresh samples can be kept under refrigeration, although they are unstable and a rapid decrease in the content of compounds usually occurs (Vallejo, Tomas-Barberan, and García-Viguera, 2003). Sample blanching and/or storage in an inert atmosphere may help to improve compound stability, although freezing and drying are the preferred options for long storage. Freezing of the sample facilitates further extraction, because the increase of sample volume and the formation of ice crystals lead to tissue disruption facilitating the release of compounds. Enzymatic and chemical reactions may, however, take place during thawing of the samples leading to changes in the flavonoid composition. Microwave thawing usually induces lesser changes than refrigerator or room temperature thawing (Robards, 2003). Freeze drying is considered less detrimental for flavonoid stability than air drying or vacuum drying, whereas oven drying may induce degradation of thermally unstable compounds (Hung and Duy, 2012).

Flavonoids can exist in plant materials as free aglycones or as glycosylated forms and are frequently associated with other cellular components such as cell walls, carbohydrates, or proteins. Although the usual interest is in the intact conjugates, in some cases, the sample can be submitted to a previous hydrolysis-digestion step to release the aglycones. This strategy sensibly reduces the number of flavonoid structures, facilitating compound extraction and increasing the possibilities of detection and quantification.

Within minor changes, the hydrolysis protocol originally described by Harborne (1965) consisting of a treatment with 2 M HCl in boiling 50% aqueous methanol is usually used for the cleavage of O-glycosylated phenolics. Nevertheless, the optimal hydrolysis conditions should be determined for each plant material, and the use of different methods might even be necessary for the same plant material when the accurate hydrolysis of different compounds is required (Nuutila, Kammiovirta, and Oksman-Caldentey, 2002). Mild hydrolysis conditions using organic acids, such as 10% of acetic acid, might offer some structural information, based on the identification of intermediate products from the partial cleavage of the conjugation moieties (Santos-Buelga et al., 2011).

Alkaline hydrolysis is recommended for the release of wall-bound flavonoids. This process is achieved with NaOH or KOH 2–10 N using incubation times from a few minutes up to 6 h, sometimes under nitrogen and protecting from light. Alkaline hydrolysis can also be used analytically to assist in the identification of acylated flavonoids; ester linkages are cleaved, and the released products (acyl residues and base flavonoid) can be further identified by HPLC or other suitable techniques. However, some compounds are unstable during alkaline hydrolysis, such as those containing o-diphenols that can be oxidized to the corresponding quinones (Robards, 2003). Another example of instability is flavanones that are easily converted to isomeric chalcones in alkaline media provided that there is a hydroxyl substituent at position 2′ or 6′ of the chalcone (Tomás-Barberán et al., 2003). The use of inert atmospheres and addition of antioxidants such as ascorbic acid or EDTA can prevent or at less moderate these losses or transformations (Nardini et al., 2002).

Although in less extension, enzymatic cleavage has also been used to release flavonoid aglycones and/or to improve the yield of flavonoid extraction from plant materials. Kuhnle et al. (2007) used a combination of β-glucuronidase, cellulase, and β-glucosidase to facilitate the extraction of isoflavones from different fruits and vegetables. Treatment with pectinase or pectinase/cellulase combinations has been used to enhance the extraction of flavones from celery (Zhang et al., 2011) and anthocyanins from grape skins (Maier et al., 2008). Chen et al. (2011), to improve the extraction of flavonoids from Ginkgo biloba L., used a bifunctional cellulase from Penicillium decumbens that, besides activity for cellulose

degradation, shows transglycosylation activity, thus facilitating both cell wall degradation and transformation of aglycones into more polar glycosides with higher solubility in the ethanol–water extractant. A propolis β-glycosidase was used by Zhang, Liu, and Hu (2012) to induce the hydrolysis of glycosides previous to the HPLC analysis of flavonoid aglycones.

3 EXTRACTION AND CLEANUP

3.1 Solvent Extraction

Conventional solid–liquid extraction is still the most common procedure for the extraction of flavonoids from plants. However, because the solubility of flavonoids varies significantly owing to their conjugation status and their association with the sample matrix, it is not possible to recommend a unique method for extraction of all classes of flavonoids with a single solvent system. Methanol or ethanol has been generally found to be efficient in extraction of lower molecular weight polyphenols or flavonoid glycosides, although, to increase the yield of the extraction, frequently, mixtures with water are used. Aglycones or less polar flavonoids (e.g., isoflavones, flavanones, or flavonols) can be extracted with organic solvents such as chloroform, dichloromethane, ethyl acetate, or diethyl ether. When the plant material is rich in flavonoids with different polarity, it is frequent to carry out a sequential solvent extraction combining a first step using less polar solvent with a subsequent extraction with an alcohol or alcohol–water mixture (Escarpa and Gonzalez, 2001). Acidification of the solvent may increase the ability to extract some flavonoids, especially when protic polar solvents, such as methanol or ethanol, are used. By acidifying the medium, the phenol–phenolate equilibrium shifts toward the less polar phenyl form, thus facilitating extraction with organic solvents. Acidification is even necessary for the extraction of anthocyanins that are structurally dependent on the pH of the medium, which modifies their characteristics of solubility and affects their stability. Soft acidic conditions must be used to prevent hydrolysis of flavonoids containing labile conjugated residues during extraction and subsequent concentration steps. The use of weak organic acids, such as formic or acetic acids, instead of inorganic acids for solvent acidification and addition of water before concentration might minimize these losses (Gonzalez-Paramás et al., 2011).

The efficiency of the extraction process is known to be a function of several conditions such as temperature, liquid–solid ratio, and particle size (see **Extraction Methodologies: General Introduction** for further details). Although high temperatures improve the efficiency of any extraction process, excessive temperature may degrade labile flavonoids; therefore, temperatures under 25°C are recommended (Escribano-Bailón and Santos-Buelga, 2003). It is also habitual working in inert atmosphere and with addition of an antioxidant (e.g., BHT, sulfites, and *tert*-butylhydroquinone) to prevent phenolic oxidation, although the use of ascorbic acid is not advised as it is recognized to cause anthocyanin degradation and may also induce the browning of flavan-3-ols (Santos-Buelga et al., 2012).

Quite usually, solvent extraction is facilitated by the use of ultrasounds or microwaves. Ultrasound-assisted extraction (UAE) constitutes, together with Soxhlet, the most accepted conventional extraction technique. Compared with other extraction techniques such as microwave-assisted extraction (MAE) or supercritical fluid extraction (SFE), UAE is an inexpensive and efficient alternative, and its operation is much easier. It requires simple equipment and dramatically reduces extraction times and solvent consumption. It also enables operation at ambient temperature and pressure, which is particularly favorable for extraction of thermally unstable components. Normal operation conditions comprise frequencies ranging 20–60 Hz, sonication powers around 90–150 W, and times of 2–30 min for one to five cycles of extraction, although these variables should be optimized for the specific type of compounds and matrix. The extraction efficiency can be enhanced by increasing the temperature, although values higher than 40°C may lead to thermal degradation of flavonoids, in addition to high power and/or large sonication times (Ma et al., 2009). Examples about applications of the UAE can be found in recent reviews (Dai and Mumper, 2010; Santos-Buelga et al., 2012; Wijngaard et al., 2012).

MAE is based on the absorption of microwave energy by polar molecules, such as water, which generates heat and results in higher extraction rates, with a significant reduction in the extraction time (quite often less than 10 min are sufficient) and solvent consumption, compared with conventional solvent

extraction techniques (see **New Trends in Extraction of Natural Products: Microwave-Assisted Extraction and Pressurized Liquid Extraction** in this handbook for further information). High temperatures produced in MAE not only enhance the extraction rates but also increase the risk of degradation of thermolabile flavonoids, especially catechins, flavonols, and compounds with greater number of hydroxyl substituents (Liazid *et al.*, 2007; Biesaga, 2011). When selecting a solvent for MAE, consideration should be given to its microwave-absorbing properties. In general, the higher the dielectric constant, the higher is the capacity of the solvent to absorb microwave energy, which can lead to faster rate of heating of the solvent with respect to the plant material (Routray and Orsat, 2012). Acetone and methanol are the most common solvents for the extraction of flavonoids, usually with some percentage of water as polar modifier. When nonpolar solvents such as hexane or toluene are used, a certain percentage of a polar solvent, most commonly water (e.g., 10%), is normally required to improve the heating rate, although approaches using mixtures of two organic solvents (e.g., mixtures of hexane–acetone) can also be used depending on the target analytes (Eskilsson and Bjorklund, 2000). The use of ionic liquids, such as 1-butyl-3-methylimidazolium bromide, has also been explored as a greener alternative to organic solvents in MAE (Du, Xiao, and Luo, 2009). MAE methods have been largely applied to the extraction of flavonoids from many different plant materials using a variety of solvents and operation conditions, as recently reviewed by Routray and Orsat (2012).

Crude extracts from solvent extraction often contain unwanted substances, such as other plant phenolics, sugars, fats, terpenes, waxes, or pigments, which can interfere with flavonoid analysis. Liquid–liquid extraction using nonpolar solvents (petroleum ether, chloroform, and hexane) is a common strategy for removal of lipohilic compounds (Santos-Buelga *et al.*, 2012), in addition to a cleanup step, quite usually by SPE.

3.2 Solid-Phase Extraction

SPE is the more usual technique applied not only for sample cleanup but also for fractionation and/or concentration of compounds from plant extracts. Chemically bonded silica, usually with a C_{18} organic group, is by far the most commonly used SPE sorbent in flavonoid analysis. Pretreatment of the SPE material with an activating solvent (such as methanol, acetone, or acetonitrile) must be used to obtain better surface contact with the aqueous solution being extracted, and sample solution and solvents are usually slightly acidified to prevent ionization of the flavonoids (de Rijke *et al.*, 2006). Silica sorbents present, however, some disadvantages, such as low recovery in extracting polar compounds, instability at extreme pH, and presence of some residual silanol groups (Poole, 2003). Porous polymers overcome some of these problems as they are stable throughout a greater pH range and the compounds adsorbed are easily eluted. The most widely used polymeric sorbent is macroporous styrene-divinylbenzene (St-DVB). StrataTMX, a surface-modified St-DVB polymer from Phenomenex, and Oasis®HLB, a macroporous poly(*N*-vinylpyrrolidonedivinylbenzene) from Waters, are two examples of widely used RP sorbents that may suit both the parent aglycones and more polar compounds (Santos-Buelga *et al.*, 2012). Polyamide-based polymers are another type of hydrophilic sorbents that are recommended for extraction of highly polar flavonoids (Waksmundzka-Hajnos *et al.*, 2008). Other SPE materials are mixed-mode cartridges that incorporate RP sorbents and ion-exchange resins, such as Oasis®MCX (cation-exchange) and Oasis®MAX (anion-exchange); the combination of strong retention with the use of efficient rinse solvents result in cleaner extracts compared with single-mode sorbent (Poole, 2003).

Different alternatives to classic SPE have arisen to improve efficiency, selectivity, and/or sensitivity, such as matrix solid-phase dispersion (MSPD), microextraction by packed sorbent (MEPS), or molecular imprinting technologies. MSPD consists of the blending of a solid support material with a solid or semisolid sample into a column to achieve extraction and sample cleanup into a single step. The principal difference with SPE is that, in MSPD, the sample is not absorbed onto the top of the column packing material but dispersed throughout the column (Barker, 2007). Different MSPD packing materials have been used for selective extraction of flavonoids from plant samples, from classic C_{18} sorbents (Visnevschi-Necrasov *et al.*, 2009) to the novel Titania microspheres (Xu *et al.*, 2011). In MEPS, the packing material is inserted into a syringe

requiring minimum solvent volumes to eluting the analytes that are injected directly into an LC or GC system. The commercially available MEPS uses the same sorbents as conventional SPE columns and, so, is suitable to use with most existing methods by scaling the reagent and sample volumes (Silva, Haesen, and Camara, 2012). Although this technique is extensively reported in bibliography for the analysis of different drugs in biological samples, fewer applications have been described for the extraction of flavonoids from plant samples (Silva, Haesen, and Camara, 2012).

Molecular imprinting technology uses a target analyte as a template molecule to produce molecularly imprinted polymers (MIP) or molecularly imprinted membranes (MIM), which exhibit a high affinity toward the compound, so that they are expected to selectively adsorb it (and/or a group of structural analogs) in preference to molecules with other shape or size, although nonspecific adsorption of other compounds might also be produced (e.g., by hydrophobic interactions). There are many examples in the bibliography concerning the use of MIPs for extraction of particular flavonoids. Quercetin or derivatives (e.g., rutin) have been used as template molecules to obtain MIPs in view to their extraction from different samples (Xie *et al*., 2001; Weiss *et al*., 2002; O'Mahony *et al*., 2006; Xia *et al*., 2006). Recently, an MIP showing good binding affinity toward quercetin, catechin, and epigallocatechin gallate has been prepared using quercetin as template and 4-vinylpyridine as functional monomer (Castro-López *et al*., 2012). There are also examples in the literature describing the preparation of different types of MIPs for selective extraction of catechins, which, in general, showed high molecular recognition ability for the respective template molecule (e.g., epicatechin, epigallocatechin, epicatechin gallate, or epigallocatechin gallate), although the competitive adsorption of other catechins was usually observed (Blahova, Lehotay, and Skacani, 2004; Ding *et al*., 2006; Haginaka *et al*., 2007). Other molecularly imprinted materials developed for the extraction of flavonoids have used template molecules such as the flavone luteoline for the preparation molecularly imprinted silica microspheres (Zhang, Qin, and Tu, 2007), or the flavanone naringin using β-cyclodextrine (Ma *et al*., 2011) or acrylamide (Tasselli, Donato, and Drioli, 2008) as functional monomers.

3.3 Pressurized Solvent Extraction

SFE and pressurized liquid extraction (PLE) are the most popular techniques based on pressurized fluids (see also chapters **Supercritical Fluid Extraction** and **New Trends in Extraction of Natural Products: Microwave-Assisted Extraction and Pressurized Liquid Extraction** in this handbook). Applications of the SFE to the extraction of bioactive compounds including flavonoids from different plant-derived sources have been recently reviewed (Pereira and Meireles, 2010; Wijngaard *et al*., 2012). Carbon dioxide (CO_2) is the supercritical fluid most commonly used, although other solvents such as propane, ethane, hexane, pentane, and butane have also been explored for polyphenol extraction (Pereira and Meireles, 2010). CO_2 is the solvent of choice to extract nonpolar to low polarity compounds; however, highly polar flavonoids are not extracted by 100% CO_2, and certain percentages of solvent modifiers have to be added in order to improve the extraction efficiency. Methanol is a usual cosolvent, although its toxicity limits its applications in view to further use of the compounds/extracts in food and drug industries; ethanol has also been broadly used as a cosolvent in the extraction of polyphenols owing to its low toxicity compared with other options (Castro-Vargas *et al*., 2010; Liza *et al*., 2012).

Comparison of classical and SFE procedures yields contradictory results. In general, in the case of the nonpolar compounds, such as flavonoid aglycones, isolation by SFE provided identical or somewhat better yields. However, in the case of water-soluble glycosides or polar analytes, yields remain low, even when high pressures and more than 20% of polar solvent modifiers are used (Pinelo *et al*., 2007). Polar analytes require either long extraction times or unusual flow rates of extraction fluid to be extracted quantitatively in a reasonable time; thus, some of the benefits of classical SFE are lost: solvent consumption is higher, the method is less environmentally friendly, and, typically, the extracts must be preconcentrated before the analyses (Klejdus *et al*., 2005). Selection of the most sustainable extraction technique for plant phytochemicals depends on the type of compound to be extracted. In general, supercritical (SC)-CO_2 is adequate for apolar compounds, whereas PLE is more suited for more polar compounds, such as polyphenols (Wijngaard *et al*., 2012). Furthermore, SC-CO_2 allows the extraction of thermolabile compounds at low temperature in an

oxygen-free environment, whereas PLE is less suitable for heat-sensitive compounds (Sticher, 2008).

PLE, also known as *accelerated solvent extraction*, subcritical-fluid extraction, or high pressure and temperature extraction, is a solid–liquid extraction process, partly derived from SFE, in which extractions are carried out under elevated pressure (4–20 MPa) in order to maintain the solvent in its liquid state, even at temperatures above boiling point.

A large number of PLE applications have been reported for the extraction of different flavonoid families from diverse plant materials. Some recent examples are summarized in Table 2.

Water and hydroalcoholic mixtures are the most usual liquids used in PLE of flavonoids (Mustafa and Turner, 2011). The use of water as a PLE solvent is known as *subcritical water extraction* (*SWE*, *pressurized hot water extraction*, and *pressurized low polarity water extraction*). Water is normally used at a temperature above 100°C but below its critical value (374°C) and at a pressure sufficiently high (>40 bar, 4 MPa) to maintain the liquid state. Although the extraction efficiency of water is considered lower than that of less polar organic solvents, it must be taken into account that water polarity markedly decreases at high temperatures and under pressure, thus increasing the range of analytes that can be extracted. Extraction with acidified subcritical water was shown to be as efficient as acidified 60% methanol in extracting anthocyanins from grape skins (Ju and Howard, 2003). SWE has been indicated as a suitable alternative for extracting some flavonoid groups, such as less polar flavonols. Yields obtained by SWE in the extraction of quercetin from onion skin were over eight-, six-, and four-fold greater than those using ethanol, methanol, and water-at-boiling-point extraction methods, respectively (Ko *et al*., 2011).

4 ANALYSIS OF TOTAL FLAVONOIDS

UV or UV–vis spectroscopy has for years been used for quantification of total phenolics and flavonoids. The method that uses the Folin-Ciocalteu (FC) reagent (a mixture of tungsten and molybdenum oxides) is the most frequently used for the determination of total polyphenols. It is based on a chemical reduction of the reagent by phenolic-rich compounds, including flavonoids, to give a blue product with maximum absorption at 765 nm. However, the method is not specific as other reducing substances present in the sample also react with the FC reagent, thus usually leading to an overestimation of the phenolic content. The most problematic interference may well be sugars, but also components such as proteins, aromatic amines, sulfur dioxide, ascorbic acid, and other enediols and reductones, organic acids, or Fe(II) (Prior, Wu, and Schaich, 2005). Methods based on the reduction of Fe(III) in the Prussian blue reaction have also been used for determination of total phenolics (Graham, 1992), although they have similar principle as the FC reaction and, therefore, show similar drawbacks.

Other spectrophotometric methods take advantage of the structural characteristics of some flavonoids. Colorimetric measurement at 430 nm after reaction with Al(III) have been used for quantification of total flavonoids, although it does not determine equally all flavonoid groups but rather more specifically flavonols, owing to their ability to form acid-stable complexes between the keto group in position 4 and neighboring hydroxyls (Markham, 1982). Flavan-3-ols are able to react with aromatic aldehydes in strongly acidic media to form colored adducts; the reagents usually used are vanillin (Broadhurst and Jones, 1978) and 4-dimethylaminocinnamaldehyde (DMACA) (Delcour and Varebeke, 1985). The

Table 2 Examples of extraction of flavonoids using pressurized liquid solvents.

Target flavonoids	Source	Extraction conditions	References
Anthocyanins	Purple-fleshed potato	1500 psi, 80–100°C, 75% methanol in acidified water	Truong *et al*. (2012)
	Grape pomace	6.8 MPa, 100°C, 50% ethanol	Monrad *et al*. (2010a)
Flavonols	Apple pomace	10.3 MPa, 102°C, 5 min, 60% ethanol	Wijngaard and Brunton (2009)
	Onion	1500 psi, 40°C, 60% aqueous methanol	Søltoft *et al*. (2009)
	Onion	9–13.1 MPa, 160°C, water	Ko *et al*. (2011)
Flavones	Olive oil by-products	200°C, water	Herrero *et al*. (2011)
	Aromatic plants	200°C, water	Miron *et al*. (2011)
Procyanidins	Grape pomace	6.8 MPa, 140°C, 50% ethanol	Monrad *et al*. (2010b)

reaction takes place with both catechins and PAs, although the reactivity of the compounds varies depending on their structure and degree of polymerization (DP), which makes it difficult to compare the results obtained with samples of different composition. Methods based on the depolymerization of PAs by heating in acidic medium and colorimetric determination of the anthocyanidins released can be used for the analysis of total condensed tannins (Porter, Hrstich, and Chan, 1986). The yield of the reaction depends on the conditions used and on the structure of the PAs present in the sample, and it is usually low owing to the formation of polymeric side products (phlobaphenes).

Method based on the acidification to highly acidic values (pH < 1) are used for selective determination of anthocyanins; at these pH values, all structural forms of the anthocyanins shift to the red-colored flavylium cation that can be measured at 520–540 nm (Ribereau-Gayon and Stonestreet, 1965).

Spectroscopic measurements in the near infrared (NIR) wavelength region of the electromagnetic spectrum have become one of the most attractive and used technique in the food industry to evaluate the composition of foods due to the ease of use in routine operations and minimal sample preparation (see **Near Infrared Spectroscopy (NIR) in Natural Product Research** for further information). Several authors have reported the feasibility of NIR spectroscopy to determine total phenolics or specific flavonoid classes such as anthocyanins, flavonols, or PAs, using diverse regression models for the calibration procedure. Thus, NIR spectral datasets were successfully correlated with the contents of total phenols, total flavonoids, and total anthocyanins in blueberries using partial-least squares (PLS) regression (Sinelli *et al.*, 2008). In addition, good calibration statistics were obtained for the prediction of total procyanidin oligomers in cocoa by applying modified partial least squares (mPLS) algorithms (Whitacre *et al.*, 2003). Synergy interval partial-least squares (siPLS) revealed superiority compared with PLS for the analysis of total polyphenols in green tea (Chen *et al.*, 2008) or total flavonoids in *G. biloba* L. leaf (Zhou *et al.*, 2007; Shi *et al.*, 2012). However, NIR spectroscopy models (PCA, PLS, and mPLS) were not successful to estimate flavonoid content in rice grains, although they could effectively predict total phenolic contents and antioxidant capacity (Zhang *et al.*, 2008a). PLS regression has also been used for the rapid quantification of individual

compounds such as hyperforin in *Hypericum perforatum* L. (Rager *et al.*, 2002), epigallocatechine gallate in green tea (Luypaert, Zhang, and Massart, 2003), or flavonol aglycones (quercetin, kaempferol, and isorhamnetin) in *G. biloba* L. leaf (Geng and Xiang, 2008). Application of principal component analysis to the NIR spectral data was also useful to predict the content of the dihydrochalcone aspalathin in unfermented rooibos and to discriminate between fermented and unfermented rooibos (Schulz, Joubert, and Schütze, 2003). NIR technology has also been used for the discrimination of *Scutellaria baicalensis* Georgi (Li *et al.*, 2011b) or propolis samples (Cai *et al.*, 2012) according to the geographic origin based on flavonoid quantification.

5 SEPARATION AND DETECTION

5.1 High Performance Liquid Chromatography

HPLC was introduced for flavonoid analysis in the mid-1970s (Adamovich and Stermitz, 1976; Wulf and Nagel, 1976), and since then, it has become the technique of choice for the individual analysis, both qualitative and quantitative, of all classes of flavonoids. There is not a universal HPLC strategy that can be applied for flavonoid separation in every situation, and, in general, laboratories tend to develop their own methods according to their particular analytical needs. In this section, only a brief outline of the more general patterns used is offered, once detailed information on the principles that govern the HPLC separation of flavonoids, in addition to practical examples for different flavonoid groups, can already be found in many reviews and book chapters (Daigle and Conkerton, 1983, 1988; Merken and Beecher, 2000, 2002; Santos-Buelga, Garcia-Viguera, and Tomas-Barberan, 2003; de Rijke *et al.*, 2006; Marston and Hostettmann, 2006).

RP-HPLC is the predominant technique for the chromatographic separation of flavonoids. In RP mode, the stationary phase is less polar than the mobile phase, and thus compounds are eluted in order of decreasing polarities. Typically, for a given group of flavonoids, glycosides will be eluted before their aglycones, and among these, those possessing more hydroxyl groups will be eluted before those with less. Normal phase (NP) chromatography is much less common than RP-HPLC for flavonoid analysis and

has been usually restricted to separations of nonpolar or weakly polar aglycones, such as polymethoxylated flavones, flavanones, or isoflavones (Marston and Hostettmann, 2006), in addition to the separation of flavan-3-ols (catechins and PAs) according to their DP (Rigaud *et al*., 1993; Hammerstone *et al*., 1999).

C18-bonded silica phases are the most commonly used for flavonoid separations in RP-HPLC. Columns ranging from 100 to 250 mm in length, with internal diameter 3.9–4.6 mm, and particle sizes of 3–5 μm have been classically used. End-capped columns with reduced proportion of residual silanol groups on the silica surface usually improve peak symmetry and provide better chromatographic resolution, especially of the most polar analytes that are not sufficiently retained on conventional RP columns (Marston and Hostettmann, 2006). Separations are usually performed at room temperature, but thermostated columns give more repeatable elution times, and moderately higher temperatures up to 40°C may be recommended to reduce the time of analysis.

Both isocratic and gradient elution can be applied for the separation of flavonoids depending on the number and type of the analytes and the nature of the matrix. In RP-HPLC, gradient elution is generally performed using binary solvent systems between an aqueous polar solvent and a less polar organic solvent, more often acetonitrile or methanol. Acetonitrile normally leads to better resolution in a shorter analysis time than methanol and, generally, gives sharper peak shapes, resulting in a higher plate number. Acetonitrile is also preferred to methanol when electrochemical detection (ECD) is used, as it produces lower baseline noise (Rehova, Skerikova, and Jandera, 2004). By contrast, methanol is preferable to acetonitrile for (semi)preparative purposes, due to its higher volatility. Occasionally, tetrahydrofuran, isopropanol, or *n*-propanol has also been used as less polar solvents (Marston and Hostettmann, 2006). Flavonoids contain ionizable hydroxyl groups, and the use on an acid modifier in the solvent system is important to suppress ionization and prevent the interactions of these groups with residual traces of metals in the stationary phase. Keeping a low pH minimizes peak tailing and improves the resolution and reproducibility of the retention characteristics. Aqueous acidified solvents such as acetic, formic, and trifluoroacetic acids and ammonium acetate, ammonium formate, and ammonium chloride buffers at low pH are among the most used for flavonoid separation; phosphate buffers have become less popular due to the dreaded contamination of the ion sources in MS detection (de Rijke *et al*., 2006). Typical flow rates for analytical separation are in the range 0.8–1.5 mL/min, although there has been a tendency to decrease them in order to save solvents and permit the coupling to mass or NMR detectors. This has been facilitated by the advances in packing materials (e.g., monolithic or superficially porous stationary phases) and the use of equipments providing ultrahigh pressures [ultrahigh pressure liquid chromatography (UHPLC)], which allow reducing the time of the HPLC runs and the amount of sample injected without compromising and even increasing peak resolution and sensitivity. Significant improvements in separation and selectivity have also been provided by the introduction of newer technical developments such as high temperature HPLC or multidimensional liquid chromatography (MDLC).

5.2 Recent Developments in HPLC

5.2.1 Improvements in Speed

5.2.1.1 Ultrahigh Pressure Liquid Chromatography

UHPLC has emerged in recent years as a suitable alternative to conventional HPLC. This novel approach offers higher separation efficiency through the use of phases with small particle size (<2 μm) and diameter (typically 2.1–1 mm), which dramatically increases the number of theoretical plates and provides faster separations in the order of three to nine times compared with conventional HPLC. Although conventional HPLC equipments may allow the use of shorter sub-2 μm columns (under 50 mm), owing to the generated backpressure, optimized instrumentation capable of operating at high pressures above 500 bar is required for longer columns and/or demanding separations (Majors, 2012). An interesting advantage of UHPLC over other strategies is the theoretical possibility of transferring existing methods from conventional HPLC using basic calculations as for the scaling up from analytical to preparative chromatography (Guillarme *et al*., 2010).

Over the past 8 years, more than 100 articles have been published dealing with the application of UHPLC to both screening and profiling of flavonoids

in different sample matrices. In most cases, MS is the technique used for detection combined or not with photodiode array detection. Significant achievements were obtained by Hanhineva *et al.* (2008) in the nontargeted metabolomic analysis of secondary metabolites in flower organs of strawberry (*Fragaria x ananassa*) using a UHPLC-PDA-qTOF/MS method; up to 109 phytochemicals, including flavonoids, were separated in less than 25 min using a Waters Acquity BEH C18 column 1.7 µm (100 × 2.1 mm). More recently, Lin *et al.* (2011) used a UHPLC-PDA-electrospray ionization (ESI)/HRMS/MS method for profiling phenolic compounds in red mustard greens (*Brassica juncea* L.) using for the separation a Hypersil Gold AQ RP-C18, 1.9 µm (200 × 2.1 mm), which allowed them the identification of 67 anthocyanins, 102 flavonol glycosides, and 40 hydroxycinnamic acid derivatives in a running time of 60 min. In a recent article, Vrhovsek *et al.* (2012) applied a UHPLC-QqQMS/MS method for the targeted analysis of phenolic compounds in fruits. The method used a Waters Acquity HSS T3 column 1.8 µm (100 × 2.1 mm) and was optimized to achieve the separation of 135 phenolic metabolites, including more than 60 flavonoids over a period of 15 min. More information about applications of UHPLC to the analysis of flavonoids and other related phenolics can be found in the reviews by Eugster *et al.* (2011), Kalili and de Villiers (2011), or Qiao *et al.* (2011).

5.2.1.2   Monoliths

Monolithic HPLC columns consist of a single piece of a continuous, porous material that is hermetically sealed against the wall of a tube, so that the stream of mobile phase cannot bypass any significant length of the bed but must percolate through it. Silica monolithic columns usually possess greater surface area than polymer monoliths and are more suitable for separation of small molecules such as flavonoids, whereas polymers have been more applied to separations of larger molecules (e.g., peptides, proteins, or nucleic acids). Monolithic columns require higher flow rates to achieve separations comparable to particle-packed columns of similar length, which induces more solvent consumption, and it is not always compatible with MS detection.

There are very few articles reporting the use of monolithic columns for the chromatographic separation of plant flavonoids. Glycosides from six different flavonoid aglycones (apigenin, chrysoeriol,

isorhamnetin, kaempferol, luteolin, and quercetin) were separated and identified by Vukics *et al.* (2006) in extracts of *Viola tricolor* L. using a monolithic poly(*p*-methylstyrene-*co*-1,2-bis(*p*-vinylphenyl) ethane) capillary column (260 × 0.2 mm) and tandem MS detection. Repollés, Herrero-Martiinez, and Raflos (2006) used a Chromolith Performance C18 endcapped monolithic column (100 × 4.6 mm) and diode array detection for the analysis of up to 11 flavonoid aglycones in a running time less than 15 min from orange peel, *G. biloba* L., and propolis samples. A similar approach was used for the analysis of flavonol aglycones in tomato samples (Biesaga and Pyrzynska, 2009) and *Rhus coriaria* L. (Mehrdad *et al.*, 2009). Two C18 monolithic columns (100 × 4.6 mm) linked in series were used by de Sousa *et al.* (2012) for the separation of flavonols in extracts of the leaves of *Copaifera langsdorffii* Desf.

5.2.1.3   Superficially Porous Particle-Packed Columns

Superficially porous stationary phases (SPP), also referred to as *core-shell* and *fused-core particle phases*, consist of a solid fused core and a porous particle outer layer. These phases offer lower mass transfer resistance and higher permeability than totally porous particles, and their homogeneous packed bed minimizes the diffusion of the mobile phase and solute, thus decreasing the degree of band spreading. SPP columns show similar advantages as sub-2 µm particle columns in terms of resolution and speed, although they generate lower backpressure at the same flow rate, which allows their use not only in UHPLC equipments but also with conventional HPLC instrumentation. Until now, few applications of these columns to flavonoid separation have been reported. A 2.7-µm Ascentis Express C18 column was used by Dugo *et al.* (2009) for the separation of phenolic compounds (chlorogenic acids and flavonols) from the dried leaves of *Ilex paraguayensis* Hook. The same column was used for the separation of phenolic acids and flavonoids in *Thymus* L. species (Boros *et al.*, 2010) and to separate flavone and flavonol aglycones in hydrolyzed extracts of plant samples (Olszewska, 2012). Manchon *et al.* (2010) used a 2.6-µm Kinetex C18 SPP column for the separation of up to 12 standard isoflavones in a time as short as 6 min.

As SPP columns can provide fast and efficient separations on conventional HPLC instrumentation, they have been used for RP separation of phenolic

compounds in the second dimension of comprehensive chromatographic (LC × LC) systems (Section 5.2.2.2).

5.2.1.4   High Temperature Liquid Chromatography

High temperature liquid chromatography (HTLC) combines the use of standard columns with temperatures higher than 60°C; in these conditions, the viscosity of the solvents is decreased, reducing the backpressure and allowing an increase in the velocity of phase mobile, which results in faster and more efficient separations. The development of this technique has been constrained by the instrumentation requirements and the reduced availability of stable high temperature-resistant packing materials. Concerns also exist regarding the potential degradation of thermally unstable compounds (Novakova and Vlckova, 2009). The thermal stability of four hydrosilated silica-based columns and the effects of the temperature on the elution behavior of different phenolic standards were studied by Soukup and Jandera (2012). Among the stationary phases tested, Cogent UDC cholesterol™ column (modified with cholesterol as hydrophobic ligand) showed the highest temperature stability (up to 100°C) and provided the most selective and efficient separations of phenolic standards both in aqueous NP [hydrophilic interaction chromatography (HILIC)] and in RP modes with almost reversed elution order. As expected, the retention times and peak widths decrease at increasing temperature. HTLC using a polymer-based RP column operated at 120°C was successfully used by Reichelt *et al.* (2010) for the fractionation at semipreparative scale of benzoic acid derivatives and flavonoids from *Eriodictyon* spp.

5.2.2   Improvements in Resolution

5.2.2.1   Hydrophilic Interaction Chromatography

HILIC has been associated to NP-LC, although it involves more complicated separation mechanisms. Both modes use polar stationary phases, but unlike NP-LC, where nonpolar solvents are used and the water content in the mobile phase is usually kept at minimal levels, HILIC typically uses water-miscible polar organic solvents such as acetonitrile or methanol, and a water content higher than 5% is crucial for the retention mechanism. In contrast to RP-LC, gradient elution in HILIC begins with a low polarity organic solvent and elutes polar analytes by increasing the polar aqueous content. HILIC represents an alternative mode for separating polar compounds that are weakly retained in common RP-HPLC and elute near the void, and it also overcomes the drawbacks of the poor solubility often encountered in NP-LC. Although some companies have marketed columns specific for HILIC, most columns used with NP-LC such as pure silica columns, amino columns, or cyano (CN) columns can be operated under HILIC conditions. The highly volatile organic mobile phase used in HILIC can also provide increased sensitivity with ESI-MS. In addition, the organic extracts from protein precipitation, liquid–liquid extraction, or SPE are compatible with the higher organic content in HILIC mobile phases and can often be directly injected onto a HILIC column (Jian *et al.*, 2010). There are not many examples of the application of HILIC to the analysis of flavonoids. Yanagida *et al.* (2007) developed a method for the separation of oligomeric PAs from apple extracts according to the DP (up to decamer) by HILIC using an amide-silica column and aqueous acetonitrile as mobile phase. The orthogonality of HILIC and RP-HPLC has led some authors to integrate both approaches into two-dimensional HPLC setups for the separation of complex phenolic and flavonoid mixtures (Section 5.2.2.2). Zhang *et al.* (2008b) used HILIC with preparative purposes for the isolation of flavonoids from licorice and kudzu fractions obtained by RP-HPLC. Similarly, Feng *et al.* (2010) separated flavonoids from Herba *Hedyotis diffusae* L. using a Click β-cyclodextrin column successively operated under RP-HPLC and HILIC conditions. More information about HILIC characteristics and applications can be found in a recent review by Buszewski and Noga (2012).

5.2.2.2   Multidimensional Liquid Chromatography

In recent times, MDLC systems have been implemented for the separation of samples containing a large number of closely eluting compounds and/or within a range of polarities that are difficult to resolve by means of a conventional HPLC technique. Three types of MDLC arrangements can be distinguished: (i) bidimensional LC without instrumental coupling (offline 2D-LC), in which fractions collected from the

first chromatographic separation are manually transferred to the second separation system; (ii) online or "heart-cutting" 2D-LC, where only selected fractions are automatically transferred from the first to the second dimension; and (iii) comprehensive LC (LC × LC) techniques, in which all sample compounds are automatically and continuously transferred from the first to the second dimension. In order to obtain maximum peak capacity, orthogonal systems with independent separation mechanisms should be used. Silica-based monolithic columns and solid-core superficially porous columns with diameters lower than 3 μm are usually chosen for the second RP dimension, as they offer higher permeability and can provide fast efficient separations at high flow rates and moderate column pressures, which allows their use with conventional HPLC instrumentation (Jandera *et al.*, 2012).

Comprehensive LC systems have been explored for the separation of complex polyphenol mixtures containing flavonoids. In these systems, it is preferable to use the more polar of the two columns in the first dimension and the more retentive (usually a bonded alkylsilica) in the second dimension, so as to suppress band broadening connected with the sample transfer (Jandera, Hajek, and Cesla, 2010). Hajek *et al.* (2008) combined a polyethylene glycol (PEG) column in the first dimension with different alkylsilica stationary phases in the second dimension for the separation of 27 antioxidant polyphenols including various flavonoids, concluding that monoliths and SPP columns provided lower hold-up time and improved resolution in the second dimension than totally porous columns of similar length. However, the better permeability of monolithic columns allowed using higher flow rates, thus leading to significant reduction in the 2D separation time compared with particle-based columns. Similar conclusions were obtained by Jandera *et al.* (2008) for the LC × LC separation of a mixture of 33 phenolic compounds, obtaining the best results using a PEG microcolumn in the first dimension and a short monolithic C18 column in the second one.

The use of comprehensive HILIC × RP systems for the separation of flavonoids has also been explored. In principle, the selectivity differences between both types of separation should provide better orthogonality in two-dimensional separations compared with RPLC × RPLC systems. However, there are compatibility problems derived of the differences in the elution strengths in the two modes. Mobile phases with high concentrations of organic solvents used in the first, HILIC, dimension are usually too strong for RP separation and may cause detrimental effects on the second dimension. To overcome these solvent incompatibilities, Jandera *et al.* (2012) used microbore or capillary polar columns and low flow rates in the HILIC dimension, with gradients of decreasing concentration of acetonitrile in buffered aqueous–organic mobile phases, coupled in line with short nonpolar or weakly polar monolithic or porous shell columns, with fast RP gradients of increasing acetonitrile concentration in the second dimension. With this arrangement, low but enough fraction volumes were transferred from the first to the second dimension, for separation using 1–2 min fast gradients, reducing the band broadening effect caused by the incompatibility between mobile phases. Using this approach, in the two-dimensional chromatograms, those authors obtained the separation of the polyphenols in seven classes according to structural features related to compound polarity: I, most polar phenolic acids; II, phenolic acids with two –OH groups; III, phenolic acids with a single –OH and two methoxy groups; IV, flavone aglycones; V, flavone glycosides; VI, other phenolic acids with a single –OH group; and VII, catechins.

Offline HILIC × RP-LC approaches were used by Kalili and de Villiers (2009, 2010) for the separation of procyanidins in cocoa, apple extracts, and green tea. Compounds were separated in a DIOL microcolumn in the first dimension, and 1-min fractions were collected and rechromatographed on a porous particle C18 column. Offline 2D-LC systems consisting of a CN column in the first dimension and a C18 column in the second dimension were used by Liu *et al.* (2008) and Zhou *et al.* (2008) for the separation of flavonoids in extracts of *Swertia franchetiana* Smith. and *Lobelia chinensis* Lour., respectively. Heart-cutting 2D-LC mode combining size exclusion chromatography in the first dimension and RP-LC in the second dimension was used by de Souza *et al.* (2009) for the separation of flavonol glycosides in extracts of the leaves of *Maytenus ilicifolia* Mart. ex Reissek. Centrifugal partition chromatography and RP-HPLC using a fused core HALO C18 column were hyphenated by Michel, Destandau, and Elfakir (2011) for the preparative fractionation of flavonoids from *Hippophae rhamnoides* L. berries and online analysis of collected fractions.

5.2.2.3 Chiral Separations

Some flavonoid classes are characterized for possessing one or more chiral carbons (e.g., flavanones, dihydroflavonols, catechins, or PAs). In those cases, separation of enantiomer or diastereoisomer compounds can be accomplished using chiral stationary phases (CSPs) or the addition of chiral additives to the mobile phase on conventional stationary phases. CSPs have been mainly used for the enantiomeric separation of flavanones (Yañez *et al.*, 2007), although chiral HPLC has been also applied for the separation of catechin diastereomers (Rinaldo *et al.*, 2010) or taxifolin enantiomers (Vega-Villa *et al.*, 2009). Chiral additives have been used to separate flavonoid enantiomers mainly by capillary electrophoresis (Cao, Qu, and Cheng, 2010; Kwon and Jung, 2011).

5.3 Other Separation Techniques

High speed counter current chromatography (HSCCC) is as an all-liquid chromatographic technique of great interest as an alternative to column chromatography for the isolation of flavonoids at preparative and semipreparative scale. It consists of a two-phase solvent system made up of a pair of immiscible solvents, one used as the stationary phase and the other as the mobile phase. An auxiliary solvent, miscible in both phases, may also be added to aid in the partitioning of the analytes between the two immiscible phases. HSCCC provides an advantage over the conventional column chromatography by eliminating the use of a solid support, thus avoiding compounds loss due to irreversible interactions with the solid stationary phase. Other advantages include

higher loading capacity, the ability to load and fractionate crude samples, the easy scale-up to larger fractionation, and the ability for complete sample recovery by pumping out the liquid stationary phase from the coil system. By contrast, the separation process is usually much longer than in HPLC. It was firstly introduced by Ito (1980a,b) and further applied by him and his coworkers for the isolation and purification of different types of flavonoids (Han *et al.*, 2007; Yang *et al.*, 2008; Wei *et al.*, 2009; He, Huang, and Ayupbek, 2010; Liang *et al.*, 2012) or their direct analysis coupled to different detection systems (Yanagida, Shoji, and Shibusawa, 2006; Chen *et al.*, 2005). Another active group on the application of this technique to flavonoid analysis has been Winterhalter and his coworkers, which applied it to the separation of catechins (Degenhardt *et al.*, 2000a,b), PAs (Köhler and Winterhalter, 2005; Köhler, Wray, and Winterhalter, 2008), flavonols (Du *et al.*, 2004a; Gutzeit *et al.*, 2007), flavanones (Du, Jerz, and Winterhalter, 2004b), or anthocyanins (Du, Jerz, and Winterhalter, 2004c; Montilla *et al.*, 2010). An exhaustive compilation on HSCCC applications to polyphenol separation was made by Sutherland and Fisher (2009), which reviewed more than 198 publications describing the isolation of 354 relatively pure different molecules (almost 50% of which were polyphenols), covering 108 different plant species from 56 plant families. A summary of more recent applications is included in Table 3.

Capillary electrophoresis (CE) has only been occasionally used for the separation and analysis of flavonoids. This technique has high separation efficiency and resolution power, using short times of analysis and low consumption of sample and reagents. Nevertheless, it seems unreliable that CE

Table 3 Examples of recent applications of HSCCC to flavonoid isolation from plant materials.

Plant material	Compounds	Solvent systems	Reference
Caesalpinia sappan L.	Homoisoflavonoids	Chloroform–methanol–water (4:3:2, vol/vol/vol)	Xu *et al.* (2012)
Hordeum vulgare L.	Lutonarin and saponarin	Ethyl acetate–*n*-butanol–water (3:2:5)	Chen *et al.* (2012)
Daphne genkwa Sieb. and Zucc.	Luteolin, apigenin, and genkwanin	*n*-Hexane–ethyl acetate–methanol–water (5:7:5:5)	Liang *et al.* (2012)
Alpinia katsumadai Hayata.	Alpinetin, pinocembrin, and cardamomin	Hexane–ethyl acetate–methanol–water (3:7:6:4)	Xiao *et al.* (2011)
Poacynum hendersonii Woodson.	Isoquercitrin and quercetin derivatives	*n*-Butanol–petroleum ether–0.5% acetic acid (5:3:5)	Shi *et al.* (2011)
Pogostemon cablin Benth.	Methoxyflavones	*n*-Hexane–ethyl acetate–methanol–water (11:5:11:5)	Li *et al.* (2011a)

can displace HPLC for the analysis of flavonoids. The repeatability of retention/migration times still is better in HPLC than in CE, and there is no dramatic difference of run times between both techniques, especially if compared with more recent LC approaches such as UHPLC (de Rijke *et al*., 2006). Furthermore, compared to HPLC, it shows low sensitivity in terms of solute concentration and worse reproducibility, caused by the short optical path length of the capillary used as detection cell and by the small volumes that can be introduced into the capillary (Liu *et al*., 2008). Actually, there are not many examples of CE applications to separate and determine naturally occurring flavonoids in plant-derived materials (see Yang, Zhao, and Li, 2010, for a review).

Gas chromatography (GC) is another separation technique with limited applicability in flavonoid analysis. These compounds do not usually meet requirements of volatility and have to be derivatized previous to their analysis by GC, which may yield various derivatives in flavonoids with various hydroxyl substituents, making quantification difficult. Another limitation is that applications are basically restricted to the analysis of aglycones, and flavonoid glycosides have to be hydrolyzed to the aglycones before GC analysis. An early procedure for the GC separation of up to 36 different flavonoids, including, flavones, flavonols, flavanones, isoflavones, and chalcones in the form of their methyl ethers, was described by Narasimhacha and Von Rudloff (1962), although, most usually, flavonoids are analyzed after derivatization to trimethylsilyl derivatives (Bankova *et al*., 1992; Stevens *et al*., 1996; Deng and Zito, 2003; Fuzfai and Molnar-Perl, 2007). ESI-MS in the selected ion-monitoring mode is often used for detection, typically using the molecular ion, $[M+H]^+$, and/or fragments formed by the loss of CH_3 and/or CO and retro-Diels-Alder (RDA) cleavage (de Rijke *et al*., 2006). Methods for the analysis of underivatized flavonoids by GC-MS have also been reported in the literature (Schulten, Simmleit, and Mueller, 1989; Schmidt, Merfort I, Matthiesen 1993; Schmidt, Merfort I, Willuhn, 1994). These approaches are usually based on the analysis of the pyrolysis fragments of the flavonoids, although some aglycones are stable and volatile enough at operation temperatures (350–350°C) and can be occasionally detected as their molecular ions. Christov and Bankova (1992) were also able to detect underivatized flavonoids in propolis using GC with electron-capture detection.

Supercritical fluid chromatography (SFC) is a relatively recent chromatographic technique that separates analytes using a mobile phase in conditions relatively close to its critical temperature and pressure. Typical SFC separations are carried out around room temperature and at pressures of 100–200 bar; so, the system is inherently subcritical. Therefore, in most situations, the term "supercritical" does not strictly refer to the actual state of the fluid (Rajendran, 2012). While several fluids (e.g., NH_3, SF_6, ethanol, and propane) have been used, SC-CO_2 is the most usual choice for the mobile phase. Similar to SFE, a modifier fluid, generally an alcohol, is used to improve the solvating ability of the supercritical fluid and the selectivity of the separation. SFC involves three types of gradient elution methods: temperature gradient such as GC, modifier programming such as LC, and pressure programming. In the same way, two types of analytical columns can be used, that is, packed columns such as HPLC and capillary columns such as GC (Taylor, 2010). Although, during the early days of this technology, SFC was widely expected to replace HPLC, nowadays, it is clear that HPLC continues playing a pivotal role in separation science; however, SFC is an attractive alternative for some applications such as chiral separation, fatty acids, triglycerides, and lipids analyses, or in pharmaceutical industry. However, not many applications of this technique can be found in the literature for the analysis of flavonoids. Among others, it has been applied to the analysis of polymethoxylated flavones (Morin *et al*., 1991; Hadj-Mahammed *et al*., 1993), the enantioselective separation of chiral flavanones (Gaggeri *et al*., 2011; Nguyen *et al*., 2012), or the separation and identification of grape seed polyphenols (Kamangerpour *et al*., 2002).

5.4 Detection Systems

5.4.1 Spectrophotometric Detection

Classical methods for HPLC analysis of flavonoids have mainly been based on UV–vis detection. Detection at 280 nm is most generally used for single recording in flavonoid mixtures. However, the various flavonoid subclasses possess characteristic spectra shapes with absorption maxima at particular wavelengths at which they can be selectively

recorded. Isoflavones typically exhibit their maximum at 255–265 nm, catechins and PAs at 270–280 nm, flavanones at 280–290 nm, flavones at 330–350 nm and 250–270 nm, flavonols at 330–380 nm and 250–270 nm, chalcones at 340–390 nm, and anthocyanins at 500–530 nm and 270–280 nm (Markham, 1982). Thus, multiple monitoring at different wavelengths (e.g., 260, 280, 290, 340, 370, and/or 520 nm) can be used for flavonoid classification and quantification. When diode array spectrophotometers are used as detectors (HPLC-DAD), the UV–visible spectra of the peaks can also be obtained, which notably improves the chances for compound identification. Peak purity can also be assessed by examination of the spectra and calculation of the ratio of absorbances at different wavelengths: If this ratio is constant over the whole width of the peak, it can be assumed that it was pure, unless two peaks with the same spectrum shape were overlapped. Patterns for the identification of different flavonoid classes based on their chromatographic behavior and online spectra characteristics can be found in different publications (Sakakibara *et al*., 2003; Santos-Buelga, Garcia-Viguera, and Tomas-Barberan, 2003; Campos and Markham, 2007; Abad-Garcia *et al*., 2009). In order to increase the possibilities of compound identification, Hostettmann *et al*. (1984) developed an HPLC-DAD method using detection after postcolumn reaction with UV shift reagents (0.5 M Na_2HPO_4, 0.3 M KOH, 0.3 M $AlCl_3$, and 0.7 M H_3BO_3 + 0.1 M NaO_2CCH_3), as classically used for the identification of pure flavonoids (Mabry, Markham, AND Thomas, 1970).

Despite its advantages, the unambiguous identification of compounds based on the information obtained by HPLC-DAD is not always possible, even when compared with reference substances or library databases, as UV–visible spectra of different flavonoids can be very similar and their shape may change depending on the separation conditions. Today, the use of HPLC-UV-DAD continues being useful for flavonoid screening, although its value for structural analysis has diminished compared to the level of information gained by other modern spectroscopic techniques such as NMR and MS (Fossen and Andersen, 2006). As for quantitative purposes, spectrophotometric detection is usually restricted to majority compounds, as it lacks the sensitivity required when dealing with minority flavonoid metabolites, for which other techniques such as fluorescence, electrochemical, and mass spectrometric detections offer better performance.

5.4.2   Postcolumn Derivatization

Photometric detection after a chemical postcolumn reaction has been used for selective monitoring of particular flavonoid groups. A procedure based on the postcolumn reaction with *p*-DMACA was developed by Treutter (1989) and Treutter *et al*. (1994) for the analysis of flavan-3-ols (catechins and PAs) in complex mixtures after HPLC separation. The DMACA reagent shows a high specificity for these flavonoids by forming colored adducts that show maximum absorption around 640 nm, which allows their selective detection and prevents the interference of other coeluting phenolic compounds. In addition to its selectivity, this derivatization method increases the sensitivity, when compared with the detection at 280 nm, improving the possibilities of flavan-3-ol quantification.

A method based on the postcolumn reaction after HPLC separation with the stable free radical DPPH$^\bullet$ (2,2-diphenyl-1-picrylhydrazyl) was developed by Koleva, Niederlander, and van Beek (2000) for the online detection of antioxidants, including flavonoids, in complex mixtures. This radical strongly absorbs in the visible region at λ_{max} 517 nm, whereas its reduced hydrazine derivative does not show significant absorbance above 450 nm. Postcolumn reaction of the antioxidants present in the sample with the radical produces a decrease in the absorbance at 517 nm that is detected in the form of negative peaks. The same authors (Koleva, Niederlander, and van Beek, 2001) described another online HPLC method based on the reaction with the ABTS$^{\bullet+}$ radical [2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid)]; in this case, the reduction of the radical also leads to a colorless product (ABTS), which is detected as a negative peak at 734 nm. The response in these postcolumn assays is greatly dependent on experimental variables such as reagent concentration, pH, dissolved oxygen content, reaction coil dimensions, and mobile phase composition. Those factors have to be optimized for a suitable detection, which have led to discrepancies between the methodologies used by different research groups (McDermott *et al*.,

2010). A similar method using chemiluminescence detection following a luminal-based postcolumn reaction was used by Ding *et al*. (2009) for the analysis of (antioxidant) flavonoids from *G. biloba* L.

5.4.3   Fluorometry

Fluorescence detection has been used occasionally in the flavonoid analysis. The number of flavonoids that exhibit significant native fluorescence is limited and include isoflavones and especially those that lack an OH group in the 5-position (de Rijke *et al*., 2002), 3-hydroxyflavone derivatives such as quercetin (Sengupta and Kasha, 1979), catechins (Arts and Hollman, 1998), and methoxylated flavones (Roussef and Ting, 1979). In those cases, fluorescence may be useful for detection as it minimizes background interference and shows greater selectivity and increases sensitivity up to by 10-fold compared to UV detection (Robbins *et al*., 2009).

Metal cations may bind to flavonoids through the hydroxyls at 3- or 5- position and 4-keto group and/or the hydroxyls at 3′- and 4′-, producing fluorescent complexes. Thus, complexation with metal ions can be used to enhance flavonoid fluorescence. Hollman, Van Trijp, and Buysman (1996) described an HPLC-fluorescence method for the analysis of flavonols after postcolumn derivatization with Al^{3+}. In that system, only compounds that contain both a 3-hydroxyl and a 4-keto group show sufficient fluorescence, whereas compounds lacking a free 3-hydroxyl group, such as flavonol glycosides substituted at that position, escape this detection, and for their analysis, deconjugation to the aglycone is previously required.

5.4.4   Electrochemical Detection

HPLC with ECD is another sensitive technique that can be applied for compounds able to be oxidized or reduced at low voltage potentials such as flavonoids, whose phenolic groups confer them electroactive properties. Nevertheless, there is no always a linear relation between the antioxidant activity and the electrochemical response, because other structural parameters may also play a role. For instance, glycosylation of the 3-OH group decreases the antioxidant activity in flavonols but not their electrochemical behavior (de Rijke *et al*., 2006). The

magnitude of the reduction potential for a given type of flavonoid is roughly related to the substitutions of the ring-B following the sequence catechol > methoxycatechol > monohydroxyl > methoxyl (Robards and Antolovich, 1997). Both voltammetric and coulometric detectors can be used, although the latter show higher sensitivity as they provide an almost complete oxidation/reduction of the analytes. Typical detection limits for ECD of flavonoids are about 10- to 100-fold higher than with UV detection and 10-fold higher than with FL detection (Manach, 2003).

Flavonoids present various waves of oxidation corresponding to different moieties capable of undergoing oxidation. Usually, only the lowest wave of oxidation is used for quantification, but with multielectrode detection systems (HPLC-CoulArray), the electrochemical behavior of a compound across the array of the applied potentials can be used with qualitative purposes and for peak purity assessment (Morand *et al*., 2001). Gradient elution was considered not suitable for electrochemical detection, because the change in the mobile phase causes a great baseline drift, but HPLC-CoulArray equipments incorporate software containing a gradient-correcting algorithm that removes the baseline drift and allow the use of gradient elution. Thus, chromatographic conditions do not differ from those used with UV detection, except that mobile phases must contain an electrolyte to facilitate the current flux, typically 30–100 m*M* sodium or lithium phosphate. The pH must also be carefully controlled as it influences the current/voltage curve of the compounds, and thus their optimal oxidation potential. A comprehensive review on the fundamentals and applications of this technique to flavonoid analysis was published by Manach (2003).

5.4.5   Nuclear Magnetic Resonance

Hyphenation of liquid chromatography to NMR spectroscopy (LC-NMR) emerged in the mid-1990s as a promising approach for the separation and structural elucidation of unknown compounds in mixtures. However, this technique shows some drawbacks that have prevented its growth, such as problems for solvent suppression, differences in chemical shifts compared to usual offline NMR analysis, lack of reference of online NMR spectra, long run times, or relatively low sensitivity, thus requiring more amount

of sample and higher solvent volumes, which affects the chromatographic resolution and separation leading to peak broadening (Silva Elipe, 2003; Wolfender, Ndjoko, and Hostettmann, 2003). The information provided by LC-NMR consists mainly of ^{1}H-NMR spectra; access to ^{13}C-NMR is possible, but it is restricted to a very limited number of cases where the concentration of the chromatographic peak is high (Marston and Hostettmann, 2006). A possibility for increasing the sensitivity is performing experiments in stop-flow mode, in which the flow is stopped when a compound is detected by another detection system (usually a photometric detector) so that it can remain in the NMR probe for enough time (even for days) for acquisition of the spectrum. This mode may also offer the possibility to perform homo- and heteronuclear 2D NMR correlation experiments (Wolfender, Ndjoko, and Hostettmann, 2001). In most cases, LC-NMR data are supported by results obtained by LC-MS. Although the direct connection of an HPLC with both spectrometers in one system is possible, it is not a frequent arrangement because of the need of use of deuterated water as a solvent for NMR. In these conditions, some protons in the flavonoid molecules are exchanged for deuterium atoms, causing the increase of the apparent molecular weights of analytes (Stobiecki and Kachlicki, 2006).

Despite the advances made concerning the improvement in NMR flow probe technology, introduction of new techniques for solvent suppression or coupling to an online SPE unit for concentrating and focusing analyte peaks before spectra recording (LC-SPE-NMR), which have contributed to improve the sensitivity, this hyphenated technique has not yet become popular among others owing to the high price of the equipments. Indeed, the LC-NMR equipments are not available to many laboratories; so, only few articles have been published applying this technique to the analysis of flavonoids (Le Gall et al., 2003; Wolfender, Ndjoko, and Hostettmann, 2003; Moco et al., 2006; Timmers and Urban, 2012).

5.5 LC Coupled to MS

Over the past 20 years, with the generalization of atmospheric pressure ionization systems, LC-MS has become a routine technique for the analysis of flavonoids, being ESI and atmospheric pressure chemical ionization (APCI), the most widely used interfaces. Thermospray ionization, initially used by

some authors (Pietta et al., 1994; Gabetta et al., 2000), has been progressively displaced, as it requires higher temperatures in the ion source for efficient ionization of the molecules to be analyzed, which affects thermal stability of the compounds (Stobiecki, 2000). Although there is not a clear tendency in practice, ESI may suit better for the analysis of polar, ionized, or ionizable molecules and APCI that of less polar compounds. Predominant LC-MS configurations are based on quadrupole (Q), ion-trap (IT), and time-of flight (TOF) mass analyzers. Q and IT are low resolution analyzers, with upper limits for mass-to-charge ratio (m/z) detection reaching up to 3000 and 6000, respectively, which is appropriate for most flavonoids. TOF analyzers show much better mass accuracy and higher sensitivity than Q and IT analyzers, within an m/z range up to 20,000 (Holcapek, Jirasko, and Lisa, 2012). Fragmentation of the compounds in view to increase the possibilities of identification can be produced by collision-induced dissociation (CID) in IT and triple quadrupole (TQ) mass analyzers. In the ITs, ions generated in the ionization source can be selectively trapped and fragmented a number of times so as to perform sequential fragmentation of parent and daughter ions, thus providing MS^n spectra. TQs consist of a sequence of three quadrupole cells that successively act as a filter for selected precursor ions, a collision cell, and a detector for the product ions, yielding MS/MS spectra (i.e., tandem-MS). High resolution MS/MS or MS^n analyses can also be performed using hybrid Q-TOF or IT-TOF equipments, in which the fragment ions produced in the Q or IT are injected into the TOF analyzer. In recent years, equipments based on Fourier transform, such as Orbitrap and ion cyclotron resonance (ICR) mass analyzers that offer ultrahigh resolution power and mass accuracies up to 1 ppm and below, have been introduced, which open the possibility of very narrow widths for precursor ion isolation and leads to increased selectivity and sensitivity.

Both positive and negative ion modes can be used for flavonoid detection. The negative mode usually provides best sensitivity for most flavonoid families, except anthocyanins. The selection of the mobile phase is a critical factor in achieving not only good chromatographic separation but also appropriate ionization. HPLC separations of flavonoid mixtures are most usually performed on RP C18 columns with acidified mobile phases, quite usually with acetic or formic acids. In these conditions, in positive ion mode, strong suppression of cationic adducts

([M+Na]$^+$ or [M+K]$^+$ ions) is observed in mass spectra although an increase in sensitivity may be achieved owing to the presence of single [M+H]$^+$ ions (Stobiecki and Kachlicki, 2006). Stronger acids such as trifluoroacetic acid are less advisable for LC-MS analysis, especially in negative mode, as it suppresses the ionization owing to ion-pairing and surface tension effects (de Rijke *et al.*, 2006). Rauha, Vuorela, and Kostiainen (2001) found that, in negative ion mode, the optimal ionization efficiency for flavonoids was achieved using ammonium acetate buffer at pH 4.0 as the aqueous component of the LC solvent system, whereas 0.4% formic acid (pH 2.3) gave the best results in positive ion mode, both in ESI and APCI. Comparable effects may be achieved using ammonium acetate, in which case strong ammonium adduct ions [M+NH$_4$]$^+$ are observed (de Rijke *et al.*, 2003).

Soft ionization methods used in LC-MS do not typically produce much fragmentation of flavonoid compounds. Typical mass spectra, as obtained in APCI or ESI equipments, usually provide information about the molecular weight of the flavonoid (detected as a protonated molecule, [M+H]$^+$, adduct, [M+Na]$^+$, or as a deprotonated molecule, [M−H]$^−$) and the nature of aglycone and substituting (acyl)glycosidic residues if some fragmentation is produced. APCI usually produces more fragments than ESI, and, therefore, intact flavonoid glycosides are normally best observed under ESI conditions. Additional information for the identification of the aglycone and the type and location of substituents can be obtained from the analysis of MS/MS and MS^n spectra, although no information can be usually obtained about the stereochemistry of the glycosidic linkages and the identity of diasteromeric sugar units. Differentiation between *O*-glycosides, *C*-glycosides, and *O,C*-diglycosides can be made by examining first-order ion spectra or low energy CID spectra (Bravo and Mateos, 2008), but fragmentation of the aglycone so as to obtain data for its characterization requires higher energy levels. In flavonoid *O*-glycosides, the cleavage at the glycosidic *O*-linkages with a concomitant H-rearrangement leads to Y$_n$ type ions (Figure 2) from the elimination of monosaccharide residues [i.e., −162 u (hexose), −146 u (deoxyhexose), −132 u (pentose), or −176 u (uronic acid)], which can be used for the determination of the sugar substitution pattern and carbohydrate sequence (Cuyckens and Claeys, 2004). When acyl substituents exist, they remain attached to the sugar residues,
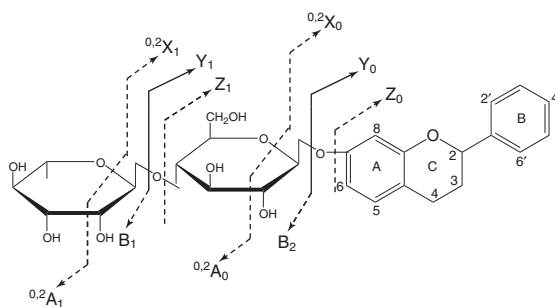
Figure 2 Ion nomeclature used for *O*-glycosides. *Source*: Adapted from H. Cuyckens and M. Claeys (2004). © John Wiley & Sons, Ltd.

so that no particular fragments from their loss are usually observed. Particular care should be taken as some acyl residues possess the same mass units as sugar moieties [e.g., caffeoyl (162 u), coumaroyl (146 u), and feruloyl (−176 u)]. In the case of *O*-diglycosides, differentiation between disaccharides containing 1,2- and 1,6-interglycosidic linkage, such as in *O*-neohesperidosides and *O*-rutinosides, can be achieved based on the relative abundances of Y$_0$ and Y$_1$ ions formed by fragmentation at glycosidic bonds. Greater abundance of the Y$_0$ ion is observed in the case of a 1→2 linkage, whereas Y$_1$ is more abundant in the case of a 1→6 linkage. This differentiation can be more clearly made in the negative ion mode (Cuyckens *et al.*, 2001). In the fragmentation of *O*-neohesperidosides and *O*-rutinosides, an irregular Y* ion may also be produced owing to a rearrangement in the sequence of sugar moieties. The formation of Y* is favored at relatively low collision energies and has been rationalized by loss of an internal dehydrated glucose residue and a mechanism involving a mobile proton from the aglycone to the terminal rhamnose. The relative abundance of the Y* fragment is usually higher for *O*-rutinosides than for *O*-neohesperidosides (Ma *et al.*, 2001).

C-glycosides are characterized by the loss of characteristics fragments from the cleavage of the sugar pyrano ring (Figure 3) and sugar dehydrations. Guidelines for the systematic characterization of *C*-glycosyl flavones based on different diagnostic product ions obtained by fragmentation using ESI tandem MS have been reported by Abad-Garcia *et al.* (2008). According to them, the C-6 glycosides are more fragmented than the corresponding C-8 isomers, and they proposed two diagnostic product

$^{0,2}X$: [M ± H − 120]
$^{0,3}X$: [M ± H − 90]
$^{1,5}X$: [M ± H − 134]

$^{0,2}X$: [M ± H − 90]
$^{0,3}X$: [M ± H − 60]
$^{1,5}X$: [M ± H − 104]
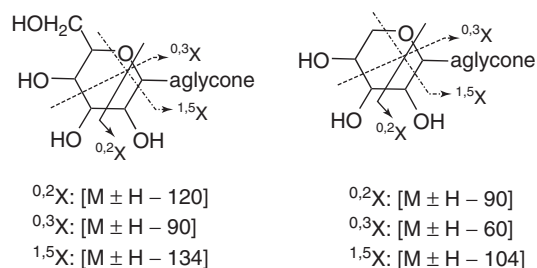
Figure 3 Characteristic product ions formed by cross-ring cleavages in a hexose and pentose residue. *Source*: Adapted from H. Cuyckens and M. Claeys (2004). © John Wiley & Sons, Ltd.

Flavonols

Flavones

Flavan-3-ols

Anthocyanins

Figure 4 Schematic summary of principal fragment patterns for various types of flavonoids.

ions for C-6 ($[M+H-4H_2O]^+$) and C-8-glycosides ($[^{0,3}X]^+$). Di-*C*-glycosides fragment in the same way as mono-*C*-glycosides, but, as the number of possibilities increases, the spectra become more complex.

Most characteristic ions for flavonoid algycones are due to fragmentations in the ring-C, such as RDA cleavage, heterocyclic ring fission (HRF), or quinone methide (QM) cleavage. A schematic summary of principal fragment patterns for different types of flavonoids in presented in Figure 4. In addition to ring-C cleavage, the loss of small neutral molecules or radicals, such as CH_3 (15 u), H_2O (18 u), CH_2O (30 u), C_2H_2O (42 u), CO_2 (44 u), CH_2CHOH (46 u), COCO (56 u), or CO_2CO (72 u), is commonly observed and can be useful to identify the presence of specific functional groups. The identification of prenylated flavonoids can be deduced from the loss of isobutene (C_4H_8, 56 u), although it is not observed in the negative ion mode (Bravo and Mateos, 2008). Generally, fragments issued from the ring-C cleavage are easier to observe on a Q-TOF instrument, whereas losses of small molecules are favored in IT-MS (Marston and Hostettmann, 2006).

Flavan-3-ols are a specific type of flavonoids that do not usually occur as glycosylated forms but can be found as condensed structures (i.e., PAs) that show characteristic fragmentation behavior (Figure 5). HRF and RDA cleavage have been reported to occur fundamentally in the upper subunit of the PAs so that the analysis of the produced fragments can provide information about the relative position of elementary units and their hydroxylation patterns (de Pascual-Teresa, Rivas-Gonzalo, and Santos-Buelga, 2000; Friedrich, Eberhardt, and Galensa, 2000; Gu *et al*., 2003). The sequence of elementary units in PA oligomers can be further deduced through diagnostic ions derived from the QM cleavage of the
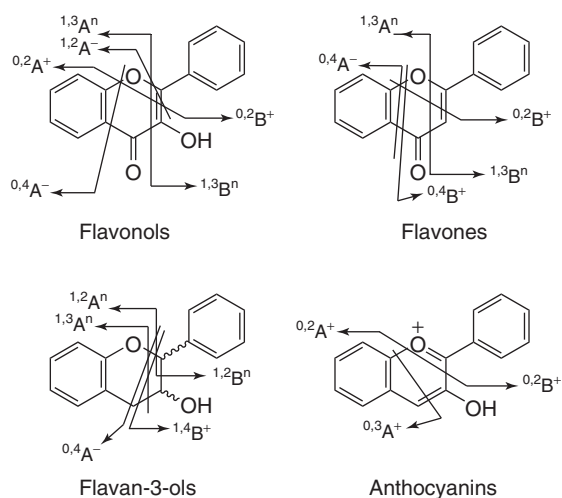
interflavan bond, as lower units are released intact, whereas the upper units suffer a structural rearrangement, yielding ions 2 Da lower than the original flavanol constituents (Friedrich, Eberhardt, and Galensa, 2000; Gu *et al*., 2003). Fragmentation patterns are similar for PA dimers to tetramers, although spectra become more complex as the number of elementary units increases due to the alternative cleavage of the different interflavan bonds. Mass spectra do not allow, however, differentiating isomeric catechins [e.g., (gallo)catechin/epi(gallo)catechin, or afzelechin/epiafzelechin] nor establishing the position of the linkage between flavanol units (i.e., C4-C8 or C4-C6). Similar principles as for PAs have been observed in the case of anthocyanin–flavanol condensed pigments found in some plant natural sources (Gonzalez-Paramás *et al*., 2006) and in aged red wines (Alcalde-Eon *et al*., 2007).

Polymeric PAs (from tetramers on) are not well separated by HPLC and cannot be detected as individual chromatographic peaks for their identification. Methods for the characterization of PA polymers and complex PA mixtures in terms of size distribution, mean DP, and/or type of constituting flavan-3-ol units have been described using MALDI-MS and ESI-MS, either by direct infusion or following LC injection (Hayasaka *et al*., 2003; Gonzalez-Manzano *et al*., 2006; Monagas *et al*., 2010; Mouls *et al*., 2011; Perez-Jimenez and Torres, 2012). It should be taken into account that PA oligo/polymers can appear in
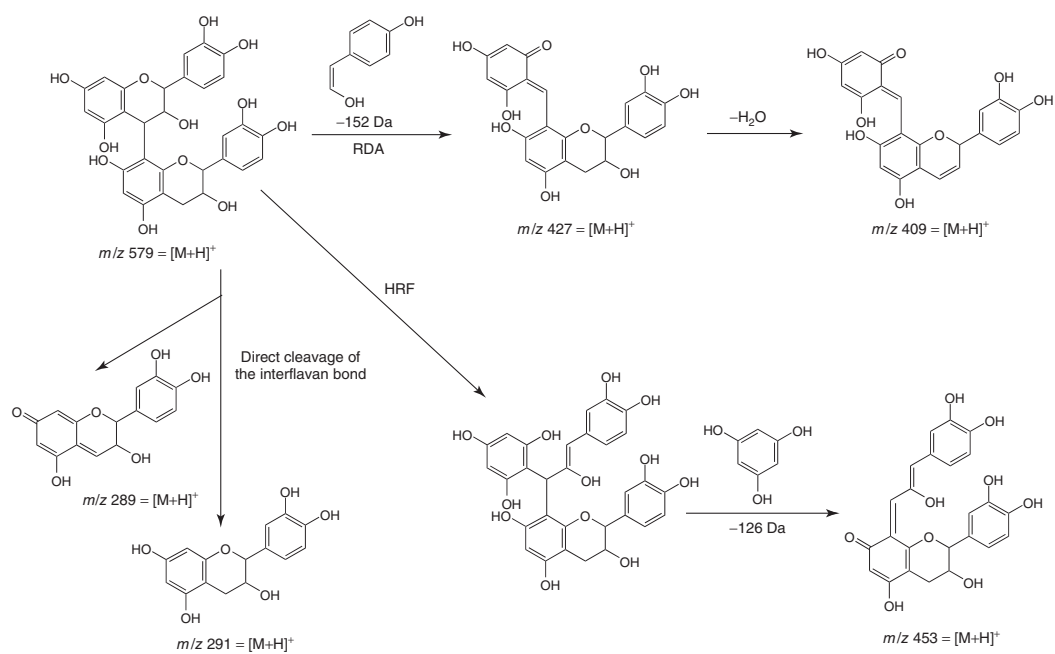
Figure 5 Fragmentation patterns of proanthocyanidin oligomers.

mass spectra as multiple charged species (e.g., doubly and triply charged ions), which can be detected by their carbon isotope distribution using zoom mass scan recording. In the case of a single-charged ion, the distance between the isotopic peaks of the carbon would be m/z 1, in the case of a double charged ion m/z 0.5, and so forth (Cheynier *et al.*, 1997; Wollgast *et al.*, 2001). Mouls *et al.* (2011) found that the sensitivity for ions corresponding to PAs with higher DPs is increased in the positive ionization mode compared to negative ionization. Those authors also reported that, although MALDI-MS allows the detection of slightly higher DPs than ESI-MS, better discrimination of the lower molecular ions (DP < 15) is observed in ESI-MS.

Structural identification of flavonoids based on MS analysis has been object of many good articles that can be consulted for more detailed information (Ma *et al.*, 1997, 2001; Justesen, 2000; Stobiecki, 2000; Cuyckens and Claeys, 2004; Abad-Garcia *et al.*, 2009; Justino, Borges, and Florencio, 2009; Vukics and Guttman, 2010).

Quite usually, HPLC-MS is used in combination with other detection systems and specially diode array detection, which facilitates the assignation of the flavonoid class based on the UV–vis spectra of the compounds. Additional information can also be obtained from the chromatographic retention behavior of the compounds, as influenced by the type of flavonoid and substitution pattern. Differentiation of isobaric compounds (with the same nominal mass but different elemental composition) can also be made from their exact mass values, as provided by TOF detectors. A library of phenolic compounds, including flavonoids, identified according to their retention time, MS, and MS/MS spectra obtained using HPLC interfaced to an ESI-Qq-TOF mass spectrometer, which provides mass accuracy and true isotopic pattern, has been recently published by Gomez-Romero *et al.* (2011).

6 CONCLUSION

The large structural diversity of flavonoids greatly complicates their analysis, and there is no standardized protocol that can be recommended for all type of compounds and plant materials. Procedures have to be optimized depending on the matrix, the nature of the analytes, and the object of the analysis (e.g., structure elucidation, quantification of individual or total flavonoids, and metabolic profiling). Sample preparation and extraction are critical steps to ensure adequate recovery of the target flavonoids and prevent chemical modifications. Solvent extraction is still commonly used, but assisted extraction methods, as those using ultrasounds, microwaves, or pressurized fluids, are currently gaining place. Crude extracts usually contain unwanted compounds, including other phenolic compounds, and a purification step is usually required. SPE is the most common technique for the purification of extracts and/or isolation of compounds, although more modern approaches, such MSPD, MEPS, or molecular imprinting technologies, have arisen as alternatives for selective compound extraction.

HPLC is the technique that has dominated the field of flavonoid separation and analysis for the past 30 years. The introduction of newer technical developments, such as UHPLC, HTLC, MDLC, or novel packing materials for columns, has provided notable improvements in resolution, efficiency, and/or speed of separation, which has further increased the scope of the HPLC technique. Other separation techniques such as GC or electrophoresis have found less application in flavonoid analysis, whereas HSCCC has emerged as a suitable methodology for preparative isolation of flavonoids. The coupling of HPLC to different detection systems (e.g., spectrophotometry, MS, fluorometry, NMR, or ECDs) allows integrating separation, characterization, and quantification steps without requiring high purity of extracts. In particular, the use of LC-MS has become a routine for the analysis of flavonoids, so that it can be considered the technique of choice both for compound identification and quantification. ESI is the most widely used LC-MS interface, with predominant configurations based on Q, IT, and TOF mass analyzers. More recently, equipments based on Fourier transform, such as Orbitrap and ICR mass analysers, have been introduced, which offer ultrahigh resolution power and greater mass accuracies, increasing both selectivity and sensitivity. LC-NMR is another promising powerful technique for the separation and structural elucidation of compounds that has not yet offered its full potential, mostly because of the existence of technical limitations and the cost of the instruments, although an increase in its applications should be expected in coming years with the introduction of improved and cheaper equipments, making the technique available to more laboratories.

ACKNOWLEDGMENT

REFERENCES

Abad-Garcia, B., Garmon-Lobato, S., Berrueta, L. A., *et al*. (2008) *Rapid Commun. Mass Spectrom.*, **22**, 1834–1842.

Abad-Garcia, B., Berrueta, L. A., Garmon-Lobato, S., *et al*. (2009) *J. Chromatogr.*, **1216**, 5398–5415.

Adamovich, J. and Stermitz, F. R. (1976) *J. Chromatogr.*, **129**, 464–465.

Alcalde-Eon, C., Escribano-Bailón, M. T., Santos-Buelga, C., *et al*. (2007) *J. Mass Spectrom.*, **42**, 735–748.

Andersen, O. M., and Markham, K. R. eds. (2006) *Flavonoids: Chemistry, Biochemistry and Applications*, Taylor & Francis 1237 pp, Boca Ratón.

Arts, I. C. W. and Hollman, P. C. H. (1998) *J. Agric. Food Chem.*, **46**, 5156–5162.

Bankova, V., Christov, R., Stoev, G., *et al*. (1992) *J. Chromatogr.*, **607**, 150–153.

Barker, S. A. (2007) *J. Biochem. Biophys. Methods*, **70**, 151–162.

Biesaga, M. (2011) *J. Chromatogr. A*, **1218**, 2505–2512.

Biesaga, M. and Pyrzynska, K. (2009) *Crit. Rev. Anal. Chem.*, **39**, 95–107.

Blahova, E., Lehotay, J. and Skacani, I. (2004) *J. Liq. Chromatog. Rel. Techn.*, **27**, 2715–2731.

Boros, B., Jakabova, S., Dornyei, A., *et al*. (2010) *J. Chromatogr. A*, **1217**, 7972–7980.

Bravo, L. and Mateos, R. (2008) Analysis of flavonoids in functional foods and nutraceuticals, in *Methods of Analysis for Functional Foods and Nutraceuticals*, 2nd edn, ed. W. J. Hurst, CRC Press, Boca Raton, FL, pp. 145–204.

Broadhurst, R. B. and Jones, W. T. (1978) *J. Sci. Food Agric.*, **29**, 788–794.

Buszewski, B. and Noga, S. (2012) *Anal. Bioanal. Chem.*, **402**, 231–247.

Cai, R., Wang, S., Meng, Y., *et al*. (2012) *Anal. Meth.*, **4**, 2388–2395.

Campos MG, Markham KR. 2007. Structure Information from HPLC and On-line measured Absorption Spectra. Imprensa de Universidad de Coimbra, Portugal, 118 pp.

Cao, J., Qu, H. B. and Cheng, Y. Y. (2010) *Electrophoresis*, **31**, 1689–1696.

Castro-López, M. M., Cela-Pérez, M. C., Dopico-García, M. S., *et al*. (2012) *Anal. Chim. Acta*, **721**, 68–78.

Castro-Vargas, H. I., Rodríguez-Varela, L. I., Ferreira, S. R. S., *et al*. (2010) *J. Supercrit. Fluids*, **51**, 319–324.

Chen, L. J., Song, H., Du, Q. Z., *et al*. (2005) *J. Liq. Chromatogr. Relat. Technol.*, **28**, 1549–1555.

Chen, Q., Zhao, J., Liu, M., *et al*. (2008) *J. Pharm. Biomed. Anal.*, **46**, 568–573.

Chen, S., Xing, X. H., Huang, J. J., *et al*. (2011) *Enzyme Microb. Technol.*, **48**, 100–105.

Chen, T., Wang, P., Du, Y., *et al*. (2012) *J. Liq. Chromatogr. Relat. Technol.*, **35**, 2524–2532.

Cheynier, V., Doco, T., Fulcrand, H., *et al*. (1997) *Analysis*, **25**, M32–M37.

Christov, R. and Bankova, V. (1992) *J. Chromatogr.*, **623**, 182–185.

Cuyckens, H. and Claeys, M. (2004) *J. Mass Spectrom.*, **39**, 1–15.

Cuyckens, H., Rozenberg, R., de Hoffmann, E., *et al*. (2001) *J. Mass Spectrom.*, **39**, 1203–1210.

Dai, J. and Mumper, R. J. (2010) *Molecules*, **15**, 7313–7352.

Daigle, D. J. and Conkerton, E. J. (1983) *J. Liq. Chromatogr.*, **6**, 105–118.

Daigle, D. J. and Conkerton, E. J. (1988) *J. Liq. Chromatogr.*, **11**, 309–332.

Degenhardt, A., Engelhardt, U. H., Lakenbrink, C., *et al*. (2000a) *J. Agric. Food Chem.*, **48**, 3425–3430.

Degenhardt, A., Engelhardt, U. H., Wendt, A. S., *et al*. (2000b) *J. Agric. Food Chem.*, **48**, 5200–5205.

Delcour, J. A. and Varebeke, D. J. (1985) *J. Inst. Brew.*, **91**, 37–40.

Deng, F. and Zito, S. W. (2003) *J. Chromatogr. A*, **986**, 121–127.

Ding, S. J., Dudley, E., Chen, L. J., *et al*. (2006) *Rapid Comm. Mass Spectrom.*, **20**, 3619–3624.

Ding, X. P., Qi, J., Chang, Y. X., *et al*. (2009) *J. Chromatogr. A*, **1216**, 2204–2210.

Du, Q., Chen, P., Jerz, G., *et al*. (2004a) *J. Chromatogr. A*, **1040**, 147–149.

Du, Q., Jerz, G. and Winterhalter, P. (2004b) *J. Liq. Chromatogr. Relat. Technol.*, **27**, 3257–3264.

Du, Q., Jerz, G. and Winterhalter, P. (2004c) *J. Chromatogr. A*, **1045**, 59–63.

Du, F. Y., Xiao, X. H. and Luo, X. J. (2009) *Talanta*, **78**, 1177–1184.

Dugo, P., Cacciola, F., Donato, P., *et al*. (2009) *J. Chromatogr. A*, **1216**, 7213–7221.

Escarpa, A. and Gonzalez, M. C. (2001) *Crit. Rev. Anal. Chem.*, **31**, 57–139.

Escribano-Bailón, M. and Santos-Buelga, C. (2003) Polyphenol extraction from foods, in *Methods in Polyphenol Analysis*, eds. C. Santos-Buelga and G. Williamson, Royal Society of Chemistry, Cambridge, pp. 1–16.

Eskilsson, C. S. and Bjorklund, E. (2000) *J. Chromatogr.*, **902**, 227–250.

Eugster, P. J., Guillarme, D., Rudaz, S., *et al*. (2011) *J. AOAC Int.*, **94**, 51–70.

Feng, J. T., Guo, Z. M., Shi, H., *et al*. (2010) *Talanta*, **81**, 1870–1876.

Fossen, T. and Andersen, O. M. (2006) Spectroscopic techniques applied to flavonoids, in *Flavonoids: Chemistry, Biochemistry and Applications*, eds. O. Andersen and R. Markham, CRC Press, Boca Raton, pp. 37–142.

Friedrich, W., Eberhardt, A. and Galensa, R. (2000) *Eur. Food Res. Technol.*, **211**, 56–64.

Fuzfai, Z. and Molnar-Perl, I. (2007) *J. Chromatogr. A*, **1149**, 88–101.

Gabetta, B., Fuzzati, N., Griffini, A., *et al*. (2000) *Fitoterapia*, **71**, 162–175.

Gaggeri, R., Rossi, D., Collina, S., *et al*. (2011) *J. Chromatogr. A*, **1218**, 5414–5422.

Geng, Y. and Xiang, B. (2008) *J. Near Infrared Spectrosc.*, **16**, 551–559.

Gomez-Romero, M., Zurek, G., Schneider, B., *et al*. (2011) *Food Chem.*, **124**, 379–386.

Gonzalez-Manzano, S., Santos-Buelga, C., Perez, J. J., *et al*. (2006) *J. Agric. Food Chem.*, **54**, 4326–4332.

Gonzalez-Paramás, A. M., Lopes, F., Martin, P., *et al*. (2006) *Food Chem.*, **94**, 428–436.

Gonzalez-Paramás, A. M., Santos-Buelga, C., Dueñas, M., *et al*. (2011) *Mini Rev. Med. Chem.*, **11**, 1239–1255.

Graham, H. D. (1992) *J. Agric. Food Chem.*, **40**, 801–805.

Gu, L., Kelm, M. A., Hammerstone, J. F., *et al*. (2003) *J. Mass Spectrom.*, **38**, 1272–1280.

Guillarme, D., Ruta, J., Rudaz, S., *et al*. (2010) *Anal. Bioanal. Chem.*, **397**, 1069–1082.

Gutzeit, D., Wray, V., Winterhalter, P., *et al*. (2007) *Chromatographia*, **65**, 1–7.

Hadj-Mahammed, M., Badjah-Hadj-Ahmed, Y. and Meklati, B. Y. (1993) *Phytochem. Anal.*, **4**, 275–278.

Haginaka, J., Tabo, H., Ichitani, M., *et al*. (2007) *J. Chromatogr. A*, **1156**, 45–50.

Hajek, T., Skerikova, V., Cesla, P., *et al*. (2008) *J. Sep. Sci.*, **31**, 3309–3328.

Hammerstone, J. F., Lazarus, S. A., Mitchell, A. E., *et al*. (1999) *J. Agric. Food Chem.*, **47**, 490–496.

Han, X., Ma, X., Zhang, T., *et al*. (2007) *J. Chromatogr. A*, **1151**, 180–182.

Hanhineva, K., Rogachev, I., Kokko, H., *et al*. (2008) *Phytochemistry*, **69**, 2463–2481.

Harborne, J. B. (1965) *Phytochemistry*, **4**, 107–120.

Harborne, J. B. ed. (1988) The Flavonoids: Advances in Research since 1980, Chapman and Hall900 pp, London.

Harborne, J. B. ed. (1994) The Flavonoids: Advances in Research since 1986, Chapman and Hall676 pp, London.

Harborne, J. B., and Mabry, T. J. eds. (1982) The Flavonoids: Advances in Research, Chapman and Hall744 pp, London.

Harborne, J. B., Mabry, T. J., and Mabry, H. eds. (1975) *The Flavonoids*, Chapman and Hall1204 pp, London.

Hayasaka, Y., Waters, E. J., Cheynier, V., *et al*. (2003) *Rapid Commun. Mass Spectrom.*, **17**, 9–16.

He, D. J., Huang, Y. and Ayupbek, A. (2010) *J. Liq. Chromat. Rel. Techn.*, **33**, 615–628.

Herrero, M., Temirzoda, T. N., Segura-Carretero, A., *et al*. (2011) *J. Chromatogr. A*, **1218**, 7511–7520.

Holcapek, M., Jirasko, R. and Lisa, M. (2012) *J. Chromatogr. A*, **1259**, 3–15.

Hollman, P. C. H., Van Trijp, J. M. and Buysman, M. N. (1996) *Anal. Chem.*, **68**, 3511–3515.

Hostettmann, K., Domon, B., Schaufelberger, D., *et al*. (1984) *J. Chromatogr.*, **283**, 137–147.

Hung, P. V. and Duy, T. L. (2012) *Int. Food Res. J.*, **19**, 327–332.

Ito, Y. (1980a) *J. Chromatogr.*, **188**, 33–42.

Ito, Y. (1980b) *J. Chromatogr.*, **188**, 43–60.

Jandera, P., Cesla, P., Hajek, T., *et al*. (2008) *J. Chromatogr. A*, **1189**, 207–220.

Jandera, P., Hajek, T. and Cesla, P. (2010) *J. Sep. Sci.*, **33**, 1382–1397.

Jandera, P., Hajek, T., Stankova, M., *et al*. (2012) *J. Chromatogr. A*, **1268**, 91–101.

Jian, W., Edom, R. W., Xu, Y., *et al*. (2010) *J. Sep. Sci.*, **33**, 681–697.

Ju, Z. Y. and Howard, L. R. (2003) *J. Agric. Food Chem.*, **51**, 5207–5213.

Justesen, U. (2000) *J. Chromatogr. A*, **902**, 369–379.

Justino, G. C., Borges, C. M. and Florencio, M. H. (2009) *Rapid Comm. Mass Spectrom.*, **23**, 237–248.

Kalili, K. M., and de Villiers, A. (2009) *J. Chromatogr. A*, **1216**, 6274–6284.

Kalili, K. M., and de Villiers, A. (2010) *J. Sep. Sci.*, **33**, 853–863.

Kalili, K. M., and de Villiers, A. (2011) *J. Sep. Sci.*, **34**, 854–876.

Kamangerpour, A., Ashraf-Khorassani, M., Taylor, L. T., *et al*. (2002) *Chromatographia*, **55**, 417–421.

Klejdus, B., Lojkova, L., Lapcik, O., *et al*. (2005) *J. Sep. Sci.*, **28**, 1334–1346.

Ko, M. J., Cheigh, C. I., Cho, S. W., *et al*. (2011) *J. Food Eng.*, **102**, 327–333.

Köhler, N. and Winterhalter, P. (2005) *J. Chromatogr. A*, **1072**, 217–222.

Köhler, N., Wray, V. and Winterhalter, P. (2008) *J. Agric. Food Chem.*, **56**, 5374–5385.

Koleva, I. I., Niederlander, H. A. G., and van Beek, T. A. (2000) *Anal. Chem.*, **72**, 2323–2328.

Koleva, I. I., Niederlander, H. A. G., and van Beek, T. A. (2001) *Anal. Chem.*, **73**, 3373–3381.

Kuhnle, G. C. C., Dell'aquila, C., Low, Y. L., *et al*. (2007) *Anal. Chem.*, **79**, 9234–9239.

Kwon, C. and Jung, S. (2011) *Carbohyd. Res.*, **346**, 133–139.

Le Gall, G., DuPont, M. S., Mellon, F. A., *et al*. (2003) *J. Agric. Food Chem.*, **51**, 2438–2446.

Li, K., Zhang, H., Xie, H., *et al*. (2011a) *J. Liq. Chromatogr. Relat. Technol.*, **34**, 1617–1629.

Li, W., Xing, L., Cai, Y., *et al*. (2011b) *Vib. Spectrosc.*, **55**, 58–64.

Liang, S., Liang, Y., He, J., *et al*. (2012) *J. Liq. Chromatogr. Relat. Technol.*, **35**, 2610–2622.

Liazid, A., Palma, M., Brigui, J., *et al*. (2007) *J. Chromatogr. A*, **1140**, 29–34.

Lin, L. Z., Sun, J., Chen, P., *et al*. (2011) *J. Agric. Food Chem.*, **59**, 12059–12072.

Liu, E. H., Qi, L. W., Cao, J., *et al*. (2008) *Molecules*, **13**, 2521–2544.

Liza, M. S., Abdul, R. R., Mandana, B., *et al*. (2012) *Int. Food Res. J.*, **19**, 503–508.

Luthria, D. L. (2006) *J. Sci. Food Agric.*, **86**, 2266–2272.

Luypaert, J., Zhang, M. H. and Massart, D. L. (2003) *Anal. Chim. Acta*, **478**, 303–312.

Ma, Y. L., Li, Q. M., Van den Heuvel, H., *et al*. (1997) *Rapid Comm. Mass Spectrom.*, **11**, 1357–1364.

Ma, Y. L., Cuyckens, F., Van den Heuvel, H., *et al*. (2001) *Phytochem. Anal.*, **12**, 159–165.

Ma, Y. Q., Chen, J. C., Liu, D. H., *et al*. (2009) *Ultrason. Sonochem.*, **16**, 57–62.

Ma, X., Chen, Z., Chen, R., *et al*. (2011) *Polym. Int.*, **60**, 1455–1460.

Mabry, T. J., Markham, K. R. and Thomas, M. B. (1970) *The Systematic Identification of Flavonoids*, Springer-Verlag354 pp, New York.

Maier, T., Goppert, A., Kammerer, D. R., *et al*. (2008) *Eur. Food Res. Technol.*, **227**, 267–275.

Majors, R. E. (2012) *LC GC Europe*, **October**, 7–14.

Manach, C. (2003) The use of HPLC with coulometric detection in the analysis of flavonoids in complex matrices, in *Methods in Polyphenol Analysis*, eds. C. Santos-Buelga and G. Williamson, The Royal Society of Chemistry, Cambridge, pp. 63–91.

Manchon, N., D'Arrigo, M., Garcia-Lafuente, A., *et al*. (2010) *Talanta*, **82**, 1986–1994.

Markham, K. R. (1982) *Techniques of Flavonoid Identification*, Academic Press, London, p. 113.

Marston, A. and Hostettmann, K. (2006) Separation and quantification of flavonoids, in *Flavonoids: Chemistry, Biochemistry and Applications*, eds. O. Andersen and R. Markham, CRC Press, Boca Raton, pp. 1–36.

McDermott, G. P., Noonan, L. K., Mnatsakanyan, M., *et al*. (2010) *Anal. Chim. Acta*, **675**, 76–82.

Mehrdad, M., Zebardast, M., Abedi, G., *et al*. (2009) *J. AOAC Int.*, **92**, 1035–1043.

Merken, H. H. and Beecher, G. R. (2000) *J. Agric. Food Chem.*, **48**, 577–599.

Merken, H. H. and Beecher, G. R. (2002) *ACS Symp. Series*, **803**, 21–41.

Michel, T., Destandau, E. and Elfakir, C. (2011) *J. Chromatogr. A*, **1218**, 6173–6178.

Miron, T. L., Plaza, M., Bahrim, G., *et al*. (2011) *J. Chromatogr. A*, **1218**, 4918–4927.

Moco, S., Tseng, L. H., Spraul, M., *et al*. (2006) *Chromatographia*, **64**, 503–508.

Monagas, M., Quintanilla-López, J. E., Gómez-Cordovés, C., *et al*. (2010) *J. Pharmaceut. Biomed. Anal.*, **51**, 358–372.

Monrad, J. K., Howard, L. R., King, J. W., *et al*. (2010a) *J. Agric. Food Chem.*, **58**, 2862–2868.

Monrad, J. K., Howard, L. R., King, J. W., *et al*. (2010b) *J. Agric. Food Chem.*, **58**, 4014–4021.

Montilla, E. C., Hillebrand, S., Butschbach, D., *et al*. (2010) *J. Agric. Food Chem.*, **58**, 9899–9904.

Morand, C., Manach, C., Donovan, J., *et al*. (2001) *Methods Enzymol.*, **335**, 115–121.

Morin, P., Gallois, A., Richard, H., *et al*. (1991) *J. Chromatogr.*, **586**, 171–176.

Mouls, L., Mazauric, J. P., Sommerer, N., *et al*. (2011) *Anal. Bioanal. Chem.*, **400**, 613–623.

Mustafa, A. and Turner, C. (2011) *Anal. Chim. Acta*, **703**, 8–18.

Narasimhacha, N. and Von Rudloff, E. (1962) *Can. J. Chem.*, **40**, 1123–1129.

Nardini, N., Cirillo, E., Natella, F., *et al*. (2002) *Food Chem.*, **79**, 119–124.

Nguyen, T. B., Lozach, O., Surpateanu, G., *et al*. (2012) *J. Med. Chem.*, **55**, 2811–2819.

Novakova, L. and Vlckova, H. (2009) *Anal. Chim. Acta*, **656**, 8–35.

Nuutila, A. M., Kammiovirta, K. and Oksman-Caldentey, K. M. (2002) *Food Chem.*, **76**, 519–525.

O'Mahony, J., Wei, S., Molinelli, A., *et al*. (2006) *Anal. Chem.*, **78**, 6187–6190.

Olszewska, M. A. (2012) *J. Sep. Sci.*, **35**, 2174–2183.

de Pascual-Teresa, S., Rivas-Gonzalo, J. C. and Santos-Buelga, C. (2000) *Int. J. Food Sci. Technol.*, **35**, 33–40.

Pereira, C. G. and Meireles, M. A. A. (2010) *Food Bioprocess. Technol.*, **3**, 340–372.

Perez-Jimenez, J. and Torres, J. L. (2012) *Food Res. Int.*, **49**, 798–806.

Pietta, P., Maffei-Facino, R., Carini, M., *et al*. (1994) *J. Chromatogr. A*, **661**, 121–126.

Pinelo, M., Ruiz, R. A., Sineiro, J., *et al*. (2007) *Eur. Food Res. Technol.*, **226**, 199–205.

Poole, C. F. (2003) *Trends Anal. Chem.*, **22**, 362–373.

Porter, L. J., Hrstich, L. N. and Chan, B. G. (1986) *Phytochemistry*, **25**, 223–230.

Prior, R. L., Wu, X. L. and Schaich, K. (2005) *J. Agric. Food Chem.*, **53**, 4290–4302.

Qiao, X., Yang, W. Z., Guo, D. A., *et al*. (2011) *Curr. Org. Chem.*, **15**, 2541–2566.

Rager, I., Roos, G., Schmidt, P. C., *et al*. (2002) *J. Pharm. Biomed. Anal.*, **28**, 439–446.

Rajendran, A. (2012) *J. Chromatogr. A*, **1250**, 227–249.

Rauha, J. P., Vuorela, H. and Kostiainen, R. (2001) *J. Mass Spectrom.*, **36**, 269–1280.

Rehova, L., Skerikova, V. and Jandera, P. (2004) *J. Sep. Sci.*, **27**, 1345–1359.

Reichelt, K. W., Hartmann, B., Weber, B., *et al*. (2010) *J. Agric. Food Chem.*, **58**, 1850–1859.

Repollés, C., Herrero-Martiinez, J. M. and Raflos, C. (2006) *J. Chromatogr. A*, **1131**, 51–57.

Ribereau-Gayon, P. and Stonestreet, E. (1965) *Bull. Soc. Chim.*, **9**, 2649–2652.

Rigaud, J., Escribano-Bailon, M. T., Prieur, C., *et al*. (1993) *J. Chromatogr. A*, **654**, 255–260.

de Rijke, E., Joshi, H. C., Sanderse, H. R., *et al*. (2002) *Anal. Chim. Acta*, **468**, 3–11.

de Rijke, E., Zappey, H., Ariese, F., *et al*. (2003) *J. Chromatogr. A*, **984**, 45–58.

de Rijke, E., Out, P., Niessen, W. M. A., *et al*. (2006) *J. Chromatogr. A*, **1112**, 31–63.

Rinaldo, D., Batista, J. M., Rodrigues, J., *et al*. (2010) *Chirality*, **22**, 726–733.

Robards, K. (2003) *J. Chromatogr. A*, **1000**, 657–691.

Robards, K. and Antolovich, M. (1997) *Analyst*, **122**, 11R–34R.

Robbins, R. J., Leonczak, J., Johnson, C., *et al*. (2009) *J. Chromatogr. A*, **1216**, 4831–4840.

Roussef, R. L. and Ting, S. V. (1979) *J. Chromatogr.*, **176**, 75–87.

Routray, W. and Orsat, V. (2012) *Food Bioprocess. Technol.*, **5**, 409–424.

Sakakibara, H., Honda, Y., Nakagawa, S., *et al*. (2003) *J. Agric. Food Chem.*, **51**, 571–581.

Santos-Buelga, C., and Williamson, G. eds. (2003) *Methods in polyphenol analysis*, The Royal Society of Chemistry383 pp, Cambridge, UK.

Santos-Buelga, C., Garcia-Viguera, C. and Tomas-Barberan, F. A. (2003) On-line identification of flavonoids by HPLC coupled to diode array detection, in *Methods in Polyphenol Analysis*, eds. C. Santos-Buelga and G. Williamson, The Royal Society of Chemistry, Cambridge, pp. 92–124.

Santos-Buelga, C., Gonzalez-Manzano, S., Duenas, M., *et al*. (2011) Analysis and characterization of flavonoid metabolites, in *Recent Advances in Polyphenols Research*, eds. S. Quideau and V. Cheynier, Wiley-Blackwell, Oxford, vol. **III**, pp. 249–286.

Santos-Buelga, C., Gonzalez-Manzano, S., Dueñas, M., *et al.* (2012) *Methods Mol. Biol.*, **864**, 427–464.

Schmidt, T. J., Merfort, I. and Matthiesen, U. (1993) *J. Chromatogr.*, **634**, 350–355.

Schmidt, T. J., Merfort, I. and Willuhn, G. (1994) *J. Chromatogr.*, **669**, 236–240.

Schulten, H. R., Simmleit, N. and Mueller, R. (1989) *Anal. Chem.*, **61**, 221–227.

Schulz, H., Joubert, E. and Schütze, W. (2003) *Eur. Food Res. Technol.*, **216**, 539–543.

Sengupta, P. K. and Kasha, M. (1979) *Chem. Phys. Lett.*, **68**, 382–385.

Shi, J., Li, G., Wang, H., *et al.* (2011) *Phytochem. Anal.*, **22**, 450–454.

Shi, J. Y., Zou, X. B., Zhao, J. W., *et al.* (2012) *Spectrochim Acta Part A*, **94**, 271–276.

Silva Elipe, M. V. (2003) *Anal. Chim. Acta*, **497**, 1–25.

Silva, C. L., Haesen, N. and Camara, J. S. (2012) *J. Chromatogr. A*, **1260**, 154–163.

Sinelli, N., Spinardi, A., Di Egidio, A., *et al.* (2008) *Postharvest. Biol. Technol.*, **50**, 31–36.

Søltoft, M., Christensen, J. H., Nielsen, J., *et al.* (2009) *Talanta*, **80**, 269–278.

Soukup, J. and Jandera, P. (2012) *J. Chromatogr. A*, **1245**, 98–108.

de Sousa, J. P. B., Brancalion, A. P. S., Junior, M. G., *et al.* (2012) *Nat. Prod. Comm.*, **7**, 25–28.

de Souza, L. M., Cipriani, T. R., Sant'Ana, C. F., *et al.* (2009) *J. Chromatogr. A*, **1216**, 99–105.

Stevens, J. F., Hart, H., Elema, E. T., *et al.* (1996) *Phytochemistry*, **41**, 503–512.

Sticher, O. (2008) *Nat. Prod. Rep.*, **25**, 517–554.

Stobiecki, M. (2000) *Phytochemistry*, **54**, 237–256.

Stobiecki, M. and Kachlicki, P. (2006) Isolation and identification of flavonoids, in *The Science of Flavonoids*, ed. E. Grotewold, Springer, New York, pp. 47–70.

Sutherland, I. A. and Fisher, D. (2009) *J. Chromatogr. A*, **1216**, 740–753.

Tasselli, F., Donato, L. and Drioli, E. (2008) *J. Membr. Sci.*, **320**, 167–172.

Taylor, L. T. (2010) *Anal. Chem.*, **82**, 4925–4935.

Timmers, M. and Urban, S. (2012) *Nat. Prod. Comm.*, **7**, 551–560.

Tomás-Barberán, F. A., Gil-Izquierdo, A., Ferreres, F., *et al.* (2003) Analysis and purification of flavanones, chalcones and dihydrochalcones, in *Methods in Polyphenol Analysis*, eds. C. Santos-Buelga and G. Williamson, The Royal Society of Chemistry, Cambridge, pp. 359–371.

Treutter, D. (1989) *J. Chromatogr.*, **467**, 185–193.

Treutter, D., Santos-Buelga, C., Gutmann, M., *et al.* (1994) *J. Chromatogr. A*, **667**, 290–297.

Truong, V. D., Hu, Z., Thompson, R. L., *et al.* (2012) *J. Food Compos. Anal.*, **26**, 96–103.

Vallejo, F., Tomas-Barberan, F. and García-Viguera, C. (2003) *J. Agric. Food Chem.*, **51**, 3029–3034.

Vega-Villa, K. R., Remsberg, C. M., Ohgami, Y., *et al.* (2009) *Biomed. Chromatogr.*, **23**, 638–646.

Visnevschi-Necrasov, T., Cunha, S. C., Nunes, E., *et al.* (2009) *J. Chromatogr. A*, **1216**, 3720–3724.

Vrhovsek, U., Masuero, D., Gasperotti, M., *et al.* (2012) *J. Agric. Food Chem.*, **60**, 8831–8840.

Vukics, V. and Guttman, A. (2010) *Mass Spectrom. Rev.*, **29**, 1–16.

Vukics, V., Ringer, T., Kery, A., *et al.* (2006) *J. Chromatogr. A*, **1206**, 11–20.

Waksmundzka-Hajnos, M., Wianowska, D., Oniszczuk, A., *et al.* (2008) *Acta Chromatogr.*, **20**, 475–488.

Wei, Y., Xie, Q. Q., Dong, W. T., *et al.* (2009) *J. Chromatogr. A*, **1216**, 4313–4318.

Weiss, R., Molinelli, A., Lakusch, M., *et al.* (2002) *Bioseparation*, **10**, 379–387.

Whitacre, E., Oliver, J., Van Den Broek, R., *et al.* (2003) *J. Food Sci.*, **68**, 2618–2622.

Wijngaard, H. and Brunton, N. (2009) *J. Agric. Food Chem.*, **57**, 10625–10631.

Wijngaard, H., Hossain, M. B., Rai, D. K., *et al.* (2012) *Food Res. Int.*, **46**, 505–513.

Wolfender, J. L., Ndjoko, K. and Hostettmann, K. (2001) *Phytochem. Anal.*, **12**, 2–22.

Wolfender, J. L., Ndjoko, K. and Hostettmann, K. (2003) *J. Chromatogr. A*, **1000**, 437–455.

Wollgast, J., Palalroni, L., Agazzi, M. E., *et al.* (2001) *J. Chromatogr. A*, **926**, 211–220.

Wulf, L. W. and Nagel, C. W. (1976) *J. Chromatogr.*, **116**, 271–279.

Xia, Y. Q., Guo, T. Y., Song, M. D., *et al.* (2006) *React. Funct. Polym.*, **66**, 1734–1740.

Xiao, X., Si, X., Tong, X., *et al.* (2011) *Sep. Purif. Technol.*, **81**, 265–269.

Xie, J., Zhu, L., Luo, H., *et al.* (2001) *J. Chromatogr. A*, **934**, 1–11.

Xu, L., Shi, H., Liang, T., *et al.* (2011) *J. Sep. Sci.*, **34**, 1347–1354.

Xu, P., Guan, S., Feng, R., *et al.* (2012) *Phytochem. Anal.*, **23**, 228–231.

Yanagida, A., Shoji, T. and Shibusawa, Y. (2006) *J. Chromatogr. A*, **1112**, 195–201.

Yanagida, A., Murao, H., Ohnishi, M., *et al.* (2007) *J. Chromatogr. A*, **1143**, 153–161.

Yañez, J. A., Preston, K., Andrews, P. K., *et al.* (2007) *J. Chromatogr. B*, **848**, 159–181.

Yang, Y., Wu, T., Yang, W. X., *et al.* (2008) *J. Liq. Chromatogr. Relat. Technol.*, **31**, 1523–1531.

Yang, F. Q., Zhao, J. and Li, S. P. (2010) *Electrophoresis*, **31**, 260–277.

Zhang, Y. Q., Qin, Z. and Tu, Z. Y. (2007) *Chem. Eng. Technol.*, **30**, 1014–1019.

Zhang, C., Shen, Y., Chen, J., *et al.* (2008a) *J. Agric. Food Chem.*, **56**, 8268–8272.

Zhang, H., Guo, Z., Zhang, F., *et al.* (2008b) *J. Sep. Sci.*, **31**, 1623–1627.

Zhang, Q., Zhou, M. M., Chen, P. L., *et al.* (2011) *J. Food Sci.*, **76**, C680–C685.

Zhang, C. P., Liu, G. and Hu, F. L. (2012) *Nat. Prod. Res.*, **26**, 270–273.

Zhou, X., Xiang, B., Wang, Z., *et al.* (2007) *Anal. Lett.*, **40**, 3383–3391.

Zhou, Y., Wang, Y., Wang, R., *et al.* (2008) *J. Sep. Sci.*, **21**, 2388–2394.

Coumarins – Analytical and Preparative Techniques

Krystyna Skalicka-Woźniak and Kazimierz Głowniak

Department of Pharmacognosy with Medicinal Plant Unit, Medical University of Lublin, Lublin, Poland

1 INTRODUCTION

Coumarins (2H-1-benzopyran-2-one derivatives) are, nowadays, an important group of organic compounds from natural and synthetic sources, which are useful in a number of fields. Coumarins comprise a large group of phenolic substances, which occur in plants as secondary metabolites. Most coumarins occur in higher plants, with the richest sources being the Rutaceae and Umbelliferae. Although distributed throughout all parts of the plant, coumarins are detected at the highest levels in the fruits, followed by the roots, stems, and leaves. Environmental conditions and seasonal changes have great influence over the occurrence in diverse parts of the plant (Lacy and O'Kennedy, 2004).

There are four main coumarin subtypes (Figure 1):

- the simple coumarins (hydroxylated, alkoxylated, and alkylated derivatives),
- furanocoumarins (consist of a five-membered furan ring attached to the coumarin nucleus, divided into linear or angular types – psoralen and angelicin derivatives, respectively),
- pyranocoumarins (contain a six-membered ring: xanthyletine type, seseline type, and alloxanthyletine type), and

- isocoumarins (Murray, Mendez, and Brown, 1982; Waksmundzka-Hajnos and Hawrył, 2008).

Coumarins typically appear as colorless or yellow crystalline substances and are well soluble in organic solvents (chloroform, diethyl ether, and ethyl alcohol) and in fats and fatty oils. Coumarins and its derivatives exhibit sublimation on heating to 100°C (Lozhkin and Sakanyan, 2006).

2 PHARMACOLOGICAL PROPERTIES OF COUMARINS

Coumarins are known for their different pharmacological properties. They interact with benzodiazepine receptor binding site. The pharmacological effects on activation of the benzodiazepine receptor range from sedative-hypnotic, anticonvulsant, and anxiolytic, depending on the intrinsic activity of the compounds (Dekermendjian *et al.*, 1996). Imperatorin, in a dose-dependent manner, increased the threshold for electroconvulsions in mice. The maximum anticonvulsant effect produced by imperatorin in the maximal electroshock seizure threshold test was observed at 30 min after its systemic (i.p.) administration. It resulted because of inhibition of gamma-aminobutyric acid (GABA)

Simple coumarins
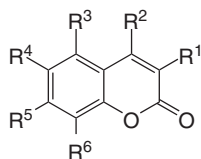
Name	R^1	R^2	R^3	R^4	R^5	R^6
Coumarin	H	H	H	H	H	H
Fraxetin	H	H	H	OCH$_3$	OH	OH
Herniarin	H	H	H	H	OCH$_3$	H
Isofraxidin	H	H	H	OCH$_3$	OH	OCH$_3$
Osthenol	H	H	H	H	OH	CH$_2$CHC(CH$_3$)$_2$
Osthol	H	H	H	H	OCH$_3$	CH$_2$CHC(CH$_3$)$_2$
Scoparone	H	H	H	OCH$_3$	OCH$_3$	H
Scopoletin	H	H	H	OCH$_3$	OH	H
Umbelliferone	H	H	H	H	OH	H

Furanocoumarins - psoralen type
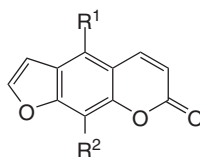
Name	R^1	R^2
Imperatorin	H	OCH$_2$CHC(CH$_3$)$_2$
Phellopterin	OCH$_3$	OCH$_2$CHC(CH$_3$)$_2$
Byakangelicol	OCH$_3$	OCH$_2$CHOC(CH$_3$)$_2$
Bergapten	OCH$_3$	H
Psoralen	H	H
Xanthotoxin	H	OCH$_3$
Isopimpinellin	OCH$_3$	OCH$_3$
Xanthotoxol	H	OH

Figure 1   Structures of the most popular coumarins.

Furanocoumarins – angelicin type
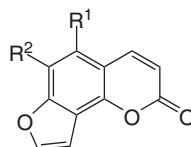
Name	R^1	R^2
Angelicin	H	H
Isobergapten	OCH_3	H
Pimpinellin	OCH_3	OCH_3

Pyranocoumarins
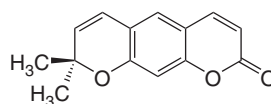
Xanthyletin
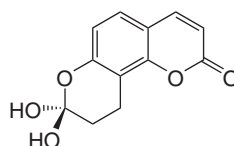
Seselin

Dihydropyranocoumarins
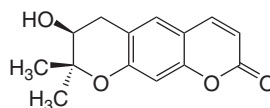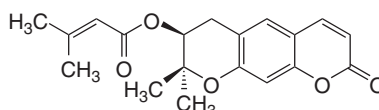
Decursinol

Decursin

Figure 1   *(continued)*

transaminase activity and being a partial agonist of GABA–benzodiazepine receptor complex (Deker-mendjian *et al.*, 1996; Choi *et al.*, 2005; Luszczki, Głowniak, and Czuczwar, 2007a). In addition, imperatorin enhanced the anticonvulsant effects of known and popular synthetic drugs (Luszczki, Głowniak, and Czuczwar, 2007b; Łuszczki *et al.*, 2008). Anticonvulsant activity was demonstrated for bergapten and simple coumarin – osthol, which inhibits L-type

Ca^{2+} channels present in neurons (Cooper and White, 2000; Tosun *et al.*, 2008; Luszczki *et al.*, 2009). Four new dihydropyranocoumarins – decursinol derivatives isolated from *Angelica gigas* roots, such as: 4″-hydroxytigloyldecursinol, 4″-hydroxydecursin, (2″S,3″S)-epoxyangeloyldecursinol, and (2″R,3″R)-epoxyangeloyldecursinol, together with decursinol and decursin, exhibited protective activities

against glutamate-induced neurotoxicity, at concentrations ranging from 0.1 to 10 μM (So *et al*., 2005).

Coumarins exhibit potent acetylcholinesterase (AChE) inhibitory activity and provide protection to neurons against β-amyloid plaques (Aβ)-induced oxidative stress and free radicals. AChE is the enzyme that degrades acetylcholine, but, furthermore, it has been shown to accelerate amyloid-β peptide formation. All that cause that AChE inhibitors are used in Alzheimer's diseases (Hostettmann *et al*., 2006; Anand, Singh, and Singh, 2012). Among five isolated coumarins from *Peucedanum ostruthium*, ostruthol was the most effective as AChE inhibitor in thin layer chromatography (TLC) bioautographic assay. Very interesting was the fact that ostruthol was about 10-fold more active than the commercial AChE inhibitor galanthamine and as strong as huperzine A, which is currently one of the most powerful natural AChE inhibitors. In addition, a chromone derivative (peucenin) was found to inhibit AChE activity (Urbain, Marston, and Hostettmann, 2005). Imperatorin, xanthotoxin, and bergapten isolated from fruits of *Angelica officinalis* displayed a high inhibition (over 80%) against butyrylcholinesterase (BChE), which is comparable with galanthamine activity (Senol *et al*., 2011). Orhan *et al*. have tested AChE and BChE inhibitory activities of seven coumarin derivatives including umbelliferone, 4-methylumbelliferone, 4-hydroxycoumarin, scopoletin, 8-methoxypsoralen, bergapten, and isobergapten, in addition to the furanocoumarin mixture obtained from *Heracleum crenatifolium*. Among them, the furanocoumarin mixture, bergapten, scopoletin, and 4-methylumbelliferone showed notable inhibition toward AChE, whereas umbelliferone, scopoletin, and 8-methoxypsoralen had a prominent anti-BChE effect. Compounds were tested by the spectrophotometric method of Ellman using an ELISA microplate reader (Orhan, Tosun, and Sener, 2008).

Coumarins are known to exert antitumor effects and can cause significant changes in the regulation of immune responses, cell growth, and differentiation. Their antiproliferative effect has been used in the treatment of malignant melanoma, leukemia, renal cell carcinoma, and prostate and breast cancers (Lacy and O'Kennedy, 2004). Coumarins isolated from *A. officinalis* fruits had the proliferation inhibition activity on HeLa cells as follows: osthol > xanthotoxol > 4-methylaesculetin > isopimpinellin > bergapten >

xanthotoxin > imperatorin > coumarin > umbelliferone > 4-hydroxycoumann (Gawron and Głowniak, 1987).

Psoralen molecules, when activated by UVA radiation, form crosslinks between adjacent strands of DNA, thus interfering with DNA and cellular replication. Although it has been assumed that this is the mechanism of action of psoralen ultraviolet A (PUVA) therapy in disorders associated with increased cell division (such as psoriasis), PUVA also has other important actions on the skin, including induction of pigmentation (treatment of vitiligo) and epidermal hyperplasia or suppression of certain components of the immune system (Conforti *et al*., 2009).

Among 21 tested coumarins with different structure, simple coumarins such as aesculetin, fraxetin, and daphnetin were found to possess significant scavenging properties and, in dose-dependent manners, were more efficient than α-tocopherol against both Fe^{2+}/ascorbic and AAPH-induced lipid peroxidation (Thuong *et al*., 2010).

One of the characteristic pharmacological properties of coumarin derivatives is antimicrobial activity. They displayed antimicrobial activity, and some of them possessed resistance-modifying activity (Smyth, Ramachandran, and Smyth, 2009). Melliou *et al*. (2005) studied the antibacterial activity of 26 pyranocoumarins and found that they have a broad spectrum of activity. Calanolides, a tetracyclic dipyranocoumarins, is active against strains of HIV-1 in early stage reproduction. Several coumarins isolated from plant extracts, that is, oxypeucedanin hydrate, oxypeucedanin, paburenol, oxypeucedanin methnolate, heraclenol, heraclenin, imperatorin, osthol, psoralen, bergapten, and heraclenin, exhibit anti-HIV IIIB activity *in vitro* in H9 lymphocytes. Suksdorfin, a dihydroseselin-type angular pyranocoumarin, suppressed viral replication and has been chosen as a lead compound for various structural modifications (Yu *et al*., 2003).

It is also well known that a range of dicoumarins have strong anticoagulant activity and that bergapten also has strong antiplatelet aggregation properties (Chen *et al*., 1996).

3 EXTRACTION OF COUMARINS

One of the most important tasks for the analytical chemist is to transfer as many of the analyte molecules as possible from the matrix without

interfering substances; therefore, finding the proper solvent, in addition to optimal extraction conditions, is the first step of analysis. Coumarins are usually isolated from plants by extraction with solvents such as ethanol, methanol, benzene, chloroform, diethyl and petroleum ethers, or their combinations. Extraction with petroleum ether provides good yields of furanocoumarins, which can be isolated in crystalline form; however, medium polar compounds are soluble in chloroform, ethyl acetate, and acetone, whereas hydroxycoumarins or glycosides prefer alcohols (Lozhkin and Sakanyan, 2006).

Härmälä *et al*., in order to find the most suitable solvent for extraction of coumarins from roots of *Angelica archangelica* (*A. officinalis*), tested 20 organic solvents. After comparing of physicochemical properties, chloroform was chosen as the most efficient one (Härmälä *et al*., 1992). Seven different extraction procedures have been used for quantifying coumarin and corresponding glycosides from *Melilotus officinalis* and furanocoumarins from *Psoralea cinerea*. In the first stage, extraction with the polar solvent such as ethanol, methanol, and water seems to be the most efficient. Water extraction gave the highest total coumarin concentration. Free coumarin could also be recovered in large quantity when the extraction was performed with water at room temperature instead of at 100°C. For furanocoumarins, the best results were obtained with boiling methanol using either Soxhlet extraction or refluxed solvent (Bourgaud, Poutaraud, and Guckert, 1994).

Different ethanol concentrations were tested to optimize condition of direct reflucence extraction of coumarins from the Chinese traditional medicine – roots of *Angelica pubescens*. The sample was extracted at 85°C, and a volume ratio of solvent to sample was 10 : 1. In order to obtain the optimal extraction condition, the relationship between solvent concentration and extraction time was studied with uniform design. The sum peak area of coumarins separated with high performance liquid chromatography (HPLC) was considered as detection index. The optimal extraction conditions were obtained when 95% solution of ethanol was used during the extraction time of 3.6 h (Guo *et al*., 2006). Extract obtained according to this procedure was further used for purification of imperatorin, bergapten, and cnidilin (Wang *et al*., 2007).

Waksmundzka-Hajnos *et al*., in order to choose the optimal conditions for the extraction method of furanocoumarins from *Pastinaca sativa* fruits,

have tested exhaustive extraction in Soxhlet apparatus, ultrasonification (USAE, ultrasound-assisted extraction) at 25 and 60°C, microwave-assisted solvent extraction (MAE) in open and closed systems, and accelerated solvent extraction [ASE, pressurized liquid extraction (PLE)]. In most cases, the yield of furanocoumarins was highest by the use of ASE method (especially more hydrophobic furanocoumarins such as bergapten, imperatorin, and phellopterin) and by ultrasonification at 60°C (mostly for more hydrophilic furanocoumarins: xanthotoxin and isopimpinellin). Both methods were better than exhaustive extraction in Soxhlet apparatus (Waksmundzka-Hajnos *et al*., 2004a). The same conditions were tested for the extraction of furanocoumarins from the fruits of *A. officinalis*. Similar results were achieved. ASE, again, was the most suitable method. By increasing the temperature from 100 to 130°C, the amount of umbelliferone and xanthotoxin extracted increased. For bergapten and isopimpinellin, the increase in the extraction yield was moderate, whereas imperatorin and phellopterin showed a decreasing tendency (Waksmundzka-Hajnos *et al*., 2004b). Pressurized MAE gives high yield of extraction of more polar furanocoumarins, but it cannot be recommended as a leaching method of furanocoumarin fraction because of the probable change of analytes during the leaching process (Waksmundzka-Hajnos *et al*., 2004a). Despite this fact, Martino *et al*. proved that short extraction time required by MAE allowed efficient extraction of coumarin without any degradation problem. Two successive irradiations within closed vessel system of 5 min each at 50°C with 50% aqueous ethanol, with a cooling step in between extraction, were chosen as the proper method for coumarins extraction. In addition, authors reported that the extraction yield with Soxhlet is not satisfactory. This is probably due to the thermal degradation of coumarins. In addition, the reproducibility of this method is unsatisfactory. In order to find the best method the kinetics of extraction depending on solvent, evaluating the yields of any compounds at different time in USAE was studied. The experiments were performed using extraction solvents such as 50% aqueous EtOH, 50% aqueous MeOH, MeOH, and boiling water. For all solvents, the extraction yield increased up to 60 min and then became fairly constant. The best results were obtained using 50% aqueous ethanol (Martino *et al*., 2006).

The stability of hydroxycoumarins such as esculetin, umbelliferone, scopoletin, 4-hydroxy-coumarin has been studied under conditions of MAE, with the same solvent (methanol) at different temperatures: 50, 75, 100, 125, 150, and 175°C. Only umbelliferone was found to be stable up to the maximum working temperature. The other coumarins can only be extracted without significant degradation at temperatures up to 125°C. In this case, the differences due to the methoxyl substituents are reflected in a reduced degradation at 175°C; that is, degradation of 9.5% for scopoletin and 57% for the esculetin, which is the compound with the largest number of hydroxyl substituents (Liazid *et al.*, 2007).

Some statistical tools were used to optimize the MAE technology for the extraction of active substances such as imperatorin and isoimperatorin from the prescription of traditional Chinese medicine containing *Angelica dahurica*. The MAE under optimal conditions (microwave power of 500 W, ethanol level of 70%, and extraction time of 27 min) can reduce extraction time, save energy, and increase extraction yields as compared with modified traditional extraction with the same extraction solvent (Liao *et al.*, 2008). The maximum extraction yield of oxypeucedanin hydrate from the roots of *A. dahurica* was developed with the same MAE method under conditions as follows: ethanol concentration of 90%, solid-to-liquid ratio of 1 : 5, and extraction time of 10 min. The plant material was first extracted with water and 30% ethanol to remove all undesired compounds (Xie *et al.*, 2010).

In order to increase the extraction yield of simple coumarins, the application of polyethylene glycol (PEG) aqueous solution as a green solvent in MAE was developed for the extraction of aesculin and aesculetin from *Cortex fraxini*, a commonly used Chinese herbal drug, which is officially listed in the Chinese Pharmacopoeia. The PEG solutions were optimized by a monofactor test, and the other conditions of MAE including the size of sample, liquid/solid ratio, extraction temperature, and extraction time were optimized. PEG–MAE, organic solvent–MAE, and conventional heating reflux extraction were evaluated. The results showed that the molecular weight of PEG had significant effects – the less the molecular weight is, the less the viscosity that PEG solution would have, which is beneficial for the mass transfer during extraction. In addition, the polarity of the PEG solutions increased as the molecular weight of PEG decreased. Therefore,

PEG-200 was chosen for further application. Among its different concentration, 60% solution was selected for further analysis. Considering time and solvent saving, the optimum experimental conditions for aesculin and aesculetin in *C. fraxini* were as follows: the size of sample was 0.30–0.15 mm, the extraction temperature was 80°C, the extraction time was 10 min, and the ratio of liquid/solid was 10 : 1. Comparing the extraction efficiency of the PEG solution with organic solvent such as methanol, it was shown that methanol gave slightly lower yields than PEG-200; therefore, this method is recommended as a simple, rapid, effective, and environmental friendly (Zhou *et al.*, 2011).

Wolski *et al.* chose petroleum ether as the best single solvent for extraction of furanocoumarin from *A. officinalis* fruits. Addition of small amounts of organic modifiers, for example, petroleum ether +30% dichloromethane, allowed decrease in the extraction time (Wolski *et al.*, 1997). Thus, petroleum ether, dichloromethane, and methanol were tested for extraction efficiency of coumarins from the fruits of *Heracleum leskowii*. As the leading method, PLE was chosen. The highest extraction yield of umbelliferone, angelicin, and xanthotoxin was achieved when methanol was used as the extraction solvent. Imperatorin, isoimperatorin, and dominant compound bergapten were also efficiently extracted by methanol but the proportion of those compounds in dichloromethane extracts was slightly higher. Petroleum ether, which was the most popular for extraction of furanocoumarins, showed the lowest extraction efficiencies. Increasing the temperature from 80 to 110°C increased the extracted amount of target compounds; isolation of more polar compounds such as umbelliferone, xanthotoxin, and angelicin with methanol was more effective at lower temperature. Therefore, those constituents are more susceptible to degradation (Skalicka-Woźniak and Głowniak, 2012). For extraction of heraclenol, heraclenin, bergapten, and psoralen from the fruits of *Heracleum candicans* pure methanol and methanol : water solutions (70 : 30, 80 : 20, and 90 : 10) were tested. The results indicate that heraclenol is found in highest level in 30% aqueous methanolic extract, whereas the maximum concentration of heraclenin and bergapten was found in the pure methanolic extract (Govindarajan *et al.*, 2007).

The ASE conditions were also established for extraction of three common *Angelica* species such

as decursin, decursinol angelate, and butylidene phthalide. *n*-Hexane efficiently extracted those three compounds compared to acetonitrile and dichloromethane. Increasing the temperature from 40 to 80°C at constant pressure drastically increased the proportions of decursin, decursinol angelate and butylidene phthalide in the volatiles, but increasing the temperature from 80 to 120°C reduced the proportions of these components. The optimum temperature was found to be 80°C. The experiment was performed at constant pressure (1500 atm), and static extraction time of 10 min was chosen (Cho *et al*., 2007).

For separation of scoparone together with two alkaloids and two bibenzyls from *Dendrobium* spp., PLE procedure was optimized. The parameters including the type of solvent, particle size, temperature, and static extraction time were studied. As the best condition, 80% methanol, 0.13–0.15 mm of particle size, temperature 140°C, and static extraction time 15 min was proposed (Xu *et al*., 2010).

Supercritical fluid extraction with carbon dioxide (SFE-CO_2) seems to be an attractive alternative offering several advantages such as speed of extraction, high mass transference, completeness of extraction, cost savings, and simplicity of analysis (Teng, Chen, and Chung, 2005). It also eliminates time-consuming process of concentration and uses no or minimal organic solvent, what makes method environmentally friendly (Skalicka-Woźniak, Widelski, and Głowniak, 2008). SFE method with modifier was effective for extraction of imperatorin, meranzin, and meranzin hydrate from the peel of *Citrus maxima*. The highest extraction efficiency was obtained at 50°C and 27.6 MPa. The presence of modifiers significantly affected the extraction efficiency, and ethanol was the best one among the tested solvent (Teng, Chen, and Chung, 2005). The same temperature was optimal for extraction of the principal bioactive constituents of *A. dahurica* such as phellopterin, isoimperatorin, imperatorin, alloimperatorin, byakangelicin, isooxypeucedanin, and pimpinellin. Other conditions were as follows: 30 MPa of pressure, 2.0 L min^{-1} of flow rate of CO_2, and 4.0 L min^{-1} of flow rate of 75% ethanol as a modifier. Before the extraction, the effects of different parameters, such as pressure, temperature, flow rate of CO_2, and the amount of modifier, were carefully studied (Chen *et al*., 2008).

An artificial neural network model was used to determine the optimum extraction conditions of coumarin from *Cuscuta reflexa*. Supercritical CO_2 with a 2% of methanol as cosolvent at 55°C and 250 bars for 150 min was described as the optimal conditions. Before that, different temperature (35–75°C), time (30–150 min), and pressure (15,160–34,450 kPa) were tested (Mitra, Barman, and Chang, 2011). However, Celeghini *et al*. demonstrated that maceration and maceration under sonication gave better than SFE extraction of coumarin from *Mikania glomerata* leaves. SF extracts contain a high level of chlorophylls. Here, addition of polar modifier (EtOH) did not present significant advantages (Celeghini, Vilegas, and Lanças, 2001).

In order to determine a suitable extraction condition for separation of isocoumarins from *Coriandrum sativum*, in a wide range with a minimum number of trials, an orthogonal test design was used where temperature, pressure, and time were considered to be three major factors for effective extraction. After 10 min of static extraction (no liquid flow), the sample was subjected to dynamic extraction by flowing gaseous carbon dioxide at a rate of 2 L min^{-1}. Experiment was performed at the temperature of 35°C, under 5 MPa for 2 h (Chen *et al*., 2009). An orthogonal test design was also used in order to determine a suitable extraction condition of psoralen and isopsoralen. The sample particle size was found to be the most important determinant of the yield. The yield of psoralen and isopsoralen significantly increased as the particle size decreased. Pressure and temperature have no significant influence on the yield of the target compounds. As the optimal condition, 26 MPa of pressure, 60°C of temperature, and 40–60 mesh of sample particle size were chosen for extraction (Wang, Lee, and Wang, 2004a; Wang *et al*., 2004b).

An interesting alternative for the extraction of coumarins from *A. dahurica* was proposed by Wang *et al*. (2011). To compare SFE with common extraction methods, solvent extraction using 95% ethanol, ultrasound extraction, and SFE were all performed. SFE showed the highest selectivity among all the three extractions, leaving the next purification easy. Liquid–liquid extraction, using two systems of petroleum ether : water and petroleum ether : 60% methanol, was performed as a further separation technique to enrich the coumarins. Combining both extraction procedures, the content of coumarins could be enriched from 0.14% in the herb to 41.68% in the obtained extract (Wang *et al*., 2011).

A method based on micelle-mediated extraction and cloud-point preconcentration was developed for the separation of hydrophobic compounds osthol and imperatorin from *Cnidium monnieri*. At the first step, the conditions for extraction osthol and imperatorin from solid herbal materials into aqueous surfactant solution were determined. Genapol X-080 at a concentration 10% was chosen as the cloud-point extraction (CPE) surfactant. Because it possesses no aromatic moiety, Genapol X-080 does not absorb above 210 nm. The amount of extracted osthol and imperatorin increases when the surfactant concentration increases from 0.5% to 10%. At higher concentration, solution becomes too sticky. Liquid/solid ratio of $200:1$ (mL g^{-1}) was used in the following experiments. Results obtained under described conditions were compared with conventional extraction solvent – methanol. The amount of the target compounds was almost two times higher. The second step was to preconcentrate target compounds by phase separation based on the cloud-point phenomenon of the surfactant. The result shows that the addition of sodium chloride facilitates the separation between the surfactant-rich phase and the aqueous phase. The extraction effect is best when the concentration of sodium chloride is 2.0 mol L^{-1} and the process was at 60°C for 30 min (Zhou, Sun, and Wang, 2008a).

4 ANALYTICAL METHODS

4.1 TLC Analysis

TLC is used not only for checking the efficiency of separation provided by column chromatography but also for further purification or for analytical and quantitative purposes. The main advantages of TLC methods are easy execution and quick response. Wide variety of literature can be found concerning TLC of coumarins using well-selected solvent systems (Tosun and Tomek, 2011).

For optimization of separation condition for preparative isolation of some coumarins from petroleum ether extract of *Libanotis dolichostyla*, Zgórka elaborated model of retention on silica and florisil layers. Mobile phases were prepared as ternary mixture containing 5–20% of polar modifiers in hexane : dichloromethane $(1:1)$. For hydrophobic coumarins (imperatorin, isoimperatorin, bergapten, xanthotoxin, osthol, and edultin), a satisfactory selectivity was achieved in system containing 20% of ethyl acetate (on silica) or butyl acetate (on florisil) in the mobile phase. Polar coumarins (umbelliferone, scopoletin, and *cis*-khellactone) were isolated with addition of 20% of methyl ethyl ketone (on florisil) or di-*n*-propyl ketone (on silica) to the mobile phase. Separation of all target compounds was achieved when dioxane was used as a polar modifier (Zgórka, 2001).

For clear differentiation of morphologically very similar umbelliferous drugs, TLC method was developed. Analysis was carried out on aluminum plates precoated with silica gel with solvent consisting of toluene : ethyl acetate : glacial acetic acid $(9:1:0.1)$. Plates were observed in UV at 254 and 366 nm and after spraying with anisaldehyde/sulfuric acid reagent. Clear discrimination by means of TLC can be done for *Angelica sinensis*, *A. dahurica*, and *A. pubescens* (Zschocke *et al.*, 1998).

Structurally, coumarins are closely related compounds possessing similar physicochemical properties, thus are difficult for separation. Often, one chromatographic system is insufficient for the isolation of pure compounds (Waksmundzka-Hajnos and Hawryl, 2008). In order to separate and identify some of the furanocoumarins present in extracts obtained from the fruits of *A. officinalis*, *Heracleum sphondylium*, and *P. sativa*, graft TLC with two distinct layers was applied. Use of two different TLC systems:

- first dimension, CN-silica with 30% acetonitrile : water (three developments); second dimension, SiO$_2$ with 35% ethyl acetate : *n*-heptane (three developments); and
- first dimension, SiO$_2$ with 35% ethyl acetate : *n*-heptane (three developments); second dimension, RP-18 with 55% methanol : water, enables complete separation (Cieśla *et al.*, 2008b).

All the systems investigated were used for quantification of selected coumarins in the extracts. The highest precision and reproducibility of results were obtained by combining SiO$_2$ with RP-18 layers (Cieśla *et al.*, 2008c).

A new procedure for analysis of a mixture of natural compounds of different polarity (a mixture of 17 coumarins and flavonoids) has been used for fingerprint construction and finding chemical similarities and differences between selected species within the

genus *Peucedanum*. This is the first time multidimensional planar chromatography has been applied for such purpose. A 10 × 10-cm glass-backed CN F$_{254}$ high performance thin layer chromatography (HPTLC) plates were used. First, the plant extracts were applied separately to the plates as spots, and plates were developed three times with careful drying each time in unsaturated horizontal DS chambers at ambient temperature in 35% ethyl acetate in *n*-heptane as mobile phase. Target compound were divided along the plates into three groups of different polarity. The plates were turned through 90°, and adsorbent layers were scraped in some places to separated fractions. The least polar fractions were developed twice with 50% (v/v) methanol in water, the medium polarity fractions were developed with 30% acetonitrile in water as mobile phase, and the most polar constituents were developed with 50% (v/v) methanol in water containing 2% acetic acid. To prevent the mobile phase from migration on adsorbent containing constituents not being chromatographed during a particular step, part of the plate, from which the adsorbent was removed, was coated with a lipophilic substance (wax). The plates were observed under UV light (254 and 366 nm) (Cieśla *et al.*, 2009).

The combination of TLC with adsorbent gradient and unidimensional multiple development technique was used for constructing of chromatographic fingerprints of furanocoumarins for distinguishing selected varieties and forms of *Heracleum* species. First, the plant extracts were applied to the silica HPTLC plates and developed with the eluent system: 35% (v/v) ethyl acetate in *n*-heptane. Then, the plates were cut into strips containing partly separated coumarin mixtures, which were connected with the second chromatographic plate (RP-18W). Coumarins were transferred to the second plate using methanol. The plates were then developed, in the second direction, using 55% (v/v) methanol in water as the mobile phase. The method, owing to its high specificity, can also be used for species authentication and quality control of coumarin containing plants (Cieśla *et al.*, 2008a).

Two-dimensional TLC on polar-bonded stationary phases (CN-silica and Diol-silica) in either nonaqueous or aqueous systems was used for separation of closely related coumarins in *A. officinalis* and *Heracelum* sp. By the use of 2D-TLC on CN-silica in systems: 30% acetonitrile : water and 35% ethyl acetate : *n*-heptane, the following coumarins can be completely separated: esculetin, fraxidin, isoscopoletin, scopoletin, umbelliferone, xanthotoxin, coumarin, xanthotoxol, byakangelicin, bergapten, and imperatorin. The identification was confirmed by comparison with UV-DAD spectra. The 2D-TLC connected with DAD-densitometry was found to be an excellent tool for the quantitative analysis of coumarins (Waksmundzka-Hajnos *et al.*, 2006).

Two-dimensional overpressured TLC (2D OPLC) method was described as a powerful tool for separation of 16 coumarins from the genus *Angelica*. For optimization of chromatographic conditions, the PRISMA model was used. The 12 solvents were tested in TLC system. The mobile phase composed of 7% ethyl acteate, 52.9% chloroform, 20% dichloromethane, and 20.1% *n*-hexane was found to give the best separation. When combination of 100% chloroform in the first direction and 30% ethyl acetate in the second direction was applied, a total baseline separation of all coumarins was achieved in 2D-TLC system (Härmälä *et al.*, 1990).

In order to standardize two commonly used herbs: *Anethum graveolens* and *Carum carvi*, an efficient HPTLC method was developed. Umbelliferone together with carvone and myristicin were chosen as biomarkers. Methanolic extracts were applied on silica gel plates and, after 30 min of saturation, were developed with the mobile phase toluene : ethyl acetate : formic acid (9.3 : 0.7 : 0.1). After that, drying plates were scanned at 331 nm for umbelliferone. The proposed method was validated according to ICH rules. The calibration plot for umbelliferone was linear with the correlation coefficient of 0.997 ± 0.016. To study the accuracy of the method, recovery studies were performed by the method of standard addition at three different levels, and the average percentage recovery was found to be 99.05% (Dhalwal, Shinde, and Mahadik, 2008).

4.2 HPLC Analysis

HPLC is, nowadays, the most widely used method for analysis of coumarin derivatives. This chromatographic technique is widely used for qualitative and quantitative analyses of coumarins.

It is worth to remember that, before HPLC analysis, easy and simple solid-phase extraction (SPE) can be performed not only for cleaning of the sample but also for preliminary separation and concentration. Several methods have been elaborated. For

the isolation of the linear furanocoumarins: xanthotoxin, bergapten, and imperatorin from methanolic extracts of fresh fruits of *L. dolichostyla*, a new SPE method was developed. The octadecyl SPE microcolumns (500 mg; 3 mL) preliminarily conditioned with methanol, followed by double-distilled water were used. Diluted methanolic (30%) extracts were passed through the conditioned octadecyl microcolumns to adsorb furanocoumarin fractions on the sorbent beds and, then, were eluted at a flow rate of 0.5 mL min^{-1} with 70% acetonitrile or 80% methanol into calibrated vials. Recovery rates were in the range of 99.6 $\pm$ 1.9 to 106.7 $\pm$ 7.1 for pure standards and 95.7 $\pm$ 2.9 to 99.5 $\pm$ 1.7 for fortified samples, depending on the solvent used. In order to establish the sorption capacity of Baker-Bond octadecyl microcolumns for the linear furanocoumarins, breakthrough volumes were determined (25 mL for xanthotoxin and bergapten and 26 mL for imperatorin). This means that the sorption capacity of the octadecyl sorbent for all compounds is approximately 40 mg g^{-1} (Zgórka and Głowniak, 1999).

Zgórka *et al*. also elaborated SPE method for simple determination of pyranocoumarin visnadin together with furanochromones khellin and visnagin in *Amni visnaga* fruits and complex pharmaceuticals. Methanolic extract (20%) of *A. visnaga* or commercial drugs were loaded onto C18 SPE columns activated as it was described above. Different solvents were tested in order to elute target compounds: 60% methanol, 40% acetonitrile, and 30% tetrahydrofuran for removing furanochromones; 80% methanol, 70% acetonitrile, and 60% tetrahydrofuran for eluting pyranocoumarins. Recoveries of visnadin obtained on octadecyl columns were as follows: 102.4 $\pm$ 4.1, 99.9 $\pm$ 5.8, 101.1 $\pm$ 4.0 for pure standard and 104.9 $\pm$ 2.4, 102.4 $\pm$ 6.0, 97.7 $\pm$ 4.7 for fortified samples when 80% methanol, 70% acetonitrile, and 60% tetrahydrofuran were used as SPE eluents, respectively (Zgórka *et al*., 1998).

In order to find optimal condition for gas chromatography (GC)-flame ionization detector (FID) analysis of psoralen, bergapten, isopimpinellin, and pimpinellin in creams and pomades used to treat vitiligo, Cardoso *et al*. tested several SPE methods. Samples were extracted either with 10 mL of chloroform or with 10 mL methanol in a sonic bath for 20 min. Two cleanup processes were performed. First, a Sep-Pak Silica (690 mg, 55–105 mm) cartridge was conditioned with 20 mL of chloroform. After loading the extract dissolved in chloroform,

the cartridge was eluted with 5 mL chloroform and then with 10 mL chloroform : methanol (90 : 10). This fraction was tested for the presence of target compounds. For similar experiments, Sep-pak C18 (360 mg, 55–105 mm) cartridge was used, previously conditioned with 20 mL of methanol. After loading the extract dissolved in methanol, the cartridge was eluted with 2 mL methanol and then with 10 mL methanol : chloroform (80 : 20). Collected fractions were directly analyzed by GC-FID. The best recoveries of the furanocoumarins in the creams analyzed were 94–97%, whereas, in the pomades, recoveries were 94–96%. The results showed that extraction with methanol afforded better chromatographic results, because chloroform also extracted many other substances, leading to lower selectivity and peaks with retention times close to those of the furanocoumarins (Cardoso, Vilegas, and Honda, 2000).

A very selective molecular imprinted polymer to 7-hydroxycoumarin was prepared and then packed into cartridges and used as a SPE sorbent. The columns were preconditioned by flushing with 3 mL methanol, 2 mL water, and 1 mL buffer solution. After loading the sample, column was washed with 2 mL of deionized water and allowed to dry for 30 min. The target substance was then eluted with 3 mL methanol (Walshe *et al*., 1997). In order to find a suitable elution solvent, a number of solvents were examined, including methanol, diethyl ether, hexane : diethyl ether (50 : 50), and ethyl acetate. Methanol gave rise to the best recoveries, with the optimal amount of 3 mL.

For HPLC analysis of umckalin cartridges of silica-C18 were conditioned with methanol (5 mL) and water (5 mL), and then, samples were introduced. The cartridge was then washed with 20% aqueous acetonitrile : water before analyte elution with methanol (Franco and de Oliveira, 2010).

Because coumarins possess distinctive chromophore groups conveniently visible under UV light, UV absorption spectra are useful for their determination. They show specific absorption bands at 274 and 311 nm, which are attributed to benzene and pyrone rings, respectively. The methyl substitutions at C-5, C-7, and C-8 lead to a bathochromic shift of the 274 nm peak but not of the 311 nm peak. The spectra of 7-oxygenated coumarins (umbelliferone) show strong absorption bands at 217 and 315–330 nm with peaks or shoulders at 240 and 255 nm. Similar are spectra of the linear or angular

dihydrofurano- and dihydropyranocoumarins. Linear furanocoumarins show four zones of absorption at 205–225, 240–255, 260–270, and 298–316 nm, whereas the angular have their maxima at 242–245 and 260–270 nm (Murray, Mendez, and Brown, 1982).

Usually a reversed phase (RP) system is the method of choice in analysis of coumarins. Isocratic mode seems to be an efficient for simple analysis. A 66 : 34 (v/v) methanol : water mixtures at a flow rate of 0.8 mL min^{-1} was used for separation of bergapten, imperatorin, and cnidilin from the roots of *A. dahurica*. A 250 × 4.6-mm; 5-μm particle, Kromasil-C18 column protected by a 5-μm particle C18 guard column was used. The compounds were detected by UV absorption at 310 nm. Calibration plots for all the coumarins had correlation coefficients close to unity. Limits of detection (LODs, S/N = 3) were <92 ng mL^{-1} and limits of quantification (LOQs, S/N = 10) were <259 ng mL^{-1}. Mean recovery of the coumarins was in the range of 96.7–101.9%, and the intraday and interday precisions, as relative standard deviation (RSD), were <2.3 and <2.9%, respectively (Wang *et al.*, 2007).

Isocratic elution with the mobile phase composed of methanol : water (75 : 25, v/v) was used for separation of osthol and xanthotoxol in 10 min on Polaris ODS C18 column (250 × 4.6 mm) at column temperature of 35°C. The flow rate was 1.0 mL min^{-1} and the effluent monitored at 254 and 322 nm by photodiode array detector (PDA and DAD) (Wei, Zhang, and Ito, 2004). Mixtures of methanol and water (55 : 45) set at a flow rate 1.0 mL min^{-1} were used for analysis of psoralen and isopsoralen on Shim-pack VP-ODS column (250 × 4.6 mm i.d.) at 247 nm and at a column temperature of 25°C. (Wang, Lee, and Wang, 2004a; Wang *et al.*, 2004b).

A Polaris ODS column (250 × 4.6 mm) at a column temperature of 35°C with the mobile phase, composed of methanol and water (60 : 40, v/v), was isocratically eluted at a flow rate of 1.0 mL min^{-1} and the effluent monitored at 254 nm by a PDA detector in order to analyze imparatorin, oxypeucedanin, and isoimperatorin (Wei *et al.*, 2009).

Very often, the mixtures of acetonitrile and water are used. For short analysis of umckalin, a Luna C18 column (Phenomenex, 150 × 4.6 mm) with 3 μm granulometry was used. The mixture of acetonitrile : water (45 : 55, v/v) was a mobile phase, pumped at 0.75 mL min^{-1}. The UV detection was performed at 330 nm. Under presented conditions, the analysis could be completed in <5 min and with a good separation of analyte peak. LOD and LOQ were 0.0098 and 0.0298 g mL^{-1}, respectively (Franco and de Oliveira, 2010). Simple elution with acetonitrile : water (6 : 4, v/v) on the Hypersil ODS column (200 × 4.6 mm i.d.; 5 μm) was used for identification of linear furanocoumarins in the fruits of *L. dolichostyla* (Zgórka and Głowniak, 1999).

Separation of scoparone was achieved within 15 min on a Diamonsil C18 column (200 × 4.6 mm i.d.; 5 μm) using the mobile phase acetonitrile : water (25 : 75, v/v). The flow rate was 1.0 mL min^{-1}, the effluent was monitored at 254 nm, and the column temperature was set at 30°C (Ma *et al.*, 2006). Coumarin was analyzed on Supelco C18 column (250 × 4 mm; 5 μm) in the isocratic mode, using acetonitrile : water (40 : 60; v/v) at a flow rate of 1 mL min^{-1}. UV detection was at 274 nm. Average standard errors for the peak areas of replicate injections were smaller than 5% showing good repeatability. The HPLC-UV detection limit was 5 μg L^{-1} (Celeghini, Vilegas, and Lanças, 2001).

Alternatively, it is possible to use some buffers or acids in small additives. Mitra *et al.* for quantification of coumarin have used C18 column (250 × 4.6 mm; 5 μm) and the gradient of mobile phase composed with potassium phosphate buffer : acetonitrile : water with pH = 3 (85 : 10 : 5, v/v/v, 50 : 45 : 5, v/v/v). The flow rate was 0.8 mL min^{-1} (Mitra, Barman, and Chang, 2011).

In order to determinate osthol in the rat plasma, the RP-HPLC method was developed. The separation was performed on a Zorbax SBC18 column (150 × 4.6 mm, 5 μm particle size; Agilent Technologies), and a gradient elution of A (0.1% phosphoric acid and 0.05% triethylamine) and B (acetonitrile) was used. The linear gradient was as follows: 10–30% B over 0–10 min, 30–90% B over 10–30 min, and then returned to 30% B at 30 min immediately. The flow rate was set at 0.8 mL min^{-1}. The detector was operated at 322 nm, and the column temperature was maintained at 25°C. Under the optimum conditions, the method was shown to be reproducible and reliable with intraday precision below 7.62%, interday precision below 6.37%, and accuracy within ±5.02% and mean extraction recovery more than 90.4%. The calibration curve for the analyte was linear in the range from 0.1 to 20 μg mL^{-1} with the correlation coefficients >0.9981. LOD (S/N = 3) was <0.03 μg

mL^{-1} and LOQ ($S/N = 10$) was <0.1 µg mL^{-1} (Zhou, Wang, and Sun, 2008b).

A µBondapack C18 analytical column (300 × 3.9 mm; 5 µm) eluted with acetonitrile, tetrahydrofuran, and water (70 : 15 : 15, v/v/v) containing 0.07% trifluoroacetic acid (TFA) was used for quatitative determination of bergapten in human blood samples. Blood samples were collected over a 2-day period before and after treatment with bergamot oil. Both fluorometric and ultraviolet detection were used. In order to establish the optimum conditions for the separation of psoralens in the essential oil, different eluants, acetonitrile : water (70 : 30, 50 : 50, 33 : 67, and 25 : 75, v/v); acetonitrile : methanol : water (20 : 60 : 20, 30 : 40 : 30, 40 : 20 : 40, and 70 : 10 : 20, v/v), and acetonitrile : tetrahydrofuran : water (25 : 25 : 50, 50 : 20 : 30, and 75 : 15 : 15 containing 0.07% TFA, v/v/v), were prepared. The acetonitrile content in the mobile phase affects the capacity factors and sensitivity. The retention time decreases and sensitivity increases of bergapten as the proportion of acetonitrile in the mobile phase is increased. Only tetrahydrofuran displayed improved sensitivity characteristics and afforded a good resolution for target compound. The fluorometric detector was operated at Ex (excitation wavelength) 325 nm and Em (emission wavelength) 470 nm. The quantification limits were 0.05 ng and was lower than detection limits normally seen with an absorbance detector (Wang and Tso, 2002).

ODS-C18 column (250 × 4.6 mm; 5 µm) was used for separation of fraxin, fraxetin, aesculin, and aesculetin in an isocratic system: methanol : 0.1% phosphoric acid (16 : 84, v/v) at a flow rate 1 mL min^{-1} and detection at 254 nm (Liu *et al.*, 2005b).

More popular is the application of gradient with methanol or acetonitrile with water. Urbain *et al.*, for successful separation of ostruthin, imperatorin, ostruthol, oxypeucedanin hydrate and peucenin from *P. ostruthium*, used a Nova-Pak C18 column (150 × 4 mm; 3.9 µm.) with MeOH-H$_2$O gradient (25 : 75 to 100 : 0) containing 0.05% TFA at a flow rate of 1.0 mL min^{-1} during 25 min, followed by 5 min with 100% MeOH. Elution was monitored at 254 nm with UV–visible detector (Urbain, Marston, and Hostettmann, 2005).

Waksmundzka-Hajnos *et al.* separated typical coumarins present in the fruits of *A. officinalis* and *P. sativa* on C18 column (150 × 4.6 mm, 5 µm) in a gradient elution as follow: 0–10 min, 45% MeOH; 10–20 min, 45–55% MeOH; 20–30 min,

55–70% MeOH; and 30–40 min, 70% MeOH in water (Waksmundzka-Hajnos *et al.*, 2004a,b). Similar compounds isolated from fruits of *C. monnieri* were analyzed with the mobile phase: methanol (solvent A)–water (solvent B) in the gradient mode as follows: 0–5 min, 60% A; 5–14 min, 60–80% A; 14–15 min, 80–60% A at the flow rate 1.0 mL min^{-1} (Li and Chen, 2005).

Isolation of series of isocoumarins was performed with a gradient elution from 50% to 80% methanol in 40 min. The effluent was monitored at 254 nm, and the flow rate was set at 1.0 mL min^{-1} constantly. The column used was a Shim-pack VP-ODS column (150 × 4.6 mm; 5 µm) (Chen *et al.*, 2009).

Teng *et al.*, for separation of imperatorin, meranzin and meranzin hydrate, have used Merck Chromolith™ RP-18 column (100 × 4.6 mm) with a mobile phase consisting of acetonitrile (solvent A) and water containing 1% (v/v) acetic acid (solvent B). The gradient program commenced at 20–80 (A : B) for 5 min, followed by a linear gradient to 28–72 in 0.5 min, maintained at 28–72 for 6.5 min, followed by a linear gradient to 30–70 in 0.5 min, maintained at 30–70 for 7.5 min, followed by a linear gradient to 32–68 in 0.1 min, and maintained at 32–68 for 40 min. The flow rate was 2 mL min^{-1}, the injected volume was 20 µL, and the coumarins were detected at 315 nm. The coumarins were detected at 315 nm (Teng, Chen, and Chung, 2005).

A monolithic column RP18 (100 × 4.6 mm) was used for separation of hydroxycoumarins together with other phenolic compounds (benzoic acids, benzoic aldehydes, cinnamic acids, catechins, stilbens, and flavonols). Two solvents were used: solvent A (10% methanol and 2% acetic acid in water) and solvent B (90% methanol and 2% acetic acid in water) at a flow rate of 5 mL min^{-1}. The gradient applied was as follows: 0 min, 100% A; 1 min, 94% A; 4 min, 94% A; 5 min, 87% A; 6 min, 60% A; 9 min, 50% A; and 14 min, 0% A. Fluorescence detection was used (excitation at 340 nm and emission at 426 nm) (Liazid *et al.*, 2007).

In order to identify inflacoumarin A, the C18, 150 × 4.6 mm i.d. column was eluted with acetonitrile–0.05% aqueous TFA at a flow rate of 0.8 mL min^{-1}. The percentage of acetonitrile in the mobile phase was programmed as follows: 20% (0 min), 40% (5 min), 50% (10 min), 50% (25 min), and 80% (35 min). Elution was monitored at 254 nm and 364 nm using PDA detection. The ultraviolet

spectra of the separated individual peaks were measured with a UV–vis spectrophotometer. Different solvents including CH_3OH, CH_3ONa, $AlCl_3$ solution, $AlCl_3$–HCl, CH_3COONa, and CH_3COONa–H_3BO_3 were used to identify their structures by shifts and strengthen of the UV spectra. For the Fourier transform infrared (FTIR) spectra, the samples were prepared as potassium bromide discs (Wang, Lee, and Wang, 2004a; Wang *et al.*, 2004b).

Because *H. candicans* has been extensively used in medicine, quantitative estimation of active furanocoumarins is essential. Thus, the validated method was elaborated. Mixtures of water : phosphoric acid (99.7 : 0.3, v/v) as solvent A and acetonitrile : water : phosphoric acid (79.7 : 20 : 0.3, v/v) as solvent B using a gradient elution in 0–5 min with 30–20% A, 5–6 min with 20–19% A, 6–7 min with 19–17% A, 7–10 min with 17–15% of A, 10–15 min with 15–12% A, 15–20 min with 12–0% A, and 20–25 min with 0–70% A was applied for separation of furanocoumarin such as bergapten, psoralen, heraclenol, and herclinin. The analysis was performed at a flow rate of 0.5 mL min^{-1}, and typical column (Merck Purospher star® RP-18, 250 × 4.6 mm i.d.; 5 µm pore size) was used. A high repeatability in the retention time was obtained with (RSD) values lower than 1.4% for both standards and extracts even at high concentration. LOD and LOQ values ranged from 1.25 to 2.92 µg mL^{-1} and from 2.96 to 4.9 µg mL^{-1}, respectively, which suggested full capacity for the quantification of each furocoumarin compound investigated. The high recovery values (close to 90%) and a high repeatability indicated a satisfactory accuracy in the proposed method. There was not much variation in the interday and intraday injections performed with the RSD being 2.78% and 1.66%, respectively (Govindarajan *et al.*, 2007).

As a useful tool for qualitative analysis, MS detectors are being applied in the analysis of coumarins. In a number of cases, LC-MS provides more accurate results because of its much lower detection threshold compared to DAD (Tosun and Tomek, 2011). Liu *et al.*, for determination of eight linear furanocoumarins including aviprin, isopimpinellin, bergapten, isooxypeucedanin, gosferol, imperatorin, phellopterin, and isoimperatorin in the rhizomes of *Angelica polymorpha*, used an Agilent Eclipse XDB-C18 analytical column (250 × 4.6 mm; 5 µm) with the mobile phase methanol : water for elution in a gradient from 5 : 95 to 95 : 5 in 75 min, followed by isocratic elution with 95 : 5 between 75

and 90 min. The flow rate was typically 1 mL min^{-1}, with the detection wavelength at 300 nm. The ionization conditions were adjusted as follows: the mass spectrometer was operated in the positive ion mode. Ultrapure nitrogen was used as nebulizer, curtain, and collision-activated dissociation (CAD) gas at 8, 11, and 4 (instrument unit), respectively. The optimized Turbo Ion-Spray voltage and temperature were set at 4000 V and 450°C, respectively. All psoralen 5-alkoxyl derivatives including aviprin, bergapten, isooxypeucedanin, gosferol, isoimperatorin, and psoralen 5,8-dialkoxyl derivatives including isopimpinellin and phellopterin showed four clear absorption maxima at about 217–224, 248–251, 265–270, and 310–315 nm. Moreover, owing to the strong effect of auxochrome 5-alkoxyl, psoralen 5-alkoxyl, and 5,8-dialkoxyl derivatives had about 8–12 nm bathochromic shift of the absorption band (310–315 nm) compared to that of psoralen and its 8-alkoxyl derivatives (302–303 nm). In addition, the ratio of maximum absorbance between the absorption band (248–251 nm) and another absorption band (265–270 nm) was quite different between psoralen 5,8-dialkoxyl derivatives and psoralen 5-alkoxyl or 8-alkoxyl derivatives. The ratio of the former was less than 1, whereas that of the latter was greater than or equal to 1. In addition, an API 3000 triple quadrupole mass spectrometer equipped with a Turbo Ion-Spray electrospray ionization (ESI) source was used for mass analysis and detection. The mass spectrometer was operated in the positive ion mode. UV and positive ESI–MS data of reference compounds are presented in Table 1 (Liu *et al.*, 2011).

Yang *et al.* used an Eclipse XDB-C18 column (150 × 4.6 mm; 5 µm) for determination of xanthotoxin, psoralen, isopimpinellin, and bergapten. The mobile phase consisted of 1 mmol ammonium acetate and methanol (30 : 70, v/v) with a flow rate of 0.8 mL min^{-1} was applied. Total eluent flow from the HPLC was diverted directly into the turbo spray source. The total analysis time was just 6 min for each run. Detection was performed using a hybrid triple quadrupole linear ion trap mass spectrometer equipped with Turbo V sources and Turbo Ion-Spray interface. The instrument was operated using ESI source in positive mode. The positive ESI could offer better sensitivity and reproducibility. In the full scan mass spectra, the protonated molecular ions [M+H]$^+$ of xanthotoxin, psoralen, isopimpinellin, and bergapten (*m/z*, 217.2, 187.1, 247.1, and 217.1, respectively) were stable and exhibited higher abundance. Under the product ion

Table 1   Characterization of the most popular coumarins by HPLC-DAD-ESI-MS (according to Liu *et al.*, 2011).

Name	UV λ_{max} (nm)	(Positive) MS ions (*m/z*) Molecular ion	Fragment
Bergapten	222.0; 250.0; 268.0; 312.6	217.2 [M+H]$^+$; 239.2 [M+Na]$^+$	202.2; 174.1
Imperatorin	217.9; 248.8; 263.5; 302.9	271.1 [M+H]$^+$; 293.0 [M+Na]$^+$	203.1; 147.2
Isoimperatorin	221.1; 250.0; 267.3; 312.8	271.2 [M+H]$^+$; 293.1 [M+Na]$^+$	203.1; 147.0
Isooxypeucedanin	220.0; 249.9; 267.3; 311.1	287.3 [M+H]$^+$; 309.4 [M+Na]$^+$	202.1; 174.1
Isopimpinellin	223.2; 248.1; 269.7; 314.1	247.3 [M+H]$^+$; 268.3 [M+Na]$^+$	232.3; 174.5
Isopsoralen	209.8; 246.1; 295.0	187.2 [M+H]$^+$; 209.4 [M+Na]$^+$	159.2; 131.0
Phellopterin	223.1; 250.5; 269.0; 313.4	301.2 [M+H]$^+$; 323.2 [M+Na]$^+$	233.2; 203.3
Psoralen	217.1; 248.8; 304.9	187.2 [M+H]$^+$; 209.4 [M+Na]$^+$	159.2; 131.0

Source: Chromatographia, Simultaneous qualitative analysis of coumarins in the rhizomes of Angelica polymorpha Maxim. by LC–DAD–ESI–MS, 73, 2011, 1121, Liu N., Yu N.J., Li J., Yang Y., Guo J.F. and Zhao Y.M., Table 1. With kind permission from SpringerScience+BusinessMedia.

scan mode, the most intensive product ions were *m/z* 202.1 from *m/z* 217.1, *m/z* 131.1 from *m/z* 187.1, *m/z* 217.0 from *m/z* 247.1, and *m/z* 202.1 from *m/z* 217. Some additional experiments were performed in order to optimize the process. It is known that electrolyte modification of mobile phase can significantly improve the ESI efficiency resulting in enhanced analyte responses; therefore, for electrolyte-free mobile phase, water : methanol (30 : 70), different buffers including acetic acid (0.01%, 0.05%, and 0.1%), ammonium acetate (0.2, 1, and 2 mmol), and formic acid (0.01%, 0.05%, and 0.1%) were added in order to identify the optimal mobile phase, which could produce the best sensitivity, efficiency, and peak shape. The addition of 0.1% acetic acid reduced about half of the response of the analytes comparing to electrolyte-free mobile phase, whereas formic acid did not produce obvious effect. The highest signal intensity was achieved when the concentration of ammonium acetate in the mobile phase reached to 1 mmol. In addition, it was noticed that lower ratio of water phase in mobile phase should be chosen to enhance signal and reduce chromatographic analysis time (Yang *et al.*, 2010).

The HPLC-DAD-UV, HPLC-ESI-MS recorded in positive and negative ion modes and the positive HPLC-APCI-MS, MS2 was developed for distinguish between "poisonous" and "nonpoisonous" chemotypes of *Ferula communis*. The method was performed on a Zorbax XDB-C8 reverse phase analytical column using a binary eluent: TFA 0.01% in water (solvent A) and TFA 0.01% in acetonitrile (solvent B), starting with 55% A for 5 min, then a linear gradient from 55% A to 10% A in 40 min. The flow rate was 1.0 mL min^{-1}; the PDA detection range was 190–400 nm. For MS detection, HCOOH (0.1%) instead of TFA (0.01%) was used. Mass

spectral analyses were performed with a Finnigan MAT LCQ ion trap mass spectrometer equipped with both ESI and APCI interface. ESI mass spectra were acquired in both positive and negative ion modes by scanning over the 100–1500 mass range. Five prenylated coumarins were identified, among which ferulenol was dominant. Its UV spectrum shows λ_{max} at 202, 282, and 307 nm, attributable to the hydroxycoumarin moiety. Both the positive ESI and APCI-MS spectra showed [M+H]$^+$ signal at *m/z* 367, whereas the [M−H]$^-$ ion at *m/z* 365 was recorded in the negative ESI-MS spectrum. The APCI MS^2 spectrum was characterized by subsequent losses of 14 amu: *m/z* 311, 297, 285, 271, 257, 243, and 231, attributable to the progressive fragmentation of an aliphatic chain. The base peak at *m/z* 175 was assigned to hydroxycoumarin core moiety (Arnoldi *et al.*, 2004).

Because acetonitrile and methanol may have different selectivities in analysis of complex matrices, thus, the mixture of acetonitrile, methanol, and water are sometimes ultimately determined as the mobile phase (Xie *et al.*, 2010). The chromatographic separation of active compounds in roots of *A. dahurica* was performed on an YMC ODS-C18 column (250 × 4.6 mm; 5 μm) with the column temperature set at 30°C. The mobile phase consisted of solutions A (acetonitrile–methanol, 5 : 3) and B (water) with a linear gradient elution at a flow rate of 1 mL min^{-1}. The elution program was as follows: 20–24% A (0–10 min); 24–40% A (10–40 min); 40–52% A (40–60 min); and 52–72% A (60–90 min). The HPLC chromatogram was monitored at 311 nm, and the UV spectra were recorded in the range 200–500 nm. In addition, the HPLC effluent was split and was introduced into the mass spectrometer. The ESI-MS spectrometer was operated in positive

mode. The electrospray capillary voltage was set to 30 V. Nitrogen was used as a drying gas for solvent evaporation. The API housing and drying gas temperatures were kept at 50°C and 380°C. The collision energy was 20 eV. The spectra were recorded in the range of m/z 80–600 for full-scan MS analysis. The scan time was 1 s, and the detector multiplier voltage was set to 1300 V. On the basis of MS fragmentation behavior, 10 coumarins were identified and classified into three types: simple coumarin (scopoletin, osthenol, and auraptenol), linear furocoumarin (xanthotoxol, oxypeucedanin hydrate, byakangelicin, imperatorin, 8-geranoxypsoralen, and 8-geranoxy-5-methoxypsoralen), and angular furocoumarin (pimpinellin) (Xie *et al*., 2010).

In addition, Liu *et al*., for separation of five coumarins: xanthotoxin, isopimpinellin, bergapten, imperatorin, and osthol, used them in gradient mode as follows: 30 : 30 : 40 to 50 : 30 : 20 in 30 minutes on SPHERIGEL ODSC18 column at room temperature and the flow rate at 1.0 mL min^{-1} (Liu *et al*., 2004a).

Even temperature higher than the standard can give promising results in HPLC analysis. The crude ethyl acetate extract of *Notopterygium forbessi* containing notopterol, isoimperatorin, and few others coumarins were analyzed by HPLC. The Shim-pack CLC-ODS column (150 × 34.6 mm) at a column temperature of 40°C was used. The mobile phase composed of methanol : acetonitrile : water (30 : 30 : 40, v/v) was isocratically eluted at a flow rate of 1.0 mL min^{-1} (Yang *et al*., 2000).

A simple RP-HPLC method coupled with a PDA detector has been developed for the separation and quantitative analysis of the main coumarins from *H. leskowii* (umbelliferone, xanthotoxin, angelicin, isopimpinellin, bergapten, imperatorin, and isoimperatorin). The 250 × 4.6-mm stainless steel column, packed with 5 μm Hypersil BDS C18, was used. The flow rate was 1 mL min^{-1}; the column temperature was 25 °C. A stepwise mobile phase gradient was prepared from methanol (A) and water (B). The gradient was 0–5 min 50–60% A; 5–25 min 60–80% A; 25–30 min isocratic 80% A; and 30–40 min 80–100% A. Detection was performed using the following wavelengths: λ = 254, 280, and 320 nm (Skalicka-Woźniak and Głowniak, 2012). The same method was also used for the successful separation of rare furanocoumarins and dihydropyranochromones from the fruits of *Peucedanum alsaticum*. Here, for the identification of isolated compounds, a HPLC-DAD/ESI-TOF-MS method was used. A Zorbax SB C18 analytical column (150 × 2.1 mm, 3.5 μm) was eluted with a mixture of two solvents system: B (90% acetonitrile in 0.1% formic acid) and A (60% acetonitrile in 0.1% formic acid), using the following gradient mode: 0–8 min, isocratic elution at 100% A, then: 8–25 min, gradient from 0% to 80% B in A, and then 25–30 min: decrease from 80% to 0% B in A. The flow rate was 0.2 mL min^{-1} (Skalicka-Woźniak *et al*., 2009a).

In order to control the content of the bioactive coumarins in formulations from genus *Angelica*, the most important genera of medicinal plants used in traditional medicinal systems of the Far East and certain Western countries, Park *et el*. elaborated simultaneous determination of five furanocoumarins: imperatorin, oxypeucedanin, byakangelicol, phellopterin, and isoimperatorin. Columns with different dimensions and with different filled were tested. Because the Develosil RPAQUEOUS C30 column showed excellent separation, performing baseline separation of each peak, this was chosen for further experiments. Various volume ratios of acetonitrile and water were used as the mobile phase. A mobile phase composed of 65 and 70% acetonitrile in water showed the best resolution of the standard solution containing the five coumarins. In order to reduce the analysis time, the mobile phase was optimized with 70% acetonitrile in water. Increased column temperatures up to 40°C decreased not only the analysis time but also the resolution of some tested compounds. Temperature of 20°C was set as a preferred one. The detection wavelength was set at 254 nm. The specific determination of the mentioned compounds was also accomplished by a triple quadrupole tandem mass spectrometer equipped with an ESI source (LC-ESI-MS/MS). Multiple reaction monitoring (MRM) in the positive mode was used to enhance the selectivity of detection (Park *et al*., 2009). All five compounds showed good linearity with correlation coefficients higher than 0.9996 and wide linear ranges (30–7000 ng mL^{-1}) under current chromatographic conditions. The LODs were 12 ng mL^{-1} of byakangelicol, 20 ng mL^{-1} of oxypeucedanin and imperatorin, 36 ng mL^{-1} of phellopterin, and 28 ng mL^{-1} of isoimperatorin. The corresponding LOQs were 30, 50, 50, 90, and 70 ng mL^{-1}, respectively, indicating the high sensitivity of the HPLC conditions used in this study. The reproducibility (RSD) of the chromatographic method established in this study was determined in intraday and interday (n = 5) periods for the five

coumarins tested. The values for intraday and inter-day precision were 0.10–5.71% and 0.04–8.74%, respectively. The recoveries (accuracies) of the five coumarins were also determined in intraday and interday ($n = 5$) periods. The values for intra-day and interday accuracy were 95.9–117.6% and 97.4–110.7%, respectively. For the LC-ESI-MS/MS analysis, the best conditions for the ionization of the five coumarins in the mass spectrometer were obtained with the ESI positive ionization mode using a mixture of 1 mM ammonium acetate and acetoni-trile (35 : 65, v/v %) containing 0.1% acetic acid as the mobile phase, 40°C column temperature, 10 µL injection volume, and 0.20 mL min^{-1} flow rate. The average run time was 5 min. All five com-pounds showed good linearity (correlation coeffi-cients higher than 0.9993) with wide linear ranges (2–150 ng mL^{-1}) under current LC-ESI-MS/MS conditions. The LODs were 0.8 ng mL^{-1} of byakan-gelicol, 2.0 ng mL^{-1} of oxypeucedanin, and 1.2 ng mL^{-1} of imperatorin, phellopterin, and isoimpera-torin. Their corresponding LOQs were 2.0, 5.0, 3.0, 3.0, and 5.0 ng mL^{-1}, respectively. The values for intraday and interday accuracy were 91.8–112.7% and 94.6–111.9%, respectively. The values for intra-day and interday precision were 1.35–8.74% and 1.26–7.38%, respectively (Park *et al.*, 2009).

The usefulness of methanol, acetonitrile, and tetrahydrofuran in RP-HPLC analysis of umbellifer-one, scopoletin, psoralen, bergapten, and xanthotoxin was assessed. The optimal elution program giving a good resolution, consisted of the solvents: A, water; B, tetrahydrofuran; C, methanol; was used in a gradient mode as follows: 5–13% B within 20 min, 13–22% B, and 5–7% C from 20 to 52 min and then 65% B for 7 min. The analysis was carried out at the temperature of 30°C with the flow rate 1.5 mL min^{-1} and detection at 300, 325, and 340 nm. During the optimization process, it was seen that mixtures of methanol and acetonitrile gave insufficient separation of some of the ana-lyzed coumarins. Involving tetrahydrofuran helps in improving selectivity (Kamiński *et al.*, 2003).

A rapid and selective ultraperformance liq-uid chromatography–tandem mass spectrometry (UPLC–MS/MS) method was developed for simulta-neous determination of three bioactive coumarins in rat plasma after administration of *Toddalia asiatica*. An ACQUITY UPLC™ BEH C18 column (50 × 2.1 mm; 1.7 µm) with an isocratic mobile phase consisting of methanol : 5 mmoL L^{-1} ammonium

acetate (65 : 35, v/v) and the flow rate set at 0.20 mL min^{-1} were used. The detection was performed on a triple quadrupole tandem mass spectrometer with ESI source in positive mode. The quantification was performed using MRM of the transitions of *m/z* 247.1→231.1 for pimpinellin, *m/z* 247.1→217.0 for isopimpinellin, and *m/z* 301.0→ 233.1 for phel-lopterin, respectively, with a scan time of 0.10 s per transition. The optimal MS parameters were as follows: capillary voltage 1.5 kV, cone voltage 40 V, source temperature 110°C, and desolvation temper-ature 350°C. Nitrogen was used as the desolvation and cone gas with a flow rate of 450 and 30 L h^{-1}, respectively. Argon was used as the collision gas at a pressure of approximately 2.81×10^{-3} mbar. The optimized collision energy for pimpinellin, isopimpinellin, and phellopterin was 23, 23, and 12 eV, respectively (Liu *et al.*, 2012). The method was linear for all analytes over investigated range with all correlation coefficients >0.9942. The lower limits of quantification (LLOQ) were 25.0 ng mL^{-1} for pimpinellin, 10.0 ng mL^{-1} for isopimpinellin, and 5.0 ng mL^{-1} for phellopterin. The intra- and interday precision (RSD%) was within 12%, and the accuracy ranged from 2.3% to 5.5% (Liu *et al.*, 2012).

Although the RP is usually the method of choice, analysis of coumarins in different *Seseli* species were performed by normal-phase HPLC. Solvents A (*n*-hexane : ethyl acetate 9 : 1) and B (ethyl acetate) was used in a gradient as follows: 0–20 min 80% A : 20% B isocratic, 20–30 min increasing the amount of solvent B, and 100% of solvent B during the last 10 min. The temperature was maintained at 40°C, the flow rate was kept at 0.8 mL min^{-1}, and the detection was set at 320 nm (Tosun, Baba, and Okuyama, 2007a). A normal-phase column Waters Spherisorb S5W (250 × 40 mm; 5 µm) was also used for the best separation and recovery results of anomalin and deltoin from the root and aerial part of *Seseli resinosum*. A mobile phase consisting of *n*-hexane and ethyl acetate (75 : 25 v/v) at a constant flow rate of 0.8 mL min^{-1} was applied (Tosun, Bahadır, and Dinc, 2007b).

In order to enhance speed, sensitivity, and res-olution, several studies have focused on UPLC development. For optimization of UPLC analysis of coumarins in fruits of *P. alsaticum* and *Peucedanum cervaria*, DryLab software was used. At the begin-ning, two linear gradients in times of 10 and 20 min were performed using UPLC Waters equipment. Retention times and area of obtained peaks were used

as input for the further gradient optimization using the DryLab. The predicted linear gradients (from 26% to 45% of acetonitrile in 10 min for *P. alsaticum* sample and from 16% to 49% in 12 min for *P. cervaria*) proposed by the software were manually segmented to give a gradient from 16% to 49% ACN in 12 min followed by gradient to 100% ACN in 15 min for both samples (Skalicka-Woźniak *et al.*, 2009b).

For identification of coumarin-type compounds, some software were applied. Owing to the lack of the most important standards of compounds presented in petroleum ether extracts of *Peucadnum*, the ChromSword software was applied. ChromSword, based on relationship between the retention parameters and descriptors such as partial molecular volume of structural fragments in water and energies of electrostatic interactions of bond dipols with water, claim to be able to predict the best initial separation condition for a defined molecular structure of compounds. For samples and owned standards linear gradients were performed: 5–100% ACN and MeOH with gradient times of 40 and 80 min. Obtained retention times together with the structures of coumarins were entered in ChromSword software and chemometric values for searching coumarins were calculated (*V* – calculated molecular volumes; *E*, ΔG – energy of electrostatic interaction of bond dipols with water). The Statistica software was used for calculation of linear correlation for examined standards. On the basis of the calculated retention times and by comparison with spectra of the standards from literature and spectra library, the presence of columbianadin, ostruthin, 8-metoxypeucedanin, oxypeucedanin, and isoimperatorin were confirmed (Skalicka-Woźniak *et al.*, 2009b).

4.3 GC-MS Analysis

Coumarins are also detectable on GC. In order to determination and qualitative analysis of imperatorin, alloimperatorin, isoimperatorin, and phellopterin as the main components of volatile fraction of *A. dahurica* obtained by SFE method, a simple GC-MS assay method has been developed. A DB-5 MS capillary column (30 m × 0.32 mm; 0.25 μm film thicknesses) was used. The inlet temperature was maintained at 280°C, and the conditions were as follows: the oven temperature was initially held at 60°C for 2 min and then programmed to 280–10°C min^{-1}

where it was held constant for 6 min. Helium was used as a carrier gas at a constant flow rate of 2.0 mL min^{-1}. The samples were analyzed by GC-MS-MS with the pulsed split (scan) or splitless (SIM) injection mode, and split ratio was 50 : 1. The source and electrodes of the quadrupole mass filter were both set to 200°C. Ionization was carried out in ESI mode at 70 eV. The calibration curves showed good linearity in the concentration ranges tested. The recoveries were higher than 85%, with RSD <10% (Chen *et al.*, 2008).

Decursin, decursinol angelate, columbianetin, seselin, lomatin, butylidene phthalide, dihydrophthalide, and other secondary volatile metabolites were extracted from *A. dahurica* and analyzed by GC-MS. A HP-5MS column (30 m × 0.25 mm × 0.25 μm) coated with 5% phenyl and 95% methylpolysiloxane was used. Helium was run as a carrier gas at a constant flow rate of 1 mL min^{-1}, and the injector and transfer line temperatures were 250°C and 300°C, respectively. The oven temperature was maintained at 50°C for 4 min and then increased to 280°C at the rate of 5°C min^{-1}, after which the temperature of the column was maintained for 10 min. A split ratio of 1 : 10 was used, and the mass spectrometer was operated in an ionizing energy of 70 eV (Cho *et al.*, 2007).

GC-FID was used for determination of psoralen, bergapten, pimpinellin, and isopimpinellin. A capillary fused silica LM-5 column (15 m × 0.2 mm; film thickness 0.5 μm) and an FID were used. Helium was used as carrier gas at a flow rate 0.8 mL min^{-1}, and the injection split ratio was 1 : 20. The injection temperature was 280°C. Column temperature was programmed from 150°C to 240°C with a linear increase of 10°C min^{-1}, then 240–280°C with a linear increase of 5°C min^{-1}, and was then held for 15 min. The detector temperature was 280°C. Target compounds were analyzed in a satisfactory time interval of <8 min. The LODs were for psoralen, 1.3 mg mL^{-1}; bergapten, 0.6 mg mL^{-1}; pimpinellin, 2.9 mg mL^{-1}; and isopimpinellin, 1.0 mg mL^{-1}. The LOQs were for psoralen, 4.3 mg mL^{-1}; bergapten, 2.0 mg mL^{-1}; pimpinellin, 9.7 mg mL^{-1}; and isopimpinellin, 3.0 mg mL^{-1}. The precision of the method was tested for both intraday and interday reproducibilities in creams and pomades and were lower than 5% (Cardoso, Vilegas, and Honda, 2000).

Ayapin (6,7-methylenedioxy-coumarin) and scopoletin (6-methoxy-7-hydroxycoumarin) were

purified from *Helianthus tuberosus* by consecutive TLC and were analyzed by GC-MS on a DB4 column (30 m × 0.32 mm). The oven temperature program was set as 170°C for 1 min, 170–200°C at 15°C min^{-1}, and 200–300°C at 2 min^{-1} (Cabello-Hurtado *et al.*, 1998).

4.4 Capillary Zone Electrophoresis

A capillary zone electrophoresis became more popular in the analysis of coumarins. A 27-cm by 50-mm (I.D.) fused-silica column, with a capillary-to-detector distance of 20 cm was used for the separation of 7-hydroxycoumarin. The capillary was prepared by rinsing with 0.1 *M* sodium hydroxide for 1 min and then with buffer solution for 1.2 min. The separation was achieved with an applied voltage of 20 kV (rise time 0.2 min) at 25°C. The resultant electropherogram was monitored at 210 nm (Walshe *et al.*, 1997).

Ketai *et al.* for the separation of active compounds: imperatorin, isoimperatorin, and scopoletin in *A. dahurica*, developed a micellar electrokinetic capillary chromatography (MECC) method. The optimum separation for these analytes was achieved using 10 m*M* borate, 20 m*M* SDS, and 10% (v/v) acetonitrile, with applied voltage of 22 kV. All experiments were performed using a 50.0-cm (42.5 cm effective length) × 75-μm i.d. fused-silica capillary (Ketai *et al.*, 2001).

Five coumarins such as xanthotoxol, osthenol, imperatorin, oxypeucedanin hydrate, and byakangelicin were analyzed by a pressurized capillary electrochromatography (pCEC) system, in which a mobile phase is driven by both pressurized flow and electroosmotic flow (EOF). The separation was optimized with respect to composition of the mobile phase, ionic strength of buffers, pH, and applied voltage. The capillary column used in this study had an i.d. of 100 μm and an o.d. of 375 μm, which was packed for about 25 cm length using the electrokinetic packing method. Baseline separation was achieved in <25 min using a mobile phase of methanol : acetonitrile : phosphate buffer (pH 4.8; 15 m*M*) (22.5 : 15 : 62.5, v/v/v), under conditions of 1 kV applied voltage, 20°C column temperature, 216 nm UV detection, 50 μL min^{-1} flow rate, 13.8 MPa backpressure, and 4 nL injection. The recovery of the method was determined with the standard addition method for the five coumarins in the three sample solutions, with results of 96.3–103.6% for xanthotoxol, 95.4–104.3% for osthenol, 94.8–103.8% for imperatorin, 96.5–99.6% for oxypeucedanin hydrate, and 96.5–104.7% for byakangelicin, respectively. The method was compared to capillary HPLC and conventional HPLC and gave higher column efficiency and shorter analysis time (Chen *et al.*, 2006) (Figure 2).

A rapid CEC method with poly(butyl methacrylate-co-ethylene dimethacrylate-co-[2-(methacryloyloxy) ethyl] trimethylammonium chloride) monolithic column has been developed for separation and determination of four coumarins (isopimpinellin, bergapten, imperatorin, and osthol) in extracts from the fruits of *C. monnieri*. Under the same conditions (50% acetonitrile and 50% of a 10-m*M* sodium dihydrogen phosphate electrolyte at pH 4.95), in contrast to 25 min of analysis time in HPLC and 10 min of analysis time in pCEC, a fast separation of these analytes was achieved in <5 min in CEC. The proposed rapid, sensitive, and simple method was validated. The mixture of coumarins was analyzed six times in 1 day and for 6 days in order to obtain the intraday and interday repeatability of relative migration times and relative peak areas. The RSD of relative migration times and relative peak area with intraday were in the range of 0.16–0.87 and 1.66–3.81, respectively. The RSDs of relative migration times and relative peak area with interday were in the range of 0.83–1.69 and 2.45–4.63, respectively. Recoveries for the four coumarins were between 93.91% and 98.65% (Wang *et al.*, 2010).

5 PREPARATIVE TECHNIQUES

Coumarins can be purified from accompanying substances by means of saponification, which is based on the ability of the lactone (α-pyrone) ring to open under the action of alkalis with the formation of coumarinates and to close again on subsequent acidification. A significant disadvantage of this method is the possible formation of secondary products and the dehydration and isomerization of some hydroxycoumarins. Very often, crystallization can be used, but, because many coumarins possess close solubilities in organic solvents, even multiple recrystallizations from solution did not provide reliable results. For this reason, subsequent progress in the chemistry of coumarins led to the development of various
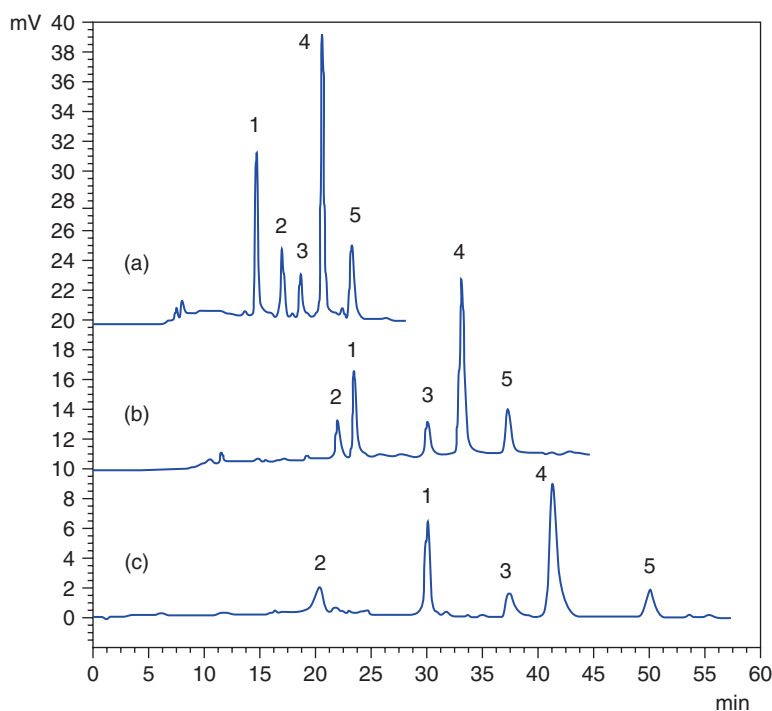
Figure 2 Separation of coumarins with (a) pCEC, (b) capillary HPLC, and (c) HPLC. Conditions for pCEC: mobile phase, methanol–acetonitrile–phosphate buffer (pH 4.8; 15 mM) (22.5 : 15 : 62.5, v/v/v); flow rate, 50 μL min^{-1}; voltage, 1 kV; backpressure, 13.8MPa; injection, 4 nL, detection, 216 nm; temperature, 20°C; and sample, standard coumarins solution. Conditions for capillary HPLC: voltage, 0 kV; other conditions as in pCEC. Conditions for HPLC: column, Ultrasphere ODS C18 (4.6 × 15 cm, 5 μm) (Beckman Coulter); mobile phase, methanol–acetonitrile–phosphate buffer (pH 4.8; 15 mM) (22.5 : 15 : 62.5, v/v/v); flow rate, 0.3 mL min^{-1}; injection, 20 μL; detection, 216 nm; temperature, 20°C. Peaks: (1) xanthotoxol, (2) osthenol, (3) imperatorin, (4) oxypeucedanin hydrate, and (5) byakangelicin. *Source*: Y. Chen *et al*. (2006). Reproduced from Elsevier.

chromatographic techniques, which are free of disadvantages inherent in the early methods (Lozhkin and Sakanyan, 2006).

5.1 Column Chromatography

Very often, the column chromatography with various sorbents and solvent systems is used at the first stage of partition. Usually, silica gel is the first choice for this purpose.

Bergapten, imperatorin, and cnidilin were isolated from 95% ethanolic extracts from the roots of *A. dahurica*. The extract after evaporation was suspended in water and partitioned three times with chloroform. The chloroform layer was subjected to silica gel column chromatography (200–300 mesh) and eluted stepwise with

petroleum ether : acetone mixtures. The fraction eluted by petroleum ether : acetone 100 : 5 (v/v) was subjected to open column ODS chromatography with methanol : water mobile phases to give three components: bergapten eluted with 45 : 55 (v/v) methanol : water, imperatorin eluted with 48 : 52 (v/v) methanol : water, and cnidilin eluted with 50 : 50 (v/v) methanol : water (Wang *et al*., 2007).

The roots of *A. dahurica* was extracted with 80% aqueous methanol, evaporated under reduced pressure, and partitioned between ethyl acetate and water. The ethyl acetate extracts obtained by removal of the solvent were chromatographed on a silica gel column with the gradient elution of *n*-hexane : ethyl acetate with increasing proportions of the ethyl acetate; 11 subfractions were collected. Silica gel column chromatography of subfraction 5 eluting with *n*-hexane : chloroform : ethanol

(20 : 20 : 1) afforded isoimperatorin. Subfraction 7 was subjected to silica gel column and eluted with *n*-hexane : chloroform : ethanol (30 : 10 : 1) to give four subfractions. The third one was chromatographed on a silica gel column with *n*-hexane : chloroform : ethanol (28 : 12 : 1) to yield oxypeucedanin hydrate-3"-butyl ether. Imperatorin and cnidilin were obtained by silica gel column chromatography of subfraction 8 using *n*-hexane : chloroform : ethanol (25 : 20 : 1) as eluting solvent, whereas oxypeucedanin hydrate was collected when subfraction 11 was eluted on silica gel column with *n*-hexane : ethyl actetate : acetone (4 : 6 : 1) (Baek *et al*., 2000).

In order to isolate umckalin (7-hydroxy-5,6-dimethoxycoumarin), a *Pelargonium sidoides* tincture was extracted with diethyl ether (3 × 1000 mL). After evaporation of the solvent, the residue was fractionated on a silica column. Elution was made with mixtures of hexane and diethyl ether of increasing polarity. The fractions containing umckalin were combined, the solvent evaporated, and the residue recrystallized from diethyl ether (Franco and de Oliveira, 2010).

In order to isolate osthol, Zhou *et al*. extracted the fruits of *C. monnieri* with hot methanol. The dry residue was partitioned between water and chloroform, ethyl acetate, and *n*-butanol in sequence. The chloroform extract was column chromatographed on silica gel and eluted with gradients of *n*-hexane : ethyl acetate (30 : 1–1 : 1). Second fraction eluted with *n*-hexane : ethyl acetate (20 : 1) was crystallized with methanol to give a colorless crystal of osthol (Zhou, Wang, and Sun, 2008b).

Osthol and corymbocoumarin were isolated from underground parts of *Seseli gummiferum*. The *n*-hexane extract was subjected to column chromatography over silica gel and was eluted with the mixture of *n*-hexane : ethyl acetate 9 : 1 (Tosun, Baba, and Okuyama, 2007a; Tosun, Bahadır, and Dinc, 2007b).

Petroleum ether extract from the roots of *Heracleum sibiricum* was chromatographed on a column filled with silica gel 60 (with 230–400 mesh granules) suspension in hexane. First, column was eluted with the mixture of *n*-hexane : ethyl acetate with increasing polarity (10–40% ethyl acetate). The fractions rich in coumarins were separated in second column with the solvent system: *n*-hexane : chloroform : ethyl acetate (1 : 1 : 5%).

Bergapten, isopimpinellin, pimpinellin, and sphondin were isolated (Bogucka-Kocka and Krzaczek, 2003).

Sediment of coumarins precipitated from petroleum ether extract from *Peucedanum tauricum* was chromatographed on silica gel 60 (230–400 mesh; E. Merck) and eluted with gradient of ethyl acetate in dichloromethane 0–20%. The fractions richest in coumarins were developed on preparative TLC plates coated with 0.5 mm layers of silica gel 60 (E. Merck) with cyclohexane : ethyl acetate 80 : 20. Separated bands were rechromatographed again on silica TLC plates with more selective system, and imperatorin and bergapten were isolated (Głowniak *et al*., 2002). Methanolic extract was also investigated. Coumarins-containing fractions were separated from ballast compounds by separation on silanized silica gel 60 RP-2 (Merck) with 60% aqueous methanol as mobile phase. Selected bands were scraped, extracted with methanol : acetone 1 : 1, and rechromatographed by normal phase TLC with the mixture of dichloromethane : acetonitrile, 99 : 1 and 7.5 : 2.5. In this way, bergapten and scopoletin were isolated (Bartnik *et al*., 2005).

Very often, instead of using silica gel, florisil (an activated magnesium silicate) is applied. Florisil had definite advantages for the separation of lipid classes by column chromatography; therefore, it is useful especially when coumarins are isolated from unpolar extracts obtained from fruits. It was successfully applied for separation of bergapten, xanthotoxin, isopimpinellin, imperatorin, pimpinellin, and, for the first time, limettin (5,7-dimethoxycoumarin) from fruits of *Heracleum mantegazzianum*. The methodology was as follows: sediment obtained from petroleum ether extract was chromatographed on a florisil column (60–100 mesh; Fluka) and eluted with gradient of dichloromethane : ethyl acetate 0–20%. First four fractions richest in coumarins were collected together and again chromatographed on florisil (60–100 mesh) with the same mobile phases. First two fractions were subjected to column filled with silica gel 60 (230–400 mesh) with a gradient of dichloromethane : ethyl acetate 0–10%. Obtained fractions were separated by preparative TLC on silica gel using benzene : ethyl acetate 9 : 1, *n*-heptane : dichloromethane : ethyl acetate 4 : 4 : 2, dichloromethane : acetonitrile 97.5 : 2.5 (Głowniak *et al*., 2000).

A series of esculin, umbelliferone, bergapten, xanthotoxin, isoimperatorin, imperatorin, psoralen, cis-khellactone, pteryxin, and epoxypteryxin were isolated from petroleum ether extract from *Peucedanum verticillare*. Precipitated sediment of coumarins was chromatographed on florisil (60–100 mesh; Fluka) and silica gel 60 (230–400 mesh; Merck) in gradient mode, followed with preparative TLC (Kozyra *et al.*, 2005).

Instead of typical adsorbents such as silica gel or florisil, adsorption by macroporous resins is considered to be superior to conventional liquid–liquid or solid–liquid extraction. It is due to their inherent characteristics such as convenience, low operational cost, lower solvent consumption, and the absence of chemical residues in the product (Li and Chase, 2010). Pulverized dried roots of *A. dahurica* was extracted by reflux with 70% alcohol twice, filtered, evaporated to dryness, and redissolved in water. The solution was then chromatographed on 1300 macroporous resin by eluting with water to remove unwanted compounds. Then, the target compounds were stepwise eluted with 10%, 30%, 50%, and 70% ethanol. At the end, 95% ethanol was used to activate the resin for another use. Xanthotoxol, osthenol, imperatorin, oxypeucedanin hydrate, and byakangelicin were eluted with 30–70% ethanol (Chen *et al.*, 2006).

A HP20 hygro-macroporous resin (Mitsubishi Chemical Corp., Tokyo, Japan) was used for isolation of eight coumarins from *Angelica polymorpha*. After suspension of an 80% alcohol extract on the sample of resin, it was separated on glass column packed with 53 mL HP20 hygro-macroporous resin and was eluted by water and 40%, 60%, 80%, and 95% ethanol in turn (Liu *et al.*, 2011).

5.2 Preparative TLC

Preparative TLC has been effectively used for further separation and purification of coumarins. Chloroform extract of rhizomes of *P. ostruthium* and petroleum ether extract from roots of *Peucedanum officinale* were subjected to gradient TLC separation on silica gel plates. Eleven development steps were performed with six different eluents. First, mixtures of chloroform and ethyl acetate with increasing polarity were used, and after sixth development, methanol was also included to the mobile phases. Imperatorin, isoimperatorin, oxypeucedanin, oxypeucedanin hydrate, and

ostruthol have been isolated and identified from *P. ostruthium* and peucedanin and oxypeucedanin from *P. officinalae* (Cisowski *et al.*, 1991).

Coumarin glucosides: esculin and fraxetin, were isolated from stems and bark of *Aesculus hippocastanum* by preparative TLC. Methanolic extract was developed on silica gel plates with *n*-butanol : acetic acid : water (4 : 1 : 2). The spots of target compounds were fluorescent in UV light (365 nm). The bands were scrapped and extracted with methanol, and spectrophotometric determination of esculin was done at 335 nm. Densitometric determination was performed with a mixture of methanol, ethyl acetate, water, and acetic acid (1 : 9 : 1 : 9), in order to decrease the viscosity of the eluent (Matysik *et al.*, 1994).

For separation of a complex, multicomponent mixtures, which are usually difficult for separation in a single chromatographic run, a combining HPLC and a TLC system was developed. HPLC optimized by the use of DryLab was performed on RP-18 column in 60% of methanol in water. Fraction from the RP column was developed in mobile phase consisting of 40% ethyl acetate in dichloromethane : heptanes (1 : 1). In this way, 12 coumarins were isolated from roots of *A. officinalis* (Hawrył, Soczewiński, and Dzido, 2000).

5.3 Prep-HPLC

For isolation of active compounds from the roots of *A. dahurica* (xanthotoxol, osthenol, oxypeucedanin hydrate, byakangelicin, and imperatorin), preparative HPLC was applied. The mobile phase consisted of a mixture of A (acetonitrile : methanol, 5 : 3) and B (water), with elution starting from 32% A, maintained for 48 min, and increased to 64% from 48 to 70 min. Optimized at an analytical scale mobile phase conditions were fine-tuned in semiprep. An YMC ODS-C18 column (250 × 4.6 mm i.d.; 5 μm) was used for analytical scale separation with a 20-μL injection volume, and the same column with different dimensions (250 mm × 10 mm i.d.; 5 μm) was used for semipreparative scale separation with a 1.5-mL injection volume. Different flow rates were used and were 0.75 mL min^{-1} for analytical and 3.5mL min^{-1} for the preparative purposes, respectively (Xie *et al.*, 2010). They also used 2D preparative separation. The first dimension of the 2D HPLC was operated.

First, Lichrosper C18 column (100 × 10 mm i.d.; 5 μm) was eluted with a mixture of 32% A (acetonitrile : methanol, 5 : 3) and 68% B (water) in the first 11 min and then increased to 64% A until 15 min at a flow rate of 10 mL min⁻¹. The second dimension using the mixture of 32% A and 68% B was performed with an YMC ODS-C18 column (250 × 10 mm i.d.; 5 μm) at a flow rate of 3.5 mL min⁻¹ (Xie et al., 2010).

Bioguided fractionation of the roots of *Citrus sinensis* led to the isolation and identification of five coumarins, namely, clausarin, suberosin, poncitrin, xanthyletin, and thamnosmonin. First, the dichloromethane extract was fractionated by column chromatography on silica gel 60, eluting with a hexane : ethyl acetate gradient. Among 18 fractions obtained, fraction 5 was applied to a MPLC column using Lichroprep 60 RP-18 (40–63 m) and eluted with methanol : water gradient. Further purification of obtained subfractions by preparative TLC (toluene : ethyl acetate, 80 : 20) led to separation of suberosin, clausarin, poncitrin, and xanthyletin (Bayet *et al*., 2007).

After separation of the crude extract of *Peucedanum praeruptorum* on high speed countercurrent chromatography (HSCCC), fraction containing *d*-laserpitin and qianhucoumarin J were purified using prep-HPLC. A CLC-ODS column (250 × 20 mm i.d.; 5 μm) was used to elute with methanol : water (78 : 22, v/v) and methanol : water (70 : 30, v/v), respectively. The injection was made through 900 μL loop.

5.4 High Speed Countercurrent Chromatography

Crude plant extracts and semipure fractions can be separated, and pure compounds can be isolated using CCC – an excellent alternative to traditional solid-phase adsorbents. The CCC technique uses a *two-phase solvent system* made of a pair of mutually immiscible solvents, one used as the *stationary phase* and the other as the *mobile phase*. The column, made of a length of tubing wound around a bobbin, rotates around its own axis and around a central axis creating a cardioid motion. The rotation of the column generates a force field. When the column is operating, a series of simultaneous mixing and settling zones along the length of the tubing appears. The

method benefits from a number of advantages when compared with the more traditional liquid–solid separation methods. CCC is proving a versatile, high resolution separation process that gives 100% sample recovery, has no nonspecific adsorption to a solid support, and has the potential to be scaled up. With these advantages, CCC is gaining popularity as a separation method for natural products, and especially in the bioassay-guided fractionation of natural products (Sutherland *et al*., 2000; Marston and Hostettmann, 2006; Guzlek *et al*., 2010).

The selection of the two-phase solvent system for the target compound(s) is the most important step in HSCCC (Ito, 2005). Searching the proper solvent system for the successful separation of particular compounds from a complex sample mixture can be very time consuming; therefore, consulting with the literature can be very helpful (Ito, 2005). In the literature, one can find a lot of examples where CCC methods were applied for separation of pure coumarins. For many years, runs have taken many hours; but, introducing HPCCC led to fast and efficient separation.

The examples of application of CCC in separation of coumarins from plant extracts are presented in the Table 2.

6 CONCLUSION

Coumarin and coumarin-related compounds have proved for many years to possess significant therapeutic potential. They have been used for their anti-inflammatory, analgesic, spasmolytic, antibacterial, antifungal, antioxidant, and anti-cancer properties. Some of the furanocoumarins are potent photosensitizers when activated by near-UV light. Hence, a good selection of extraction or other preparative-scale separation methods is very essential. In this chapter, different analytical and preparative techniques have been presented. Many efforts have been done for ages for establishing optimal analytical methods, with application of either TLC or HPLC. One can find a review of used either stationary or mobile phase, criteria of evaluation of selective analytical method, and searching a novel or not well-known techniques suitable for analysis of coumarins. Nowadays, the great development of purification experiments are conducted (especially HSCCC) in order to isolate of bioactive principals.

Table 2 Application of high speed countercurrent chromatography in separation of coumarins from plant extracts.

Plant material	Isolated compounds	Solvent system	References
Angelica dahurica	Imperatorin; oxypeucedanin; isoimperatorin	Stepwise elution, *n*-hexane : methanol : water: (a) 5 : 5 : 5; (b) 5 : 7 : 3; *Stationary phase – upper*	Liu, Li, and Sun (2004d)
Angelica dahurica	Imperatorin; oxypeucedanin; isoimperatorin	First stage *n*-hexane : ethyl acetate : methanol : water (5 : 5 : 5 : 5), second stage (5 : 5 : 4 : 6); *Stationary phase – upper*	Wei *et al.* (2009)
Angelica dahurica	Imperatorin; oxypeucedanin; isoimperatorin	Multidimensional countercurrent chromatography *n*-hexane : ethyl acetate : ethanol : water: 1 : 1 : 1 : 1 and 5 : 5 : 4.5 : 5.5; *Stationary phase – upper*	Wei and Ito (2006)
Artemisia dalailamae	Isofraxidin; scopoletin	Chloroform : methanol : water (2 : 1 : 1); *Stationary phase – upper*	Yang, Ou, and Yu (1995)
Artemisia scoparia	Scoparone	*n*-hexane : ethyl acetate : methanol : water (1 : 1 : 0.45 : 1.55); *Stationary phase – upper*	Ma *et al.* (2006)
Citrus sinensis grafted on *Citrus limonia*	Xanthyletin; seselin; xanthoxyletin	Hexane : ethanol : acetonitrile : water (10 : 8 : 1 : 1); *Stationary phase – lower*	Cazal *et al.* (2009)
Cnidium monnieri	Bergapten; imperatorin	*n*-Hexane : ethyl acetate : ethanol : water (5 : 5 : 5 : 5), increasing the flow rate of the mobile phase stepwise from 1.0 to 2.0 mL min^{-1} after 180 min; *Stationary phase – upper*	Li and Chen (2004)
Cnidium monnieri	Xanthotoxin; isopimpinellin; bergapten; imperatorin; osthol	Stepwise elution, light petroleum : ethyl acetate : methanol : water: (a) 5 : 5 : 5 : 5, 0–100 min; (b) 5 : 5 : 6 : 4, 100–250 min; (c) 5 : 5 : 6.5 : 3.5; *Stationary phase: upper*	Liu *et al.* (2004a)
Cnidium monnieri	Osthol; xanthotoxol	Stepwise elution, *n*-hexane : ethyl acetate : methanol : water: (a) 1 : 1 : 1 : 1; (b) 5 : 5 : 6 : 4; *Stationary phase – upper*	Wei, Zhang, and Ito (2004)
Cnidium monnieri	Xanthotoxin; isopimpinellin; bergapten; imperatorin; osthol	Stepwise elution, *n*-hexane : ethyl acetate : ethanol : water: (a) 5 : 5 : 4 : 6; (b) 5 : 5 : 6 : 4; *Stationary phase – upper*	Li and Chen (2005)
Coriandrum sativum	Coriandrone B; coriandrin; dihydrocoriandrin; coriandrone A	*n*-Hexane : ethyl acetate : methanol : water (3 : 7 : 5 : 5); *Stationary phase – upper*	Chen *et al.* (2009)
Diplolophium buchanani	Oxypeucedanin	*n*-Hexane : ethyl acetate : methanol : water (10 : 5 : 5 : 1); *Stationary phase – lower*	Marston and Hostettmann (1995)
Diplolophium buchanani	Oxypeucedanin hydrate	*n*-Hexane : ethyl acetate : methanol : water (10 : 5 : 5 : 1), next for one of the fraction chlorophorm : methanol : ethyl acetate : water (5 : 6 : 3 : 4); *Stationary phase – lower*	Marston and Hostettmann (1995)
Edgeworthia chrysantha	Umbelliferone; daphnoretin	*n*-Hexane : ethyl acetate : methanol : water (4 : 6 : 4 : 6); *Stationary phase – upper*	Yan *et al.* (2006)
Fraxinus officinalis	Fraxin; aesculin; fraxetin; aesculetin	*n*-Butanol : methanol : 0.5% acetic acid (5 : 1 : 5.5); *Stationary phase – upper*	Liu *et al.* (2005b)
Glycyrrhiza inflata	Inflacoumarin A	*n*-Hexane : chloroform : methanol : water (5 : 6 : 3 : 2) and rechromatography (3 : 12 : 6 : 4); *Stationary phase – upper*	Wang, Lee, and Wang (2004a)

(Continued Overleaf)

Table 2   (*Continued*)

Plant material	Isolated compounds	Solvent system	References
Notopterygium forbessi	Notopterol; isoimperatorin	Stepwise elution, light petroleum : ethyl acetate : methanol : water: (a) 5 : 5 : 4.8 : 5; (b) 5 : 5 : 5 : 4; *Stationary phase – upper*	Yang *et al.* (2000)
Peucedanum alsaticum	Alsaticol; alsaticocoumarin A; divaricatol; notoptol	*n*-Hexane : ethyl acetate : methanol : water (1 : 1 : 1 : 1); *Stationary phase – upper*	Skalicka-Woźniak *et al.* (2009a, 2012)
Peucedanum alsaticum	Ledebouriellol	*n*-Hexane : ethyl acetate : methanol : water (1 : 1 : 1 : 1); *Stationary phase – lower*	Skalicka-Woźniak *et al.* (2012)
Peucedanum decursivum	Nodakentin; Pd-C-IV; Pd-D-V; ostruthin; decursidin; decursitin C	Light petroleum : ethyl acetate : methanol : water (5 : 5 : 7 : 4); *Stationary phase – upper*	Liu *et al.* (2005a)
Peucedanum ostruthium	Ostruthin; imperatorin; ostruthol; oxypeucedanin hydrate	*n*-Hexane : ethyl acetate : methanol : water (10 : 5 : 5 : 1), then selected fractions *n*-hexane : acetonitrile : dichloromethane (10 : 7 : 3) and *n*-hexane : *t*-butyl methyl ether : acetonitrile (5 : 2 : 4); *Stationary phase – upper*	Urbain, Marston, and Hostettmann (2005)
Peucedanum praeruptorum	Qianhucoumarin D; praeruptorin A; praeruptorin C; Pd-Ib	Stepwise elution, light petroleum : ethyl acetate : methanol : water: (a) 5 : 5 : 5 : 5; (b) 5 : 5 : 6.5 : 3.5; *Stationary phase – upper*	Liu *et al.* (2004b)
Peucedanum praeruptorum	Qianhucoumarin D; Pd-Ib; (+)-praeruptorin A; (+)-praeruptorin B; peucedanocoumarin I; peucedanocoumarin II; praeruptorin E; d-laserpitin; qianhucoumarin J	Light petroleum : ethyl acetate : methanol : water (5 : 5 : 6 : 4); *Stationary phase – upper*	Hou *et al.* (2009, 2010)
Psoralea corylifolia	Psoralen; isopsoralen	*n*-Hexane : ethyl acetate : methanol : water (1 : 0.7 : 1 : 0.8); *Stationary phase – upper*	Wang *et al.* (2004b)
Psoralea corylifolia	Psoralen; isopsoralen	*n*-Hexane : ethyl acetate : methanol : water (5 : 5 : 4.5 : 5.5); *Stationary phase – upper*	Liu *et al.* (2004c)
Psoralea corylifolia	Psoralen; isopsoralen; psoralidin; corylifol A; bavachinin	*n*-Hexane : ethyl acetate : methanol : water (1 : 1.1 : 1.3 : 1); *Stationary phase - upper*	Xiao *et al.* (2010)

REFERENCES

Anand, P., Singh, B., and Singh, N. (2012) *Bioorg. Med. Chem.*, **20**, 1175–1180.

Arnoldi, L., Ballero, M., Fuzzati, N., *et al.* (2004) *Fitoterapia*, **75**, 342–354.

Baek, N. I., Ahn, E. M., Kim, H. Y., *et al.* (2000) *Arch. Pharm. Res.*, **23**, 467–470.

Bartnik, M., Głowniak, K., Macią g, A., *et al.* (2005) *J. Planar Chromatogr.*, **18**, 244–248.

Bayet, C., Faziol, C., Darbour, N., *et al.* (2007) *Phytother. Res.*, **21**, 386–390.

Bogucka-Kocka, A., and Krzaczek, T. (2003) *Acta Polon. Pharm.*, **60**, 401–403.

Bourgaud, F., Poutaraud, A., and Guckert, A. (1994) *Phytochem. Anal.*, **5**, 127–132.

Cabello-Hurtado, F., Durst, F., Jorrín, J. V., *et al.* (1998) *Phytochemistry*, **49**, 1029–1036.

Cardoso, C. A. L., Vilegas, W., and Honda, N. K. (2000) *J. Pharm. Biomed. Anal.*, **22**, 203–214.

Cazal, C. M., VdC, D., Batalhão, J. R., *et al.* (2009) *J. Chromatogr. A*, **1216**, 4307–4312.

Celeghini, R. M. S., Vilegas, J. H. Y., and Lanças, F. M. (2001) *J. Braz. Chem. Soc.*, **12**, 706–709.

Chen, I. S., Chang, C. T., Sheen, W. S., *et al.* (1996) *Phytochemistry*, **41**, 525–530.

Chen, Y., Fan, G., Chen, B., *et al*. (2006) *J. Pharm. Biomed. Anal.*, **41**, 105–116.

Chen, Q., Li, P., He, J., *et al*. (2008) *J. Sep. Sci.*, **31**, 3218–3224.

Chen, Q., Yao, S., Huang, X., *et al*. (2009) *Food Chem.*, **117**, 504–508.

Cho, S. K., Abd El-Aty, A. M., Choi, J. H., *et al*. (2007) *J. Pharm. Biomed. Anal.*, **44**, 1154–1158.

Choi, S. Y., Ahn, E. M., Song, M. C., *et al*. (2005) *Phytother. Res.*, **19**, 839–845.

Cieśla, Ł., Bogucka-Kocka, A., Hajnos, M., *et al*. (2008a) *J. Chromatogr. A*, **1207**, 160–168.

Cieśla, Ł., Petruczynik, A., Hajnos, M., *et al*. (2008b) *J. Planar Chromatogr.*, **21**, 237–241.

Cieśla, Ł., Petruczynik, A., Hajnos, M., *et al*. (2008c) *Part II. Method for quantitative analysis of selected coumarins in plant material. J. Planar Chromatogr.*, **21**, 447–452.

Cieśla, Ł., Skalicka-Woźniak, K., Hajnos, M., *et al*. (2009) *Acta Chromatogr.*, **21**, 641–657.

Cisowski, W., PałkaGudyka, E., Krauze-Baranowska, M., *et al*. (1991) *J. Planar Chromatogr.*, **4**, 471–474.

Conforti, F., Marrelli, M., Menichini, F., *et al*. (2009) *Curr. Drug Ther.*, **4**, 38–58.

Cooper, D. C., and White, F. J. (2000) *Brain Res.*, **880**, 212–218.

Dekermendjian, K., Ai, J., Nielsen, M., *et al*. (1996) *Neurosci. Lett.*, **219**, 151–154.

Dhalwal, K., Shinde, V. M., and Mahadik, K. R. (2008) *Chromatorgraphia*, **67**, 163–167.

Franco, L., and de Oliveira, B. H. (2010) *Talanta*, **81**, 1368–1372.

Gawron, A., and Główniak, K. (1987) *Planta Med.*, **53**, 526–529.

Główniak, K., Mroczek, T., Zabza, A., *et al*. (2000) *Pharm. Biol.*, **38**, 1–5.

Główniak, K., Bartnik, M., Mroczek, T., *et al*. (2002) *J. Planar Chromatogr.*, **15**, 94–100.

Govindarajan, R., Singh, D. P., Singh, A. P., *et al*. (2007) *Chromatographia*, **66**, 401–405.

Guo, F. Q., Huang, L. F., Zhou, H. F., *et al*. (2006) *J. Cent. South Univ. Technol.*, **13**, 156–160.

Guzlek, H., Baptista, I. I. R., Wood, P. L., *et al*. (2010) *J. Chromatogr. A*, **1217**, 6230–6240.

Härmälä, P., Botz, L., Sticher, O., *et al*. (1990) *J. Planar Chromatogr.*, **3**, 515–520.

Härmälä, P., Vuorela, H., Nyiredy, S., *et al*. (1992) *Phytochem. Anal.*, **3**, 42–48.

Hawrył, M., Soczewiński, E., and Dzido, T. (2000) *J. Chromatogr. A*, **886**, 75–81.

Hostettmann, K., Borloz, A., Urbain, A., *et al*. (2006) *Curr. Org. Chem.*, **10**, 825–847.

Hou, Z., Xu, D., Yao, S., *et al*. (2009) *J. Chromatogr. B*, **877**, 2571–2578.

Hou, Z., Luo, J., Wang, J., *et al*. (2010) *Sep. Purif. Technol.*, **75**, 132–137.

Ito, Y. (2005) *J. Chromatogr. A*, **1065**, 145–168.

Kamiński, M., Kartanowicz, R., Kamiński, M. M., *et al*. (2003) *J. Sep. Sci.*, **26**, 1287–1291.

Ketai, W., Huitao, L., Xingguo, C., *et al*. (2001) *Talanta*, **54**, 753–761.

Kozyra, M., Główniak, K., Zabża, A., *et al*. (2005) *J. Planar Chromatogr.*, **18**, 224–227.

Lacy, A., and O'Kennedy, R. (2004) *Curr. Pharm. Design*, **10**, 3797–3811.

Li, J., and Chase, H. A. (2010) *Nat. Prod. Rep.*, **27**, 1493–1510.

Li, H. B., and Chen, F. (2004) *J. Chromatogr. A*, **1061**, 51–54.

Li, H. B., and Chen, F. (2005) *J. Sep. Sci.*, **28**, 268–272.

Liao, Z. G., Wang, G. F., Liang, X. L., *et al*. (2008) *Sep. Purif. Technol.*, **63**, 424–433.

Liazid, A., Palma, M., Brigui, J., *et al*. (2007) *J. Chromatogr. A*, **1140**, 29–34.

Liu, R., Feng, L., Sun, A., *et al*. (2004a) *J. Chromatogr. A*, **1055**, 71–76.

Liu, R., Feng, L., Sun, A., *et al*. (2004b) *J. Chromatogr. A*, **1057**, 89–94.

Liu, R., Li, A., Sun, A., *et al*. (2004c) *J. Chromatogr. A*, **1057**, 225–228.

Liu, R., Li, A., and Sun, A. (2004d) *J. Chromatogr. A*, **1052**, 223–227.

Liu, R., Sun, Q., Shi, Y., *et al*. (2005a) *J. Chromatogr. A*, **1076**, 127–132.

Liu, R., Sun, Q., Sun, A., *et al*. (2005b) *J. Chromatogr. A*, **1072**, 195–199.

Liu, N., Yu, N. J., Li, J., *et al*. (2011) *Chromatographia*, **73**, 1121–1129.

Liu, Z., Jiang, M., Lu, X., *et al*. (2012) *J. Chromatogr. B*, **891**(892), 102–108.

Lozhkin, A. V., and Sakanyan, E. I. (2006) *Pharm. Chem. J.*, **40**, 337–346.

Luszczki, J. J., Główniak, K., and Czuczwar, S. J. (2007a) *Neurosci. Res.*, **59**, 18–22.

Luszczki, J. J., Główniak, K., and Czuczwar, S. J. (2007b) *Eur. J. Pharmacol.*, **574**, 133–139.

Luszczki, J. J., Andres-Mach, M., Cisowski, W., *et al*. (2009) *Eur. J. Pharmacol.*, **607**, 107–109.

Łuszczki, J. J., Wojda, E., Raszewski, G., *et al*. (2008) *Pharmacol. Rep.*, **60**, 566–573.

Ma, C. H., Ke, W., Sun, Z. L., *et al*. (2006) *Chromatographia*, **64**, 83–87.

Marston, A., and Hostettmann, K. (1995) *J. Nat. Prod.*, **58**, 128–130.

Marston, A., and Hostettmann, K. (2006) *J. Chromatogr. A*, **1112**, 181–194.

Martino, E., Ramaiola, I., Urbano, M., *et al*. (2006) *J. Chromatogr. A*, **1125**, 147–151.

Matysik, G., Główniak, K., Soczewiński, E., *et al*. (1994) *Chromatographia*, **38**, 766–770.

Melliou, E., Magiatis, P., Mitaku, S., *et al*. (2005) *J. Nat. Prod.*, **68**, 78–82.

Mitra, P., Barman, P. C., and Chang, K. S. (2011) *Food Bioprocess. Tech.*, **4**, 737–744.

Murray, R. D. H., Mendez, J., and Brown, S. A. (1982) *The Natural Coumarins, Occurrence, Chemistry and Biochemistry*, Chichester, John Wiley & Sons, Ltd.

Orhan, I., Tosun, F., and Sener, B. (2008) *Z. Naturforsch. Sect. C*, **63**, 366–370.

Park, A. Y., Park, S. Y., Lee, J., *et al*. (2009) *Biomed. Chromatogr.*, **23**, 1034–1043.

Senol, F. S., Skalicka-Wozniak, K., Khan, M. T. H., *et al*. (2011) *Phytochem. Lett.*, **4**, 462–467.

Skalicka-Woźniak, K., and Głowniak, K. (2012) *Molecules*, **17**, 4133–4141.

Skalicka-Woźniak, K., Widelski, J., and Głowniak, K. (2008) Plant materials in modern pharmacy and methods of their investigations, in *Thin Layer Chromatography in Phytochemistry*, eds. M. Waksmundzka-Hajnos, J. Sherma and T. Kowalska, CRC Press, Boca Raton, vol. **99**, pp. 15–35.

Skalicka-Woźniak, K., Markowski, W., Świeboda, R., *et al.* (2009a) *Acta Chromatographica*, **21**, 531–546.

Skalicka-Woźniak, K., Mroczek, T., Garrard, I., *et al.* (2009b) *J. Chromatogr. A*, **1216**, 5669–5675.

Skalicka-Woźniak, K., Mroczek, T., Garrard, I., *et al.* (2012) *J. Sep. Sci.*, **35**, 790–797.

Smyth, T., Ramachandran, V. N., and Smyth, W. F. (2009) *Int. J. Antimicrob. Ag.*, **33**, 421–426.

So, Y. K., Ki, Y. L., Sang, H. S., *et al.* (2005) *J. Nat. Prod.*, **68**, 56–59.

Sutherland, I. A., Muytjens, J., Prins, M., *et al.* (2000) *J. Liq. Chrom. Rel. Technol.*, **23**, 2259–2276.

Teng, W. Y., Chen, C. C., and Chung, R. S. (2005) *Phytochem. Anal.*, **16**, 459–462.

Thuong, P. T., Hung, T. M., Ngoc, T. M., *et al.* (2010) *Phytother. Res.*, **24**, 101–106.

Tosun, A., and Tomek, P. (2011) Application of HPLC in coumarin analyses, in *High Performance Liquid Chromatography in Phytochemical Analysis*, eds. M. Waksmundzka-Hajnos and J. Sherma, CRC Press, Boca Raton, vol. **102**, pp. 513–534.

Tosun, A., Baba, M., and Okuyama, T. (2007a) *J. Nat. Med.*, **61**, 402–405.

Tosun, A., Bahadır, O., and Dinc, E. (2007b) *Chromatographia*, **66**, 677–683.

Tosun, F., Kizilay, C. A., Erol, K., *et al.* (2008) *Food Chem.*, **107**, 990–993.

Urbain, A., Marston, A., and Hostettmann, K. (2005) *Pharm. Biol.*, **43**, 647–650.

Waksmundzka-Hajnos, M., and Hawryl, M. (2008) Application of TLC in the isolation and analysis of coumarins, in *Thin Layer Chromatography in Phytochemistry*, eds. M. Waksmundzka-Hajnos, J. Sherma and T. Kowalska, CRC Press, Boca Raton, vol. **99**, pp. 365–403.

Waksmundzka-Hajnos, M., Petruczynik, A., Dragan, A., *et al.* (2004a) *J. Chromatogr. A*, **800**, 181–187.

Waksmundzka-Hajnos, M., Petruczynik, A., Dragan, A., *et al.* (2004b) *Phytochem. Anal.*, **15**, 1–7.

Waksmundzka-Hajnos, M., Petruczynik, A., Hajnos, M., *et al.* (2006) *J. Chromatogr. Sci.*, **44**, 510–517.

Walshe, M., Howarth, J., Kelly, M. T., *et al.* (1997) *J. Pharm. Biomed. Anal.*, **16**, 319–325.

Wang, L. H., and Tso, M. (2002) *J. Pharm. Biomed. Anal.*, **30**, 593–600.

Wang, Q. E., Lee, F. S., and Wang, X. (2004a) *J. Chromatogr. A*, **1048**, 51–57.

Wang, X., Wang, Y., Yuan, J., *et al.* (2004b) *J. Chromatogr. A*, **1055**, 135–140.

Wang, T. T., Jin, H., Li, Q., *et al.* (2007) *Chromatographia*, **65**, 477–481.

Wang, J., Chen, D., Chen, Z., *et al.* (2010) *J. Sep. Sci.*, **33**, 1099–1108.

Wang, L. H., Mei, Y. H., Wang, F., *et al.* (2011) *Sep. Purif. Technol.*, **77**, 397–401.

Wei, Y., and Ito, Y. (2006) *J. Chromatogr. A*, **1115**, 112–117.

Wei, Y., Zhang, T., and Ito, Y. (2004) *J. Chromatogr. A*, **1033**, 373–377.

Wei, Y., Xie, Q., Fisher, D., *et al.* (2009) *Chromatographia*, **70**, 1185–1189.

Wolski, T., Gliński, Z., Hołderna-Kędzia, E., *et al.* (1997) *Annales UMCS Sectio DDD*, **52**, 156–165.

Xiao, G., Li, G., Chen, L., *et al.* (2010) *J. Chromatogr. A*, **1217**, 5470–5476.

Xie, Y., Zhao, W., Zhou, T., *et al.* (2010) *Phytochem. Anal.*, **21**, 473–482.

Xu, J., Zhao, W. M., Qian, Z. M., *et al.* (2010) *J. Sep. Sci.*, **33**, 1580–1586.

Yan, J., Tong, S., Sheng, L., *et al.* (2006) *J. Liq. Chromatorg. Rel. Technol.*, **29**, 1307–1315.

Yang, F., Ou, Q., and Yu, W. (1995) *J. Liq. Chromatorg. Rel. Technol.*, **18**, 395–403.

Yang, F., Zhang, T., Liu, Q., *et al.* (2000) *J. Chromatogr. A*, **883**, 67–73.

Yang, W., Feng, C., Kong, D., *et al.* (2010) *J. Chromatogr. B*, **878**, 575–582.

Yu, D., Suzuki, M., Xie, L., *et al.* (2003) *Med. Res. Rev.*, **23**, 322–345.

Zgórka, G. (2001) *J. Liq. Chrom. Rel. Technol.*, **24**, 1397–1410.

Zgórka, G., and Głowniak, K. (1999) *Phytochem. Anal.*, **10**, 268–271.

Zgórka, G., Dragan, T., Głowniak, K., *et al.* (1998) *J. Chromatogr. A*, **797**, 305–309.

Zhou, J., Sun, X. L., and Wang, S. W. (2008a) *J. Chromatogr. A*, **1200**, 93–99.

Zhou, J., Wang, S. W., and Sun, X. L. (2008b) *Anal. Chim. Acta*, **608**, 158–164.

Zhou, T., Xiao, X., Li, G., *et al.* (2011) *J. Chromatogr. A*, **1218**, 3608–3615.

Zschocke, S., Liu, J. H., Stuppner, H., *et al.* (1998) *Phytochem. Anal.*, **9**, 283–290.

Naphthoquinones and Anthraquinones: Chemical, Analytical, and Biological Overview

Nahed El-Najjar[1], Hala Gali-Muhtasib[2], Pia Vuorela[3], Arto Urtti[4] and Heikki Vuorela[5]

[1]*Department of Physiology and Biophysics, Weill Cornell Medical College, Doha, Qatar,* [2]*Department of Biology, American University of Beirut, Beirut, Lebanon,* [3]*Division of Pharmaceutical Biosciences, University of Helsinki, Helsinki, Finland,* [4]*Centre for Drug Research, University of Helsinki, Helsinki, Finland* and [5]*Division of Pharmaceutical Biology, University of Helsinki, Helsinki, Finland*

1 INTRODUCTION

Quinones are a class of natural compounds widely distributed in nature, which harbor many physiological and therapeutic effects. They are found in wide varieties of plant families, some of which are listed in Tables 1 and 2 (Saadawi *et al.*, 2012; Yao *et al.*, 2012; Kretschmer *et al.*, 2012; Bringmann *et al.*, 2008; McGaw *et al.*, 2008), fungi, bacteria (Thomson, 1991; Carrasco *et al.*, 2008; Kim *et al.*, 2008; Wijeratne *et al.*, 2008), and in small amounts in animals, specifically in echinoderms (Singh, Moore, and Scheuer, 1967; Thomson, 1991). More than 1200–1500 quinones have been described so far (Thomson, 1987, 1997; Harborne, Baxter, and Moss, 1999; Rauwald, 1990). These molecules share a common basic structure, whereby an ortho or a para-substituted dione is conjugated either to an aromatic nucleus (benzoquinones) (**1**, **2**) or a condensed polycyclic aromatic system such as naphthoquinones (**3**), anthraquinones (**4**), and anthracyclinones (**5**) as shown in Figure 1.

Quinones are a well-studied class of molecules because of their wide occurrence in nature along with their role in many physiological, therapeutic, and toxicological effects. The aim of this chapter is to cover the current knowledge about the biological/toxicological activities of naphthoquinones (**3**) and anthraquinones (**4**) and their analytical detection. However, in order to understand the mechanisms by which naphthoquinones and anthraquinones exert their effects and the problems encountered with their detection, it is important to understand their inherent chemical reactivity.

2 CHEMISTRY OVERVIEW

Being a part of the quinone family, naphthoquinones and anthraquinones share two essential properties: first, they can undergo reversible oxido-reduction reactions and, second, many of them can undergo nucleophilic attack because of their electrophilic character. In the following section, the redox

Table 1 Examples of the main plant families, genera, and the most important species containing the different classes of quinones, that is, benzoquinones and naphthoquinones.

Quinones category	Family	Genera	Species	References
Benzoquinones	Annonaceae	*Mitrella*	*M. kentii*	Saadawi *et al*., 2012; Wu *et al*., 2011c
		Lettowianthus	*L. stellatus*	
	Asteraceae	*Pentacalia*	*P. desiderabilis*	Massaoka *et al*., 2012
	Clusiaceae	*Garcinia*	*G. acutifolia*	Rukachaisirikul *et al*., 2008
	Dioscoreaceae	*Tacca*	*T. chantrieri*	Peng *et al*., 2010
	Fabaceae	*Astragalus*	*A. propinquus*	Zhang *et al*., 2011
	Lamiaceae	Leucosceptrum	*L. canum*	Luo *et al*., 2012
	Myrsinaceae	*Embelia,*	*E. ribes,*	Radhakrishnan *et al*., 2012, Pandey and Ojha
		Ardisia	*E. tsjeriam-cottam,*	2012; Liu *et al*., 2009
			A. gigantifolia	
	Ranunculaceae	*Nigella*	*N. sativa*	Salem, 2005
	Rosaceae	*Pyrus*	*P. ussuriensis*	Jeon, Yang, and Lee, 2012
Naphthoquinones	Acanthaceae	*Rhinacanthus*	*R. nasutus*	Panichayupakaranant, Charoonratana, and Sirikatitham, 2009a; Panichayupakaranant, Sakunpak, and Sakunphueak, 2009b
	Bignoniaceae	*Tabebuia*	*T. impetiginosa*	de Melo *et al*., 2011
	Boraginaceae	*Onosma*	*O. paniculata,*	Kretschmer *et al*., 2012; Sagratini *et al*., 2008
			O. echioides	
		Lithospermum	*L. erythrorhizon*	Gwon *et al*., 2012
		Arnebia	*A. euchroma, A. guttata,*	Xiao *et al*., 2011
			A. benthamii, and	Sharma *et al*., 2009
			A. hispidissima	
	Drosophyllaceae	*Drosophyllum*	*D. lusitanicum*	Grevenstuk *et al*., 2008
	Ebenaceae	*Euclea*	*E. natalensis, E. undulata*	McGaw *et al*., 2008
	Ericaceae	*Pyrola*	*P. rotundifolia*	Ptitsyn *et al*., 2011
	Iridaceae	*Eleutherine*	*E. americana*	Paramapojn *et al*., 2008
	Plumbaginaceae	*Plumbago*	*P. zeylanica*	Bothiraja *et al*., 2012

cycle and electrophilic properties of quinones are overviewed.

2.1 One- and Two-Electron Reduction

The ability of quinones to act as oxidizing or dehydrogenating agents after being reduced contributes mainly to their toxicity. In biological systems, quinones can undergo one- or two-electron reduction, leading to the formation of semiquinones and hydroquinones, respectively.

Microsomal NADPH cytochrome P450 reductase (P450R), microsomal NADH cytochrome b5 reductase (b5R), and mitochondrial NADH ubiquinone oxidoreductase can catalyze the one-electron reduction of quinones (Bartoszek and Wolf, 1992; Monks and Jones, 2002; Holtz *et al*., 2003; Yan *et al*., 2008; Samuni *et al*., 2010; Komatsu, Tanabe, and Nishimoto, 2011). Under aerobic conditions, the semiquinone formed by one-electron reduction gets

oxidized to the original quinone with the generation of superoxide anions radicals, which in aqueous solution interact with molecular oxygen and give rise to hydrogen peroxide. Interaction between iron and hydrogen peroxide leads to the formation of toxic hydroxyl radicals to which the toxicity of quinones is attributed (Kappus, 1986; Asche, 2005). As the one-electron-reducing enzymes enhance the toxicity of quinones, they can be used in the design of bioreductive chemotherapeutic agents (Celik and Arinc, 2008; Yan *et al*., 2008). Actually, it is well documented that the one-electron bioreductive activation of several quinones-based anticancer drugs enhances their covalent binding to DNA and cytotoxic activity against tumor cells (Bartoszek and Wolf, 1992; Cullinane *et al*., 1994; Skladanowski and Konopa, 1994; Patterson *et al*., 1995, 1997; Joseph, Xu, and Jaiswal, 1996; Belcourt *et al*., 1998; Chinje *et al*., 1999; Bailey *et al*., 2001; Cowen *et al*., 2003; Kostrzewa-Nowak *et al*. 2005; Wang *et al*., 2007b; Snodgrass *et al*., 2010).

Table 2 Examples of the main plant families, genera, and the most important species containing the different classes of quinones, that is, anthraquinones.

Quinones category	Family	Genera	Species	References
Anthraquinones	Aphodelaceae	*Bulbine*	*B. frutescens, B. asphodeloides, B. tortifolia, B. aphodeloides, B. latifolia, B. natalensis, B. alooides, B. asphodeloides, B. latioflia, B. alooides, B. narcissifolia, B. capitata, abyssinica*	Reviewed in Bringmann *et al.*, 2008
		Kniphofia	*K. foliosa, K. buchananii, K. parviflora, K. laxiflora, and K. rooperi, K. tysonii, K. umbrina, K. insignis, K. pumila, K. albescens*	Reviewed in Bringmann *et al.*, 2008
		Bulbinella	*B. floribunda, B. divaginata, B. elata, B. latifolia, B. nutans*	Reviewed in Bringmann *et al.*, 2008
	Fabaceae	*Senna*	*S. alata*	Panichayupakaranant, Charoonratana, and Sirikatitham, 2009a; Panichayupakaranant, Sakunpak, and Sakunphueak, 2009b
	Juglandaceae	*Juglans*	*J. mandshurica*	Yao *et al.*, 2012
	Polygonaceae	*Fallopia*	*F. multiflora, F. japonica*	Zuo *et al.*, 2008 Qian *et al.*, 2008
		Rheum	*R. palmatum*	Lee *et al.*, 2012
			R. undulatum	Lo *et al.*, 2012
	Rhamnaceae	*Berchemia*	*A. vera*	Wei *et al.*, 2008
		Rhamnus	*B. floribunda*	Genovese *et al.*, 2010
			R. alpinus	Warner, Vath, and Falvey, 2003
	Xanthorrhoeaceae	*Aloe*	*A. ferox*	Tan, Li, and Xu, 2012

On the other hand, quinones can also undergo two-electron reduction, which is a process that is catalyzed by the cytosolic flavoenzymes NAD(P)H: quinone acceptor oxidoreductases (NQO), which is also known as *DT-diaphorase* (Siegel, Yan, and Ross, 2011). DT-diaphorase is considered as a distinctive flavoenzyme because (i) it shows a broad electron acceptor specificity and reduces quinones and structurally related compounds and displays non-selective reactivity toward NADH and NADPH, (ii) it is inhibited by the NAD(P)H competitive inhibitor dicumarol and other oral anticoagulants, and (iii) it catalyzes the so-called obligatory two-electron transfers (Cadenas, 1995; Bianchet, Faig, and Amzel, 2004; Dinkova-Kostova and Talalay, 2010). DT-diaphorase competes with the one-electron-reducing enzymes to protect the cells from oxidative stress by the conversion of quinones to hydroquinones rather than semiquinones and reactive oxygen species (ROS) (Kappus and Sies, 1981; Tampo and Yonaha, 1996; Bianchet, Faig, and Amzel, 2004; Guo *et al.*, 2008; Dinkova-Kostova and Talalay, 2010).

However, in addition to the redox-stable hydroquinones formed by DT-Diaphorase, two other types of hydroquinones can be formed as well. The latter two are redox-labile hydroquinones that subsequently autooxidize with formation of ROS and hydroquinones, which readily rearrange to potent electrophiles, participating in bioalkylation reactions (Cadenas, 1995; Nishiyama *et al.*, 2010). Activation or deactivation of quinones depends, therefore, on the properties of the corresponding hydroquinones generated by DT-diaphorase. Many studies support the detoxification properties of DT-diaphorase (Joseph and Jaiswal, 1994; Tampo and Yonaha, 1996). This detoxification results from the conversion of quinones by DT-diaphorase to the more water-soluble hydroquinones that are easily excreted following sulfate or glucuronide conjugation (Lind, 1985; Hao *et al.*, 2007; Nishiyama *et al.*, 2008, 2010). In addition to its role in the detoxification of quinones, DT-diaphorase catalyzes the bioreductive activation of quinolic chemotherapeutic compounds by forming unstable hydroquinones that autooxidize
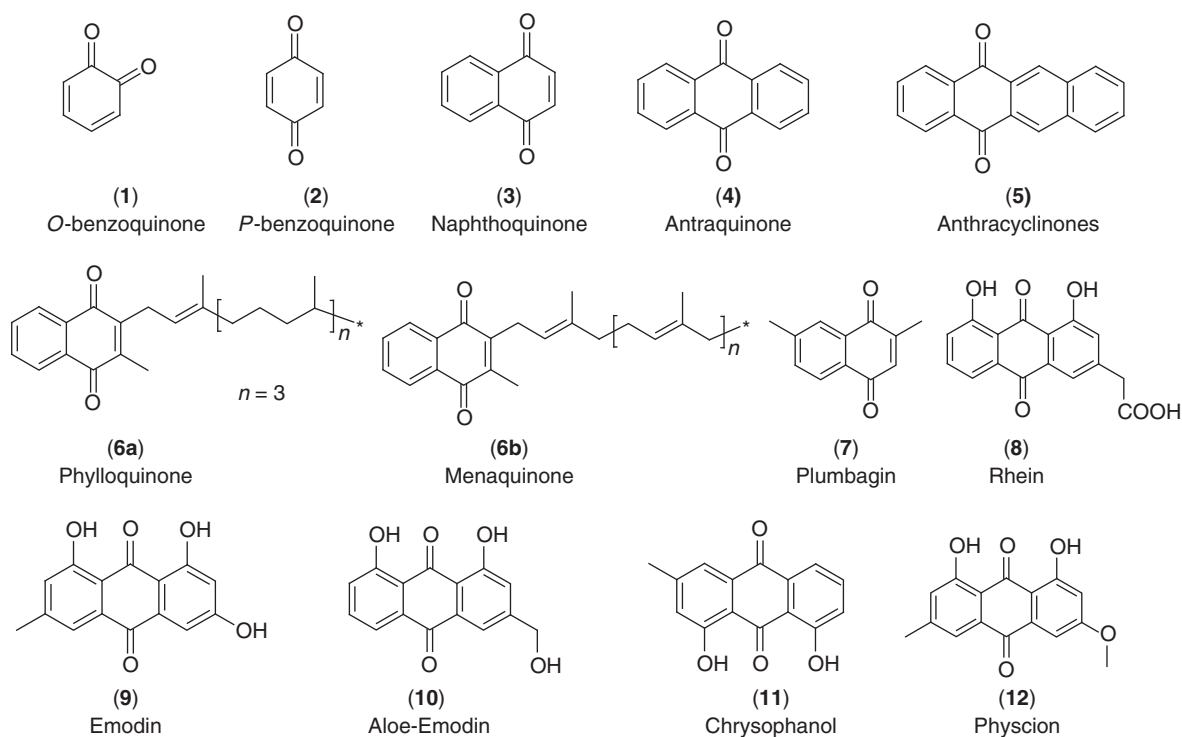
Figure 1 Chemical structures of naphtho- and anthraquinones numbered in the same order as mentioned in the text.

with the formation of ROS or that rearrange to form reactive alkylating species (Cadenas, 1995; Workman, 1994; Danson *et al*., 2011; Lee *et al*., 2011).

2.2 Nucleophilic Addition of Quinones

The electrophilic character of quinones makes them susceptible to undergo nucleophilic addition with nucleophilic species such as amine, hydroxyl, or thiol groups. In biological systems, such nucleophiles are present as side reactive groups of lysine, cysteine, and serine (Magee, 2000; Bender *et al*., 2007). The adduction of quinones with the aforementioned species occurs via the classical Michael addition and can lead to either their detoxification or enhanced toxicity (Land, Ramsden, and Riley, 2004; Li, Heinze, and Haehnel, 2005; Song and Buettner, 2010). Glutathione (GSH), an active ROS scavenger, is the most abundant non-protein antioxidant present in the cells, and its thiol group is the first to be involved in the nucleophilic addition reaction with quinones. The reaction of quinone with GSH leads to the

formation of more hydrophilic compound as compared to the parent quinone and is, therefore, considered a detoxification reaction (Belchik and Xun, 2011). Hydroquinone–glutathionyl conjugates can be formed either spontaneously via a reductive addition or is catalyzed by glutathione-*S*-transferase (Buffinton *et al*., 1989; Jakoby and Ziegler, 1990; Belchik and Xun, 2011).

In addition to its role in detoxification, the nucleophilic addition might lead to the enhanced toxicity of quinones. Such observation might be explained by the fact that in some cases the glutathionyl conjugates have a faster redox cycling potential as compared to the parent quinone (Buffinton *et al*., 1989; Jakoby and Ziegler, 1990; Van Ommen *et al*., 1992). Depletion of the reduced thiol form of GSH in the presence of high concentrations of quinones is on the basis of another mechanism of quinone-induced toxicity. Once GSH is depleted, cellular thiol-dependent proteins can be alkylated, thereby causing irreversible changes and cell death (Buffinton *et al*., 1989; Jakoby and Ziegler, 1990). The most popular mechanistic theory underlying the toxicity of quinones is their adduction with

(13)
Doxorubicin

(14)
Aloin A

(15)
Menatetrenone

(16)
Menatetrenone epoxide

(17)
Chimaphilin

(18)
Buparvaquone

(19)
Acetylshikonin

(20)
β-acetoxyisovalerylshikonin

(21)
β-Lapachone

(22)
Menadione

Figure 1 *(Continued)*

3 BIOLOGICAL ACTIVITIES

3.1 Naphthoquinones

Several biological and toxicological activities have been attributed to naphthoquinones. For instance, vitamin K **(6a, b)**, a naphthoquinone derivative, is essential to maintain life by its function in blood coagulation processes (Ahmed *et al*., 2007;

functional groups found on many cellular components, such as proteins or DNA, leading hence to disruption of biological functions (Buffinton *et al*., 1989; Jakoby and Ziegler, 1990; Endo *et al*., 2011).

Azharuddin *et al*., 2007; Benzakour, 2008; Pucaj *et al*., 2011), in preventing cardiovascular disease (Beulens *et al*., 2008; Wallin, Schurgers, and Wajih, 2008), in bone metabolism (Kudo *et al*., 1998), as well as in the prevention and treatment of osteoporosis (Bugel, 2008; Lanham-New, 2008; Weber, 2001). Moreover, several vitamin K analogs exert an anti-inflammatory activity against lipopolysaccharide (LPS)-treated human and rat cells (Ohsaki *et al*., 2010). Other naphthoquinone derivatives suppress the acute inflammation induced by 12-*O*-tetradecanoylphorbol-13-acetate in mouse ears (Maruo *et al*., 2011), reduce *in vitro* LPS-mediated TNFα release in rat primary macrophage cultures and in animal models (Lu *et al*., 2011), and inhibit the inflammatory marker leukotrienes induced

by antigen stimulation in mast cells (Kawamura, Nakanishi, and Hirashima, 2010). Naphthoquinones are also effective against several cancer types. For instance, plumbagin **(7)** has been shown to exert anti-tumor effect against human promyelocytic leukemia cells (NB4) *in vitro* and *in vivo* using NB4 tumor xenograft in NOD/SCID mice (Xu and Lu, 2010). The observed cell death is mediated via the production of ROS (Xu and Lu, 2010). Plumbagin **(7)** inhibited as well the growth of human K562 leukemia cells via ROS-induced apoptosis (Sun and McKallip, 2011). *In vivo*, plumbagin **(7)** has also been found to inhibit UV-induced skin cancer, an effect that is associated with the inhibition of proliferation and induction of apoptosis (Sand *et al.*, 2012). Other naphthoquinone derivatives inhibited the human non-small lung cancer cells via caspase-dependent apoptosis (Lim *et al.*, 2007; Acharya *et al.*, 2011) and autophagy (Acharya *et al.*, 2011). Naphthoquinones and their derivatives have been shown to inhibit colon cancer and melanoma growth using *in vitro* and *in vivo* models (Chung *et al.*, 2011; Maruo *et al.*, 2011; Subramaniya *et al.*, 2011; Vasconcellos *et al.*, 2011), suppress biliary carcinogenesis *in vivo* using a hamster biliary carcinogenetic model (Tsuchida *et al.*, 2011), inhibit the growth of human gastric cancer cells, human glioma cancer cells, human cervical cancer cells, and murine pancreatic acinar cells via ROS-mediated apoptosis (Criddle *et al.*, 2006; Cheng *et al.*, 2011; Wu *et al.*, 2011a; Kim *et al.*, 2012), and reduce the growth of breast cancer cells by inducing endoplasmic reticulum stress, JNK activation, and mitochondria-induced apoptosis (Lee *et al.*, 2011). In addition, a wide variety of naphthoquinones and their derivatives are found to have an antiparasitic effect against *Plasmodium falciparum* (Müller *et al.*, 2011), *Leishmania amazonensis* (Lima *et al.*, 2004), and *Toxoplasma gondii* (Portes *et al.*, 2011). Naphthoquinones show as well an antimicrobial effect against *Propionibacterium acnes*, which is related to the pathogenesis of the inflamed lesions in a common skin disease, acne vulgaris (Cho, Sultan, and Moon, 2009). Antifungal activities of naphthoquinones have also been reported. Several naphthoquinone derivatives have shown antifungal activity when tested against a wide variety of fungal pathogens including *Candida tropicalis*, *Fusarium oxysporum*, and *Trichophyton tonsurans* (Wittebolle *et al.*, 2006; Dzoyem *et al.*, 2011), *Candida krusei*, and *Saccharomyces cerevisiae* (Sasaki, Abe, and Yoshizaki, 2002). Other

naphthoquinone has shown a neuroprotective effect against cerebral ischemia-/reperfusion-induced brain injury by inhibiting oxidative stress (Wang *et al.*, 2010).

Although, naphthoquinones exert several beneficial effects, some studies have shown that some of the available anti-inflammatory and antimalarial naphthoquinones can induce embryolethality in pregnant rats without affecting the mother (Guerra Mde *et al.*, 1999, 2001). In addition, other *in vivo* studies show that environmental naphthoquinones can aggravate airway inflammation and therefore respiratory diseases are induced by the diesel exhaust particles (Costa *et al.*, 2010).

3.2 Anthraquinones

Anthraquinones and their derivatives exert several beneficial effects against different ailments. For instance, Rhein **(8)** has been found to be effective against nonalcoholic fatty liver disease in high fat diet-induced obesity in mice. Rhein **(8)** exerted its effect by downregulating the lipogenesis, increasing the energy expenditure and restricting the expression of proinflammatory cytokines (Sheng *et al.*, 2011). Other anthraquinones have shown a hypoglycemic effect in streptozotocin-induced diabetic rats (Shukla *et al.*, 2010). In addition, Emodin **(9)** has been found, using *in vitro* and *in vivo* models, to be a potent selective inhibitor of 11beta-hydroxysteroid dehydrogenase type 1, which is an enzyme closely associated with the development of metabolic abnormalities coupled with type 2 diabetes (Feng *et al.*, 2010). Emodin **(9)** has also improved the glucose tolerance and insulin sensitivity, and modulated the expression of metabolism-related genes in streptozotocin-induced diabetic rats (Xue, Ding, and Liu, 2010). Rhein **(8)** has been reported to attenuate hyperlipidemia and hyperglycemia in diabetic mice and protect against the potential renal damage associated with diabetes (Gao *et al.*, 2010). Emodin **(9)** has also been reported to protect the liver against acetaminophen-induced toxicity in rats (Bhadauria, 2010).

Likewise, Emodin glycoside shows protective effect against ischemia-reperfused cerebral injury in rats and against glutamate-induced toxicity in cultured cortical cells of fetal rats (Wang *et al.*, 2007a). Furthermore, Q-PCR and proteomic analysis show that Aloe-emodin metabolites protect

against *N*-methyl-D-aspartate-induced apoptosis in retinal ganglion cells mainly through modulation of the expression of the antioxidant Cu–Zn superoxide dismutase (Lin *et al*., 2007). In addition, anthraquinones have shown potent anticancer activities against several cancers. For instance Aloe-emodin **(10)** has promising effect against skin cancer, whereby it inhibited the proliferation and invasion potential and induced the differentiation of the metastatic B16-F10 melanoma murine cells (Tabolacci *et al*., 2010). Other anthraquinone/anthraquinone derivatives have been shown to inhibit the growth of human gliobastoma cells by inducing caspase and ROS-dependent apoptosis (Lu *et al*., 2010), inhibit the growth of human bladder carcinoma cell line and colon cancer cells by inducing G2/M cell cycle arrest and apoptosis (Lu, Zhang, and Qian, 2008; Tu *et al*., 2011), inhibit human cervical cancer cell line by inducing S and G2/M arrest (Lin *et al*., 2011), inhibit gastric cancer cells by inducing caspase-dependent apoptosis (Chiang *et al*., 2011), and inhibit the growth of pancreatic, liver, and breast cancers *in vitro* and *in vivo* (Chen *et al*., 2011; Liu *et al*., 2011a, b; Wu *et al*., 2011a, b). In addition to their anticancer effects, anti-inflammatory effects have been attributed to anthraquinones and their derivatives. For instance, Aloe-emodin **(10)** inhibits LPS-induced inflammatory response in RAW 246.7 macrophages through suppression of iNOS and COX-2 mRNA expression (Park, Kwon, and Sung, 2009). Emodin **(9)** suppresses proinflammatory markers in interleukin-1β (IL-1β) and LPS-treated rheumatoid arthritis synoviocytes under hypoxic conditions (Ha *et al*., 2011) and in LPS-treated human umbilical vein endothelial cells (Meng *et al*., 2010). Other anthraquinones have been found to reduce intestinal inflammation in ulcerative colitis rat model by suppressing TNF-α and IL-1β (Park, Kwon, and Sung, 2011). Furthermore, antibacterial effects have also been attributed to anthraquinones. For instance, several anthraquinones have shown antibacterial effect against *Staphylococcus aureus*, *Pseudomonas aeruginosa*, *Escherichia coli*, *Streptococcus pneumonia*, and *Helicobacter pylori* (Xiang *et al*., 2008; Zhang *et al*., 2008; Chen *et al*., 2009; Beattie *et al*., 2010; Comini *et al*., 2011). Another anthraquinone derivative isolated from cultured freshwater *Cyanobacterium eucapsis* sp has shown selective activity against *Mycobacterium tuberculosis* (Sturdy *et al*., 2010). Furthermore, an anthraquinone derivative shows antifungal activity

against several candida species mainly by inhibiting the drug efflux pumps and the mitochondrial membrane potential (Kang, Fong, and Tsang, 2010), while others has been found to be active against *Aspergillus fumigatus* and *Candida albicans* (Mishra *et al*., 2010). Some anthraquinone derivatives have also antiparasitic effects against *Leishmania major* and *P. falciparum* (Sittie *et al*., 1999; Osman *et al*., 2010). In addition to the beneficial effects of anthraquinones, toxicological effects of this class of molecules are also encountered. For instance, it is well known that sunlight photooxidation of environmental contaminants such as the polyclic aromatic hydrocarbons (PAHs) leads to the formation of toxic anthraquinones (Mallakin *et al*., 1999; Mallakin, Dixon, and Greenberg, 2000). The introduction of PAH to the aquatic environment through terrestrial runoff and offshore drilling operations might impact the ecosystem health. This can be explained by the fact that toxic anthraquinones, which are formed from PAH photooxidation, might significantly influence the bacterial growth and therefore affect the biogeochemical processes mediated by the bacteria (Vaughan *et al*., 2010).

4 ANALYTICAL DETECTION METHODS FOR NAPHTHOQUINONES AND ANTHRAQUINONES

Determination of naphthoquinones and anthraquinones from plants, pharmaceutical preparations, and biological samples has been performed using a variety of analytical methods including thin layer chromatography (TLC), gas chromatography (GC) (Raspotnig *et al*., 2010), high performance liquid chromatography (HPLC) (Sakunphueak and Panichayupakaranant, 2010), and mass spectrometry (MS) (Zhao *et al*., 2010). Although, TLC is used for the quick detection of quinones (Table 3); HPLC and HPLC/MS are among the methods most commonly used.

Sample cleanup procedures for naphthoquinones/anthraquinones include liquid–liquid extraction, solid-phase extraction (SPE), and protein precipitation. SPE is usually used to clean and concentrate the samples and the most commonly used cartridges are C18 and Oasis HLB (Karpinska *et al*., 2006; Azharuddin *et al*., 2007; Vainchtein *et al*., 2008). To disrupt the protein binding and remove interference from biological samples, protein precipitation

Table 3 Quick detection of naphthoquinones and anthraquinones by thin layer chromatography (TLC).

Quinones category	Matrix	Developing solvent	Detecting reagents/methods	References
Naphthoquinones	Cell suspension cultures of *Rudgea jasminoides*	Toluene:formic acid (99 : 1 v/v)	Spraying with 10% potassium hydroxide in methanol	De Cacia *et al*., 2007
	Cultures of *Campylobacter jejuni* and *Campylobacter fetus* subsp. *fetus*	Methanol/acetone (50 : 50, v/v)	254 nm UV	Carlone and Anet, 1983
	Plant (*Arnebia species*)	ACN/methanol/5% formic acid in water (40 : 02 : 08 v/v/v)	After drying, spots were visualized under UV (254 and 366 nm) and visible light	Sharma *et al*., 2009
	Madder lakes, reseda lakes	Toluene/acetic acid (9 : 1), acetic acid/methanol/ water (3 : 3 : 4), ethyl acetate/ tetrahydrofuran/water (6 : 35 : 47)	NH3 vapor or by UV irradiation	Grygar *et al*. 2003
	Henna powder	Chloroform–methanol (8.5 : 1.5)	334 nm UV	El-Shaer *et al*., 2007
Anthraquinones	Plant (*Cissus populnea*)	Petroleum ether (60–80 °C): ethylacetate:acetic acid (45 : 5 : 3)	25% nitric acid (heated at 110 °C for 10 min), cooled plates sprayed with 5% potassium hydroxide in 50% ethanol	Ibrahim *et al*., 2011
	Cell suspension cultures of *R. jasminoides*	Petroleum ether:ethylacetate:formic acid (75 : 25 : 1 v/v/v)	Spraying with 10% potassium hydroxide in ethanol	De Cacia *et al*., 2 007
	Plant (Ceratotheca triloba)	Hexane:ethyl acetate (9 : 1), supplemented with 1% ammonium chloride	Visualized under visible and UV (254 and 360 nm) Sprayed with ρ-anisaldehyde (5% anisaldehyde in 5% sulfuric acid in ethanol) and heated for 2–5 min at 100 °C	Mohanlall *et al*., 2011
	Leaf, fruit and root suspension cultures of *Morinda citrifolia*	Ethyl acetate:methanol:water (100 : 13.5 : 10)	254 nm UV light	Deshmukh *et al*., 2011
	Plant (*Loranthaceae*)	Ethyl acetate : toluene: Acetic acid (8 : 4 : 1)	254 nm UV light and 5% potassium hydroxide in methanol	Wahab, Ayodele, and Moody, 2010

method using methanol, ethanol, and acetonitrile is most commonly used.

Several detection methods have been coupled to HPLC for the detection of quinones.

For instance, UV (Song *et al*., 2010), chemiluminescence (CL) (Ahmed *et al*., 2007, 2009, 2011), and fluorescence (Azharuddin *et al*., 2007) have been widely used. The ability of quinones to generate hydrogen peroxide and a fluorophore when subjected to UV irradiation allows their detection by CL (Ahmed *et al*., 2007). Post-column chemical reduction using a catalyst-reduced column and a methanol–ethanol mobile phase as reductant has

been also used for the detection of the reduced form of quinones (Azharuddin *et al*., 2007). Various quinones can also be detected by GC (Lim *et al*., 2011). Derivatization of the samples with *N*,*O*-bis (trimethylsilyl)trifluoroacetamide) +1% TMCS (trimethylchlorosilane) is often used (El Sohly, Gul, and Murphy, 2004) and samples are usually separated using dimethylpolysiloxane and silica-based columns (El Sohly, Gul, and Murphy, 2004; Zuo *et al*., 2008).

The identification of quinones is usually performed by MS coupled to GC or HPLC. Different mass analyzers such as electrospray ionization (ESI), and

Table 4 Summary of selected analytical methods used for the detection of naphthoquinones.

Compound(s)	Matrix	Sample preparation	Separation	Detection	LOQ (ng mL^{-1})	References
Naphthoquinone (**3**) derivatives	Tablets	Extraction with MeOH	Nucleosil 100-5 phenyl column Mobile phase: MeOH–0.02 M sodium dihydrogen phosphate mixture (60 : 40, v/v) pH 6.0	HPLC-UV (270 nm)	96 1300–2800	Walash et al., 2011 Paramapoin et al., 2008
	Plant material (E. Americana bulbs)	Extraction with MeOH	C$_{12}$ column Mobile phase: gradient elution with H$_2$O (0.01% formic acid)–ACN	HPLC-UV (254 nm)/MS using ESI in alternating ionization mode		
Menadione (**22**)	Animal feed	Samples treated with Savinase proteinases followed with extraction with EtOH and SPE	C$_{18}$ column Mobile phase: MeOH–H$_2$O (98 : 2, v/v)	HPLC-UV 230 nm vitamin E (**9**), 265 nm Menadione (**21**)	250–4200	Xue, You, and He, 2008
Menatetrenone (**15**), Menatetrenone epoxide (**16**)	Human plasma	Protein precipitation with MeOH	C$_{18}$ column Mobile phase: MeOH	LC/MS/MS with APCI in positive ionization mode	2.5	Kang et al., 2007
Chimaphilin (**17**)	Rat plasma	Extraction with diethylether	C$_{18}$ column Mobile phase: MeOH–H$_2$O (75 : 25, v/v)	HPLC-APCI in negative ionization mode	10	Zhang et al., 2006
Buparvaquone (**18**)	Rat perfusion solution	Centrifugation	C$_4$ column Mobile phase: ammonium acetate (0.02 M)–ACN (30 : 70, v/v)	HPLC-UV (251 nm)	200	Venkatesh et al., 2007
	Human and rabbit plasma	Protein precipitation with ACN followed with SPE	C$_{18}$ column Mobile phase: ammonium acetate (0.02 M)–ACN (18 : 82, v/v)	HPLC-UV (251 nm)	50	Venkatesh et al., 2008

Table 4   (*Continued*)

Compound(s)	Matrix	Sample preparation	Separation	Detection	LOQ (ng mL^{-1})	References
β-lapachone (**21**)	Mouse plasma, tumor homogenate	Protein precipitation with ACN	SB-C$_8$ column Mobile phase: gradient elution with H$_2$O (0.1% formic acid)–ACN (0.1% formic acid)	HPLC/MS/MS using ESI in positive ionization mode	3	Savage et al., 2008
Plumbagin (**7**)	Rat plasma	Extraction with ethylacetate	C$_{18}$ column Mobile phase: H$_2$O/ACN (40 : 60,v/v)	LC–MS/MS using ESI in negative ionization mode	10	Hsieh, Lin, and Tsai, 2006
Acetylshikonin (**19**), B-acetoxyiso-valerylshikonin (**20**)	Cell culture suspension of A. euchroma	Extraction with chloroform	C$_{18}$ column Mobile phase: ACN–MeOH (95 : 5, v/v)	Prep-HPLC (520 nm)	209 (Acetylshikoni), 487 (B-acetoxy-isovalerylshikonin)	Sharma et al., 2008
Vitamin K1 (**6a, b**)	Human plasma	Protein precipitation with EtOH followed by extraction with hexane	ODS UG 120 Mobile phase: imidazol-HNO$_3$ buffer–(500 mM, pH 9.0) and ACN (10 : 90) NaOHaq as a single post-column CL reagent	HPLC–CL	0.03–0.1	Ahmed et al., 2011

Abbreviations: MeOH, methanol; EtOH, ethanol; ACN, acetonitrile; H$_2$O,water.

Table 5 Summary of selected analytical methods used for the detection of anthraquinones.

Compound (s)	Matrix	Sample preparation	Separation	Detection	LOQ (ng mL^{-1})	References
Rhein (**8**), Emodin (**9**), Aloe-emodin (**10**), chrysophanol (**11**), physcion (**12**)	Plant (*Semen Cassiae*)	Extraction with EtOH	C$_{18}$ column Mobile phase: ACN + 0.1% phosphoric acid	HPLC-UV (278 nm)	260 (Emodin), 240 (Rhein), 360 (chrysophanol), 350 (Aloe-emodin), 440 (physcion)	Xu *et al.*, 2011
Emodin (**9**)	Rat plasma	Extraction with MeOH	C$_{18}$ column Mobile phase: gradient of and 2.5 mM NH$_4$Ac, pH 7.4	UPLC-MS/MS with API in negative ionization mode	39	Liu *et al.*, 2011a, b
Rhein (**8**), Emodin (**9**), Aloe-emodin (**10**)	Rat plasma	Protein precipitation with MeOH	C$_{18}$ column Mobile phase: H$_2$O (0.1% formic acid)–MeOH (30 : 70, v/v)	LC-MS/MS using ESI in negative ionization mode	0.5 (Emodin), 0.2 (Aloe-emodin), 2 (Rhein)	Xu *et al.*, 2008
Aloe-emodin (**10**), aloin A (**14**)	Commercials Aloe-based products (liquids, gels and solids)	Extraction with ethylacetate/ MeOH followed by derivatization with BSTFA	DB-1 GC column	GS/MS	5 (Aloe-emodin), 50 (aloin A)	El Sohly, Gul, and Murphy, 2004
Rhein (**8**), Emodin (**9**), chrysophanol (**11**), physcion (**12**)	Plant (*rhisoma et radix polygoni cuspidate*)	Extraction with MeOH	C$_{18}$ column Mobile phase: gradient elution with H$_2$O (0.4% formic acid)–ACN	HPLC-UV (290 nm)	620 (Rhein, chrysophanol), 544 (physcion), 330 (Emodin)	Qian *et al.*, 2008
Rhein (**8**), Emodin (**9**), Aloe-emodin (**10**), chrysophanol (**11**), physcion (**12**)	Plant (*radix Polygoni multifori*)	Extraction with MeOH followed by derivatization with BSTFA + 1% TMCS	EC^{TM} 5 capillary column	Capillary GC–FID/MS	220 (Emodin), 600 (Rhein), 260 (chrysophanol), 520 (Aloe-emodin), 540 (physcion)	Zuo *et al.*, 2008
Rhein (**8**), Emodin (**9**), ALoe-emodin (**10**), chrysophanol (**11**), physcion (**12**)	Plant (*rhubarb*)	Extraction of the MeOH plant extract with chloroform	C$_{18}$ column Mobile phase: H$_2$O (0.1% phosphoric acid)–MeOH (31 : 69, v/v)	UPLC (254 nm)	200 (Emodin, Rhein, chrysophanol), 120 (Aloe-emodin), 400 (physcion)	Wang *et al.*, 2008

Abbreviations: MeOH, methanol; EtOH, ethanol; ACN, acetonitrile; H$_2$O,water; BSTFA, *N,O*-Bis(trimethylsilyl)trifluoroacetamide; TMCS, trimethylchlorosilane; FID, Flame ionization detection.

atmospheric pressure chemical ionization (APCI) instruments that enable tandem mass spectrometry (MS/MS) measurements are used. The limits of quantification (LOQ) reported for various quinones vary widely between HPLC and GC and this show that although all compounds are quinones their analytical detection depends on their structures and the methods used.

In Tables 4 and 5, a summary of the different HPLC/GS/MS analytical methods used for the detection of naphthoquinones and anthraquinones from different matrices is presented. The tables

include the following: compound analyzed, matrix, sample cleanup, separation, detection, and minimum quantification limits.

Identification and quantification of quinones is challenging because of their high reactivity as fast redox cycling molecules as well as their ability to bind nucleophilic groups. Therefore, there is an ongoing effort to establish efficient, accurate, and precise procedures for their quantification. The failure in studying some quinones in biological samples using conventional methods has led to the use of other approaches such as the use of radiolabeled or isotopically labeled compounds.

Isotopic or radioactive compounds are valuable tools for understanding the metabolism and disposition of both endogenous molecules and drugs, especially for compounds that are unstable or require to be detected at low concentrations.

A thorough literature search has shown that the use of labeled quinones has been instrumental in clarifying their metabolic fate and/or mode of action. For instance, vitamin K1 **(6a)** represents a perfect example for a problematic quinone. The specific challenges for its analysis in plasma result from its low concentration, interference of plasma lipid components, and the sensitivity of the molecule to degradation by light and strong alkaline conditions. Three attempts to measure vitamin K1 turnover in human subjects have been made between 1972 and 1979 using $[1', 2'-3H_2]$ vitamin K1. However, owing to the absence of a suitable method for measuring vitamin K1 in plasma, none of these studies has allowed the calculations of its body pool (Shearer *et al*., 1972; Shearer, McBurney, and Barkhan, 1974; Bjornsson *et al*., 1979). Nearly 20 years later and using tritiated vitamin K1, Olson *et al*. was able to determine the total body vitamin K1 and its turnover in human subjects at two levels of vitamin K intake (Olson *et al*., 2002). β-lapachone **(21)**, a promising anticancer compound, is another example of a problematic quinone. Miao *et al*. could not detect with conventional LC–MS the *in vitro* metabolism of β-lapachone in plasma and whole blood (Miao *et al*., 2008). However, the use of C_{14} β-lapachone has clarified the reason for the failure of its detection in blood using conventional analytical methods and allowed determining its fate. By the use of LC–MS coupled to a radioisotope counting system, it has been found that β-lapachone is extensively metabolized in whole blood under *in vitro* conditions and that the enzymatic activity is located in red blood cells. Analysis of the protein pellet prepared from whole blood spiked with C_{14} β-lapachone shows that covalent protein binding of the compound and/or its metabolites is a minor contributor in the failure of its detection in blood (Miao *et al*., 2008).

5 CONCLUSION

Naphthoquinones and anthraquinones are widely distributed in nature and are reported to have several beneficial effects. Their activity against cancer is well documented and is associated with their ability to generate ROS and/or through their nucleophilic addition to proteins/DNA, thus compromising the proper function of the cells. As anti-inflammatory agents, naphtho- and anthraquinones downregulate the proinflammatory markers both *in vivo* and *in vitro*. Other effects are also reported for naphtho- and anthraquinones such as antifungal, antibacterial, and antidiabetic, and so on. Toxicological effects of naphtho- and anthraquinones have been also documented; however, more studies are needed to increase the awareness about their harmful effects. From a chemical point of view, these molecules are unstable and therefore their analytical detection from biological matrices is problematic. However, the advances in the current analytical methods as well as the use of labeled compounds pave the way toward understanding and clarifying the mechanisms by which they exert their effects.

6 RELATED ARTICLES

Extraction Methodologies: General Introduction; Thin-layer Chromatography, with Chemical and Biological Detection Methods; HPLC and Ultra HPLC: Basic Concepts; LC and LC-MS: Techniques and Applications.

REFERENCES

Acharya, B. R., Bhattacharyya, S., Choudhury, D., *et al*. (2011) *Apoptosis*, **16**, 924–939.

Ahmed, S., Kishikawa, N., Nakashima, K., *et al*. (2007) *Anal. Chim. Acta*, **591**, 148–154.

Ahmed, S., Kishikawa, N., Ohyama, K., *et al*. (2009) *J. Chromatogr. A*, **1216**, 3977–3984.

Ahmed, S., Kishikawa, N., Ohyama, K., *et al*. (2011) *Talanta*, **85**, 230–236.

Asche, C. (2005) *Mini Rev. Med. Chem.*, **5**, 449–467.

Azharuddin, M. K., O'Reilly, D. S., Gray, A., *et al.* (2007) *Clin. Chem.*, **53**, 1706–1713.

Bailey, S. M., Lewis, A. D., Patterson, L. H., *et al.* (2001) *Biochem. Pharmacol.*, **62**, 461–468.

Bartoszek, A. and Wolf, C. R. (1992) *Biochem. Pharmacol.*, **43**, 1449–1457.

Beattie, K. D., Rouf, R., Gander, L., *et al.* (2010) *Phytochemistry*, **71**, 948–955.

Belchik, S. M. and Xun, L. (2011) *Drug Metab. Rev.*, **43**, 307–316.

Belcourt, M. F., Hodnick, W. F., Rockwell, S., *et al.* (1998) *Adv. Enzyme Regul.*, **38**, 111–133.

Bender, R. P., Ham, A. J. and Osheroff, N. (2007) *Biochemistry*, **46**, 2856–2864.

Benzakour, O. (2008) *Thromb. Haemost.*, **100**, 527–529.

Beulens, J. W., Bots, M. L., Atsma, F., *et al.* (2008) *Atherosclerosis*, **203**, 489–493.

Bhadauria, M. (2010) *Exp. Toxicol. Pathol.*, **62**, 627–635.

Bianchet, M. A., Faig, M. and Amzel, L. M. (2004) *Methods Enzymol.*, **382**, 144–174.

Bjornsson, T. D., Meffin, P. J., Swezey, S. E., *et al.* (1979) *J. Pharmacol. Exp. Ther.*, **210**, 322–326.

Bothiraja, C., Pawar, A. P., Dama, G. Y., *et al.* (2012) *J. Pharmacol. Toxicol. Methods*, **66**(1), 35–42.

Bringmann, G., Mutanyatta-Comar, J., Knauer, M., *et al.* (2008) *Nat. Prod. Rep.*, **25**, 696–718.

Buffinton, G. D., Ollinger, K., Brunmark, A., *et al* (1989) *Biochem. J.*, **257**, 561–571.

Bugel, S. (2008) *Vitam. Horm.*, **78**, 393–416.

Cadenas, E. (1995) *Biochem. Pharmacol.*, **49**, 127–140.

Carlone, G. M. and Anet, F. (1983) *J. Gen. Microbiol.*, **129**, 3385–3393.

Carrasco, I. J., Marquez, M. C., Xue, Y., *et al.* (2008) *Int. J. Syst. Evol. Microbiol.*, **58**, 1961–1967.

Celik, H. and Arinc, E. (2008) *J. Pharm. Pharm. Sci.*, **11**, 68–82.

Chen, J., Zhang, L., Zhang, Y., *et al.* (2009) *BMC Microbiol.*, **9**, 91.

Chen, P. H., Peng, C. Y., Pai, H. C., *et al.* (2011) *J. Nutr. Biochem.*, **22**, 732–740.

Cheng, H. M., Qiu, Y. K., Wu, Z., *et al.* (2011) *J. Asian Nat. Prod. Res.*, **13**, 12–19.

Chiang, J. H., Yang, J. S., Ma, C. Y., *et al.* (2011) *Chem. Res. Toxicol.*, **24**, 20–9.

Chinje, E. C., Patterson, A. V., Saunders, M. P., *et al.* (1999) *Br. J. Cancer*, **81**, 1127–1133.

Cho, S. C., Sultan, M. Z. and Moon, S. S. (2009) *Arch. Pharm. Res.*, **32**, 489–494.

Chung, H. J., Rhee, H. K., Lee, S. K., *et al.* (2011) *Chem. Biol. Interact.*, **193**, 43–49.

Comini, L. R., Montoya, S. C., Páez, P. L., *et al.* (2011) *J. Photochem. Photobiol. B*, **102**, 108–114.

Costa, S. K., Kumagai, Y., Brain, S. D., *et al.* (2010) *Arch. Toxicol.*, **84**, 109–117.

Cowen, R. L., Patterson, A. V., Telfer, B. A., *et al.* (2003) *Mol. Cancer Ther.*, **2**, 901–909.

Criddle, D. N., Gillies, S., Baumgartner-Wilson, H. K., *et al.* (2006) *J. Biol. Chem.*, **281**, 40485–40492.

Cullinane, C., Cutts, S. M., van Rosmalen, A., *et al.* (1994) *Nucleic Acids Res.*, **22**, 2296–2303.

Danson, S., Johnson, P., Ward, T., *et al.* (2011) *Ann. Oncol.*, **22**, 1653–1660.

De Cacia, O. M., Negri, G., Salatino, A., *et al.* (2007) *Rev. Bras.Bot.*, **30**, 167–172.

Deshmukh, S. R., Wadegaonkar, V. P., Bhagat, R. P., *et al.* (2011) *POJ*, **4**(1), 6–13.

Dinkova-Kostova, A. T. and Talalay, P. (2010) *Arch. Biochem. Biophys.*, **501**, 116–123.

Dzoyem, J. P., Kechia, F. A., Kuete, V., *et al.* (2011) *Nat. Prod. Res.*, **25**, 741–749.

El Sohly, M. A., Gul, W. and Murphy, T. P. (2004) *Int. Immunopharmacol.*, **4**, 1739–1744.

El-Shaer, N. S., Badr, L. M., Aboul-Ela, M. A., *et al.* (2007) *J. Sep. Sci.*, **30**, 3311–3315.

Endo, A., Sumi, D., Iwamoto, N., *et al.* (2011) *Chem. Biol. Interact.*, **192**, 272–277.

Feng, Y., Huang, S. L., Dou, W., *et al.* (2010) *Br. J. Pharmacol.*, **161**, 113–126.

Gao, Q., Qin, W. S., Jia, Z. H., *et al.* (2010) *Planta Med.*, **76**, 27–33.

Genovese, S., Tammaro, F., Menghini, L., *et al.* (2010) *Phytochem. Anal.*, **21**, 261–267.

Grevenstuk, T., Gonçalves, S., Nogueira, J. M., *et al.* (2008) *Phytochem. Anal.*, **19**(3), 229–235.

Grygar, T., Kučkova', S., Hradil, D., *et al.* (2003) *J. Solid State Electrochem.*, **7**, 706–713.

Guerra Mde, O., Mazoni, A. S., Brandão, M. A., *et al.* (1999) *Contraception*, **60**, 305–307.

Guerra Mde, O., Mazoni, A. S., Brandão, M. A., *et al.* (2001) *Braz. J. Biol.*, **61**, 171–174.

Guo, W., Reigan, P., Siegel, D., *et al.* (2008) *Drug Metab. Dispos.*, **36**, 2050–2057.

Gwon, S. Y., Ahn, J. Y., Chung, C. H., *et al.* (2012) *J. Agric. Food Chem.*, **60**(36), 9089–9096.

Ha, M. K., Song, Y. H., Jeong, S. J., *et al.* (2011) *Biol. Pharm. Bull.*, **34**, 1432–1437.

Hao, H., Wang, G., Cui, N., *et al.* (2007) *Curr. Drug Metab.*, **8**, 137–149.

Harborne, J. B., Baxter, H. and Moss, G. P. (1999) *Phytochemical dictionary. A handbook of bioactive compounds from plants*, 2nd edn, Taylor & Francis, London, UK.

Holtz, K. M., Rockwell, S., Tomasz, M., *et al.* (2003) *J. Biol. Chem.*, **278**, 5029–5034.

Hsieh, Y. J., Lin, L. C. and Tsai, T. H. (2006) *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.*, **844**, 1–5.

Ibrahim, H., Mdau, B. B., Ahmed, A., *et al.* (2011) *Afr. J. Tradit. Complement. Altern. Med.*, **8**(2), 140–143.

Jakoby, W. B. and Ziegler, D. M. (1990) *J. Biol. Chem.*, **265**, 20715–20718.

Jeon, J. H., Yang, J. Y. and Lee, H. S. (2012) *J. Food Prot.*, **75**(7), 1258–1262.

Joseph, P. and Jaiswal, A. K. (1994) *Proc. Natl. Acad. Sci. U. S. A.*, **91**, 8413–8417.

Joseph, P., Xu, Y. and Jaiswal, A. K. (1996) *Int. J. Cancer*, **65**, 263–271.

Kang, W., Jeong, J. H., Ma, E., *et al.* (2007) *J. Pharm. Biomed. Anal.*, **44**, 1178–1182.

Kang, K., Fong, W. P. and Tsang, P. W. (2010) *Med. Mycol.*, **48**, 904–911.

Kappus, H. (1986) *Biochem. Pharmacol.*, **35**, 1–6.

Kappus, H. and Sies, H. (1981) *Experientia*, **37**, 1233–1241.

Karpinska, J., Mikoluc, B., Motkowski, R., *et al.* (2006) *J. Pharm. Biomed. Anal.*, **42**, 232–236.

Kawamura, F., Nakanishi, M. and Hirashima, N. (2010) *Biol. Pharm. Bull.*, **33**, 881–885.

Kim, M. K., Park, M. J., Im, W. T., *et al.* (2008) *Int. J. Syst. Evol. Microbiol.*, **58**, 2025–2030.

Kim, J. A., Lee, E. K., Park, S. J., *et al.* (2012) *Int. J. Oncol.*, **40**, 157–162.

Komatsu, H., Tanabe, K. and Nishimoto, S. (2011) *Bioorg. Med. Chem. Lett.*, **21**, 790–793.

Kostrzewa-Nowak, D., Paine, M. J., Wolf, C. R., *et al.* (2005) *Br. J. Cancer*, **93**, 89–97.

Kretschmer, N., Rinner, B., Deutsch, A. J., *et al.* (2012) *J. Nat. Prod.*, **75**(5), 865–869.

Kudo, Y., Iwashita, M., Takeda, Y., *et al.* (1998) *Eur. J. Endocrinol.*, **138**, 443–448.

Land, E. J., Ramsden, C. A. and Riley, P. A. (2004) *Methods Enzymol.*, **378**, 88–109.

Lanham-New, S. A. (2008) *Proc. Nutr. Soc.*, **67**, 163–176.

Lee, H., Park, M. T., Choi, B. H., *et al.* (2011) *PLoS One*, **6**(6), e21533. DOI: 10.1371/journal.pone.0021533.

Lee, W., Yoon, G., Hwang, Y. R., *et al.* (2012) *BMB Rep.*, **45**(3), 141–146.

Li, W. W., Heinze, J. and Haehnel, W. (2005) *J. Am. Chem. Soc.*, **127**, 6140–6141.

Lim, E. S., Rhee, Y. H., Park, M. K., *et al.* (2007) *Ann. N. Y. Acad. Sci.*, **1095**, 7–18.

Lim, D., Ikeda, A., Vu, K. K., *et al.* (2011) *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.*, **879**, 3592–3598.

Lima, N. M., Correia, C. S., Leon, L. L., *et al.* (2004) *Mem. Inst. Oswaldo Cruz*, **99**, 757–761.

Lin, H. J., Lai, C. C., Lee Chao, P. D., *et al.* (2007) *J. Ocul. Pharmacol. Ther.*, **23**, 152–171.

Lin, Y. J., Zhen, Y. Z., Zhao, Y. F., *et al.* (2011) *Am. J. Chin. Med.*, **39**, 817–825.

Lind, C. (1985) *Arch. Biochem. Biophys.*, **240**, 226–235.

Liu, H., Zhao, F., Yang, R., *et al.* (2009) *Phytochemistry*, **70**(6), 773–778.

Liu, A., Chen, H., Wei, W., *et al.* (2011a) *Oncol. Rep.*, **26**, 81–89.

Liu, W., Zheng, Z., Liu, X., *et al.* (2011b) *J. Pharm. Biomed. Anal.*, **54**, 1157–1162.

Lo, T. C., Nian, H. C., Chiu, K. H., *et al.* (2012) *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.*, **15**, 893–894.

Lu, Y., Zhang, J. and Qian, J. (2008) *Cancer Biother. Radiopharm.*, **23**, 222–228.

Lu, H. F., Wang, H. L., Chuang, Y. Y., *et al.* (2010) *Neurochem. Res.*, **35**, 390–398.

Lu, L., Qin, A., Huang, H., *et al.* (2011) *Eur. J. Pharmacol.*, **658**, 242–247.

Luo, S. H., Liu, Y., Hua, J., *et al.* (2012) *Org. Lett.*, **14**(16), 4146–4149.

Magee, P. S. (2000) *Quant. Struct.-Act. Relat.*, **19**, 22–28.

Mallakin, A., McConkey, B. J., Miao, G., *et al.* (1999) *Ecotoxicol. Environ. Saf.*, **43**, 204–212.

Mallakin, A., Dixon, D. G. and Greenberg, B. M. (2000) *Chemosphere*, **40**, 1435–1441.

Maruo, S., Kuriyama, I., Kuramochi, K., *et al.* (2011) *Bioorg. Med. Chem.*, **19**, 5803–5812.

Massaoka, M. H., Matsuo, A. L., Figueiredo, C. R., *et al.* (2012) *PLoS One*, **7**(6), e38698.

McGaw, L. J., Lall, N., Hlokwe, T. M., *et al.* (2008) *Biol. Pharm. Bull.*, **31**, 1429–1433.

de Melo, J. G., Santos, A. G., de Amorim, E. L., *et al.* (2011) *Evid Based Complement Alternat Med.*, ID 365359, 14.

Meng, G., Liu, Y., Lou, C., *et al.* (2010) *Br. J. Pharmacol.*, **161**, 1628–1644.

Miao, X. S., Song, P., Savage, R. E., *et al.* (2008) *Drug Metab. Dispos.*, **36**, 641–648.

Mishra, B. B., Kishore, N., Tiwari, V. K., *et al.* (2010) *Fitoterapia*, **81**, 104–107.

Mohanlall, V., Steenkamp, P. and Odhav, B. (2011) *J. Med. Plants Res.*, **5**(14), 3132–3141.

Monks, T. J. and Jones, D. C. (2002) *Curr. Drug Metab.*, **3**, 425–438.

Müller, T., Johann, L., Jannack, B., *et al.* (2011) *J. Am. Chem. Soc.*, **133**, 11557–11571.

Nishiyama, T., Ohnuma, T., Inoue, Y., *et al.* (2008) *Biochem. Biophys. Res. Commun.*, **371**, 247–250.

Nishiyama, T., Izawa, T., Usami, M., *et al.* (2010) *Biochem. Biophys. Res. Commun.*, **394**, 459–463.

Ohsaki, Y., Shirakawa, H., Miura, A., *et al.* (2010) *J. Nutr. Biochem.*, **21**, 1120–1126.

Olson, R. E., Chao, J., Graham, D., *et al.* (2002) *Br. J. Nutr.*, **87**, 543–553.

Osman, C. P., Ismail, N. H., Ahmad, R., *et al.* (2010) *Molecules*, **15**, 7218–7226.

Pandey, A. K. and Ojha, V. (2012) *Indian J. Pharm. Sci.*, **73**(2), 216–219.

Panichayupakaranant, P., Charoonratana, T. and Sirikatitham, A. (2009a) *J. Chromatogr. Sci.*, **47**(8), 705–708.

Panichayupakaranant, P., Sakunpak, A. and Sakunphueak, A. (2009b) *J. Chromatogr. Sci.*, **47**(3), 197–200.

Paramapojn, S., Ganzera, M., Gritsanapan, W., *et al.* (2008) *J. Pharm. Biomed. Anal.*, **47**, 990–993.

Park, M. Y., Kwon, H. J. and Sung, M. K. (2009) *Biosci. Biotechnol. Biochem.*, **73**, 828–832.

Park, M. Y., Kwon, H. J. and Sung, M. K. (2011) *Life Sci.*, **88**, 486–492.

Parry, J. D., Pointon, A. V., Lutz, U., *et al.* (2009) *Chem. Res. Toxicol.*, **22**, 717–725.

Patterson, A. V., Barham, H. M., Chinje, E. C., *et al.* (1995) *Br. J. Cancer*, **72**, 1144–1150.

Patterson, A. V., Saunders, M. P., Chinje, E. C., *et al.* (1997) *Br. J. Cancer*, **76**, 1338–1347.

Peng, J., Jackson, E. M., Babinski, D. J., *et al.* (2010) *J. Nat. Prod.*, **73**(9), 1590–1592.

Portes, J. D., Netto, C. D., da Silva, A. J., *et al.* (2011) *Vet. Parasitol.*, **186**, 261–269, DOI: org/10.1016/j.vetpar.2011.11.008.

Ptitsyn, L. R., Nomura, K., Sklyar, I. V., *et al.* (2011) *Fitoterapia*, **82**(8), 1285–1289.

Pucaj, K., Rasmussen, H., Møller, M., *et al.* (2011) *Toxicol. Mech. Methods*, **21**, 520–532.

Qian, G., Leung, S. Y., Lu, G., *et al.* (2008) *J. Pharm. Pharmacol.*, **60**, 107–113.

Radhakrishnan, N., Gnanamani, A., Prasad, N. R., *et al*. (2012) *Int. J. Radiat. Biol.*, **88**(8), 575–582.

Raspotnig, G., Leutgeb, V., Schaider, M., *et al*. (2010) *J. Chem. Ecol.*, **36**, 158–162.

Rauwald, H. W. (1990) *PZ Wiss.*, **3**, 169–181.

Rukachaisirikul, V., Trisuwan, K., Sukpondma, Y., *et al*. (2008) *Arch. Pharm. Res.*, **31**(1), 17–20.

Saadawi, S., Jalil, J., Jasamai, M., *et al*. (2012) *Molecules*, **17**(5), 4824–4835.

Sagratini, G., Cristalli, G., Giardinà, D., *et al*. (2008) *J. Sep. Sci.*, **31**(6-7), 945–952.

Sakunphueak, A. and Panichayupakaranant, P. (2010) *Phytochem. Anal.*, **21**, 444–450.

Salem, M. L. (2005) *Int. Immunopharmacol.*, **5**, 1749–1770.

Samuni, Y., Ishii, H., Hyodo, F., *et al*. (2010) *Free Radic. Biol. Med.*, **48**, 1559–1563.

Sand, J. M., Hafeez, B. B., Jamal, M. S., *et al*. (2012) *Carcinogenesis*, **33**, 184–190.

Sasaki, K., Abe, H. and Yoshizaki, F. (2002) *Biol. Pharm. Bull.*, **25**, 669–670.

Savage, R. E., Hall, T., Bresciano, K., *et al*. (2008) *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.*, **872**, 148–153.

Sharma, N., Sharma, U. K., Malik, S., *et al*. (2008) *J. Sep. Sci.*, **31**, 629–635.

Sharma, N., Sharma, U. K., Gupta, A. P., *et al*. (2009) *J. Sep. Sci.*, **32**(18), 3239–3245.

Shearer, M. J., Mallinson, C. N., Webster, G. R., *et al*. (1972) *Br. J. Haematol.*, **22**, 579–588.

Shearer, M. J., McBurney, A. and Barkhan, P. (1974) *Vitam. Horm.*, **32**, 513–542.

Sheng, X., Wang, M., Lu, M., *et al*. (2011) *Am. J. Physiol. Endocrinol. Metab.*, **300**, E886–893.

Shukla, N., Kumar, M., Akanksha, A. G., *et al*. (2010) *Nat. Prod. Commun.*, **5**, 427–430.

Siegel, D., Yan, C. and Ross, D. (2011) *Biochem. Pharmacol.*, **83**, 1033–1040, DOI: org/10.1016/j.bcp.2011.12.017.

Singh, H., Moore, R. E. and Scheuer, P. J. (1967) *Experientia*, **23**, 624–626.

Sittie, A. A., Lemmich, E., Olsen, C. E., *et al*. (1999) *Planta Med.*, **65**, 259–261.

Skladanowski, A. and Konopa, J. (1994) *Biochem. Pharmacol.*, **47**, 2269–2278.

Snodgrass, R. G., Collier, A. C., Coon, A. E., *et al*. (2010) *J. Biol. Chem.*, **285**, 19068–19075.

Song, Y. and Buettner, G. R. (2010) *Free Radic. Biol. Med.*, **49**, 919–962.

Song, R., Xu, L., Xu, F., *et al*. (2010) *J. Chromatogr. A*, **1217**, 7144–7152.

Sturdy, M., Krunic, A., Cho, S., *et al*. (2010) *J. Nat. Prod.*, **73**, 1441–1443.

Subramaniya, B. R., Srinivasan, G., Sadullah, S. S., *et al*. (2011) *PLoS One*, **6**(4), e18695. DOI: 10.1371/journal.pone.0018695.

Sun, J. and McKallip, R. J. (2011) *Leuk. Res.*, **35**, 1402–1408.

Tabolacci, C., Lentini, A., Mattioli, P., *et al*. (2010) *Life Sci.*, **87**, 316–324.

Tampo, Y. and Yonaha, M. (1996) *Arch. Biochem. Biophys.*, **334**, 163–174.

Tan, Z. J., Li, F. F. and Xu, X. L. (2012) *Bioprocess Biosyst. Eng.*, **36**(8), 1105–1113. [Epub ahead of print].

Thomson, R. H. (1987) *Naturally Occurring Quinines. III. Recent Advances*, Chapman & Hall, London, UK.

Thomson, R. H. (1991) *Pharm. Weekbl. Sci.*, **13**, 70–73.

Thomson, R. H. (1997) *Naturally Occurring Quinines. IV. Recent Advances*, Blackie Academic & Professional, Chapman & Hall, London, UK.

Tsuchida, A., Itoi, T., Kasuya, K., *et al*. (2011) *Hepatogastroenterology*, **58**, 290–297.

Tu, H. Y., Huang, A. M., Teng, C. H., *et al*. (2011) *Bioorg. Med. Chem.*, **19**, 5670–5678.

Vainchtein, L. D., Rosing, H., Mirejovsky, D., *et al*. (2008) *Rapid Commun. Mass Spectrom.*, **22**, 462–470.

Van Ommen, B., Koster, A., Verhagen, H., *et al*. (1992) *Biochem. Biophys. Res. Commun.*, **189**, 309–314.

Vasconcellos, M. C., Bezerra, D. P., Fonseca, A. M., *et al*. (2011) *Melanoma Res.*, **21**, 106–114.

Vaughan, P. P., Novotny, P., Haubrich, N., *et al*. (2010) *Photochem. Photobiol.*, **86**, 1327–1333.

Venkatesh, G., Ramanathan, S., Mansor, S. M., *et al*. (2007) *J. Pharm. Biomed. Anal.*, **43**, 1546–1551.

Venkatesh, G., Majid, M. I., Ramanathan, S., *et al*. (2008) *Biomed. Chromatogr.*, **22**, 535–541.

Wahab, O. M., Ayodele, A. E. and Moody, J. O. (2010) *J. Pharm. Phyt.*, **2**(5), 64–70.

Walash, M. I., Belal, F., El-Enany, N., *et al*. (2011) *J. Chromatogr. Sci.*, **49**, 495–501.

Wallin, R., Schurgers, L. and Wajih, N. (2008) *Thromb. Res.*, **122**, 411–417.

Wamer, W. G., Vath, P. and Falvey, D. E. (2003) *Free Radic. Biol. Med.*, **34**(2), 233–242.

Wang, C., Zhang, D., Ma, H., *et al*. (2007a) *Eur. J. Pharmacol.*, **577**, 58–63.

Wang, S. L., Han, J. F., He, X. Y., *et al*. (2007b) *Drug Metab. Dispos.*, **35**, 176–179.

Wang, J., Li, H., Jin, C., *et al*. (2008) *J. Pharm. Biomed. Anal.*, **47**, 765–770.

Wang, Z., Liu, T., Gan, L., *et al*. (2010) *Eur. J. Pharmacol.*, **643**, 211–217.

Weber, P. (2001) *Nutrition*, **17**, 880–887.

Wei, X., Jiang, J. S., Feng, Z. M., *et al*. (2008) *Chem. Pharm. Bull.*, **56**, 1248–1252.

Wijeratne, E. M., Paranagama, P. A., Marron, M. T., *et al*. (2008) *J. Nat. Prod.*, **71**, 218–222.

Wittebolle, V., Lemriss, S., La Morella, G., *et al*. (2006) *Mycoses*, **49**, 169–175.

Workman, P. (1994) *Oncol. Res.*, **6**, 461–475.

Wu, J., Chien, C. C., Yang, L. Y., *et al*. (2011a) *Chem. Biol. Interact.*, **193**, 3–11.

Wu, Y. Y., Zhu, L., Ma, X. Y., *et al*. (2011b) *Pharm. Biol.*, **49**, 531–538.

Wu, M. M., Wang, L. Q., Hua, Y., *et al*. (2011c) *Planta Med.*, **77**(5), 481–484.

Xiang, W., Song, Q. S., Zhang, H. J., *et al*. (2008) *Fitoterapia*, **79**, 501–504.

Xiao, Y., Wang, Y., Gao, S., *et al*. (2011) *J. Chrom. B.*, **879**, 1833–1838.

Xu, K. H. and Lu, D. P. (2010) *Leuk. Res.*, **34**, 658–665.

Xu, F., Liu, Y., Zhang, Z., *et al*. (2008) *J. Pharm. Biomed. Anal.*, **47**, 586–595.

Xu, L., Chan, C. O., Lau, C. C., *et al*. (2011) *Phytochem. Anal.*, DOI: 10.1002/pca.1331.

Xue, X., You, J. and He, P. (2008) *J. Chromatogr. Sci.*, **46**, 345–350.

Xue, J., Ding, W. and Liu, Y. (2010) *Fitoterapia*, **81**, 173–177.

Yan, C., Kepa, J. K., Siegel, D., *et al*. (2008) *Mol. Pharmacol.*, **74**, 1657–1665.

Yao, Y., Zhang, Y. W., Sun, L. G., *et al*. (2012) *Apoptosis*, **17**(8), 832–841.

Zhang, C., Ondeyka, J. G., Zink, D. L., *et al*. (2008) *J. Nat. Prod.*, **71**(7), 1304–1307.

Zhang, L. J., Liu, H. K., Hsiao, P. C., *et al*. (2011) *J. Agric. Food Chem.*, **59**(4), 1131–1137.

Zhang, Y., Chen, X., Qin, S., *et al*. (2006) *Biol. Pharm. Bull.*, **29**, 2523–2527.

Zhao, Y., Qin, F., Boyd, J. M., *et al*. (2010) *Anal. Chem.*, **82**, 4599–4605.

Zuo, Y., Wang, C., Lin, Y., *et al*. (2008) *J. Chromatogr. A*, **1200**, 43–48.

FURTHER READINGS

Bakasso, S., Lamien-Meda, A., Lamien, C. E., *et al*. (2007) *Pak. J. Biol. Sci.*, **11**, 1429–1435.

Bellomo, G., Mirabelli, F., Vairetti, M., *et al*. (1990) *J. Cell Physiol.*, **143**, 118–128.

Dey, P. M. and Harborne, J. B. (1989), in *Methods in Plant Biochemistry*, ed. J. B. Harborne, Academic Press, London *Plant phenolics*, pp. 452–791.

Gong, X., Gutala, R. and Jaiswal, A. K. (2008) *Vitam. Horm.*, **78**, 85–101.

Inbaraj, J. J. and Chignell, C. F. (2004) *Chem. Res. Toxicol.*, **17**, 55–62.

Miura, T., Shinkai, Y., Jiang, H. Y., *et al*. (2011) *Chem. Res. Toxicol.*, **24**, 559–567.

Paroni, R., Faioni, E. M., Razzari, C., *et al*. (2009) *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.*, **877**, 351–354.

Zhang, F. and Bartels, M. J. (2004) *Rapid Commun. Mass Spectrom.*, **18**, 1809–1816.

Xanthones from Marine-Derived Microorganisms: Isolation, Structure Elucidation, and Biological Activities

Madalena M. M. Pinto[1], Raquel A. P. Castanheiro[1] and Anake Kijjoa[2]

[1]*Centro de Química Medicinal da Universidade do Porto (CEQUIMED-UP), CIIMAR, and Departamento de Química, Faculdade de Farmácia, Universidade do Porto, Porto, Portugal* and [2]*ICBAS-Instituto de Ciências Biomédicas de Abel Salazar and CIIMAR, Universidade do Porto, Porto, Portugal*

1 INTRODUCTION

The marine ecosystem comprises more than 70% of the Earth's surface, representing 95% of the biosphere (Jimeno *et al.*, 2004; Glaser and Mayer, 2009), being distinguished by an enormous biodiversity, mirrored by a variety of secondary metabolites found in animals, plants, and microorganisms (König *et al.*, 2006). To date, approximately 22,500 products of marine origin have been described, and hundreds of new compounds are being discovered every year. A number of bioactive compounds have been isolated from marine invertebrates such as sponges, tunicates, mollusks, and bryozoans, in addition to algae and marine microorganisms, such as cyanobacteria and fungi (Kijjoa and Sawangwong, 2004; Saleem *et al.*, 2007; Xu, 2011). Interestingly, these microorganisms are found to be in close contact with marine invertebrates, which are either symbiotic, that is, associated, or serve as food. Because many secondary metabolites isolated from marine invertebrates exhibited striking structural similarities to

those from the microbial origin, it was suggested that microorganisms are at least involved in their biosynthesis or are in fact the true sources of these respective metabolites. Even though there are some evidences for this hypothesis, it is extremely difficult to definitively state the biosynthetic source of many marine natural products because of the complexity of associations in marine organisms (König *et al.*, 2006).

Owing to their pharmacological potential, either directly as drugs or as lead structures for molecular modifications and/or drug synthesis especially in cancer research, marine natural products have aroused attention of many researchers (Blunt *et al.*, 2009). Many marine-derived natural products have already undergone clinical or preclinical trials especially with regard to, not only their antitumor properties, but also anti-inflammatory and antiinfectious properties, among others (Kosta, Jain, and Tiwari, 2008). In fact, two marine natural products have already been approved in human therapeutics: ziconotide (Prialt®), a peptide first isolated from the venom of the cone snail *Conus magus*, an analgesic used for treatment

of patients suffering from chronic pain, and trabecte-din (Yondelis®), an alkaloid originally isolated from a marine tunicate and now obtained by semisynthesis, for the treatment of soft tissue sarcomas and ovarian cancer (Jimeno *et al*., 2004; Kijjoa and Sawangwong, 2004; Baker *et al*., 2007; Glaser and Mayer, 2009).

From microorganisms described as marine nat-ural sources of the bioactive metabolites, fungi have shown great potential as suggested by the diversity of their isolated secondary metabolites. Fungi from the marine habitats are divided into obligate and facultative marine species, the former being restricted to the marine environment and the latter occurring also in freshwater or terres-trial localities or both (Bugni and Ireland, 2004). Isolation of a fungal strain from a marine sample does not prove that this fungus is actively living in the marine environment because it is possible to isolate a terrestrial fungus as being a contaminant in the marine habitat. Because most fungi isolated from marine samples are not proven to be obligate or facultative species, the more general expression "marine-derived fungi" is largely used (Bugni and Ireland, 2004). Marine-derived fungi have been recognized as a potential source of novel structures and biologically potent molecules, and a growing number of marine-derived fungi have been reported as sources of novel bioactive secondary metabolites (Bugni and Ireland, 2004; Saleem *et al*., 2007). Because of their particular living conditions, salinity, nutrition, higher pressure, temperature variations, and competition with bacteria, viruses, and other fungi, marine-derived fungi may have developed specific secondary metabolic pathways compared with their terrestrial counterparts (Bugni and Ireland, 2004; Saleem *et al*., 2007). In terms of the over-all number of secondary metabolites, polyketides, prenylated polyketides (meroterpenes), peptides, and alkaloids are the most described. With no excep-tion, xanthone derivatives that can be considered as originated from polyketide or prenylated polyketide have been largely isolated from marine-derived fungi and, to a lesser extent, from other microorganisms.

Xanthones, or 9*H*-xanthen-9-ones, encompass an important class of oxygenated heterocyclic com-pounds with the dibenzo-γ-pirone skeleton as a basic structure (Pinto, Sousa, and Nascimento, 2005) (Figure 1). The biological activities of these com-pounds are associated with their tricyclic scaffold but vary depending on the nature and/or position of the
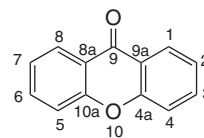
Figure 1　Xanthone's basic structure and its numbering system.

substituents. Natural xanthones can be subdivided, depending on the nature of the substituents in the dibenzo-γ-pirone scaffold, into simple oxygenated xanthones, glycosylated xanthones, prenylated xan-thones and their derivatives, xanthone dimers, xan-thonolignoids, and miscellaneous (Pinto, Sousa, and Nascimento, 2005; Vieira and Kijjoa, 2005; El-Seedi *et al*., 2009). Recently, a new classification based on structural characteristics of xanthones (monomers and dimers/heterodimers) and the level of oxidation of the xanthone C-ring (fully aromatic-, dihydro-, tetrahydro-, and hexahydroxanthone derivatives) was suggested by Masters and Bräse (2012).

Xanthones can be described as "privileged struc-tures" as they are molecular scaffolds that appear to be capable of binding to multiple targets and, consequently with appropriate structure modifica-tions, could exhibit multiple activities (Pinto, Sousa, and Nascimento, 2005; Pinto and Castanheiro, 2009; Pouli and Marakos, 2009). This class of compounds can interact with a large variety of biological targets exhibiting important activities, for instance, cancer chemopreventive and antitumor, influence on several inflammatory mediators belonging to the arachidonic acid cascade, to achieve a variety of enzymes such as kinases, proteases, and monoaminoxidases (MAOs A and B). Moreover, they exhibit antimicrobial activ-ity against a large number of human pathogenic microorganisms. Consequently, xanthonic molecules are a distinguished structural type of value to the discovery of new pharmaceutically interesting com-pounds (Pinto, Sousa, and Nascimento, 2005; Pinto and Castanheiro, 2009; Pouli and Marakos, 2009). Xanthones occur mainly in higher plants, fungi, and lichens. However, the number of xanthones isolated from marine sources, mainly from sponges- and algae-associated fungi, and form marine bacteria is not yet very high when compared with other classes of compounds.

2 MARINE SOURCES OF XANTHONES

2.1 Marine-Derived Fungi

2.1.1 From the Genus Aspergillus (Trichocomaceae)

Marine invertebrate-associated microorganisms have recently attracted attention as an important source of novel, biologically active secondary metabolites. Among these, some *Aspergillus* spp. have been reported to produce a considerable number of cytotoxic compounds and other bioactive substances (Rateb and Ebel, 2011).

In the search for new and biologically active marine natural products, Wu *et al*. have isolated more than 300 fungal strains from marine mollusks and algae (Wu, Ouyang, and Tan, 2009). Screening of extracts from cultures of selected isolates led to the discovery of several strains that showed significant antimicrobial activity in agar diffusion assays. A new difuranoxanthone, asperxanthone (**1**) (Figure 2), along with other compounds, was then isolated from a selected marine fungus of the seawater, identified as *Aspergillus* sp. (MF-93), collected in the Quan-Zhou Gulf (Wu, Ouyang, and Tan, 2009; Rateb and Ebel, 2011; Masters and Bräse, 2012). Although asperxanthone (**1**) was the most active compound in inhibiting multiplication of tobacco mosaic virus (TMV) (inhibitory rates 62.9%), its activity was lower than that of the methanol extract from which it was isolated (inhibitory rates 81.2%). These results suggested that the inhibitory activity of the methanol extract was not primarily due to any of the isolated compounds (Wu, Ouyang, and Tan, 2009; Rateb and Ebel, 2011). Lee *et al*., in their continuing search for bioactive metabolites, have isolated another two xanthones: sterigmatocystin (**2**) and dihydrosterigmatocystin (**3**), along with other compounds, from the fungus *Aspergillus versicolor* isolated from a marine sponge *Petrosia* sp. (Petrosiidae) by bioactivity-guided fractionation (Lee *et al*., 2010). Sterigmatocystin (**2**) and dihydrosterigmatocystin (**3**) were evaluated for their cytotoxicity against five human tumor cell lines, A-549 (human lung adenocarcinoma), SK-OV-3 (human ovarian adenocarcinoma), SK-MEL-2 (human skin melanoma), XF-498 (central nervous system cancer), and HCT-15 (human colon adenocarcinoma). Sterigmatocystin (**2**) was found to exhibit significant cytotoxicity against all the cell lines tested, having IC_{50}
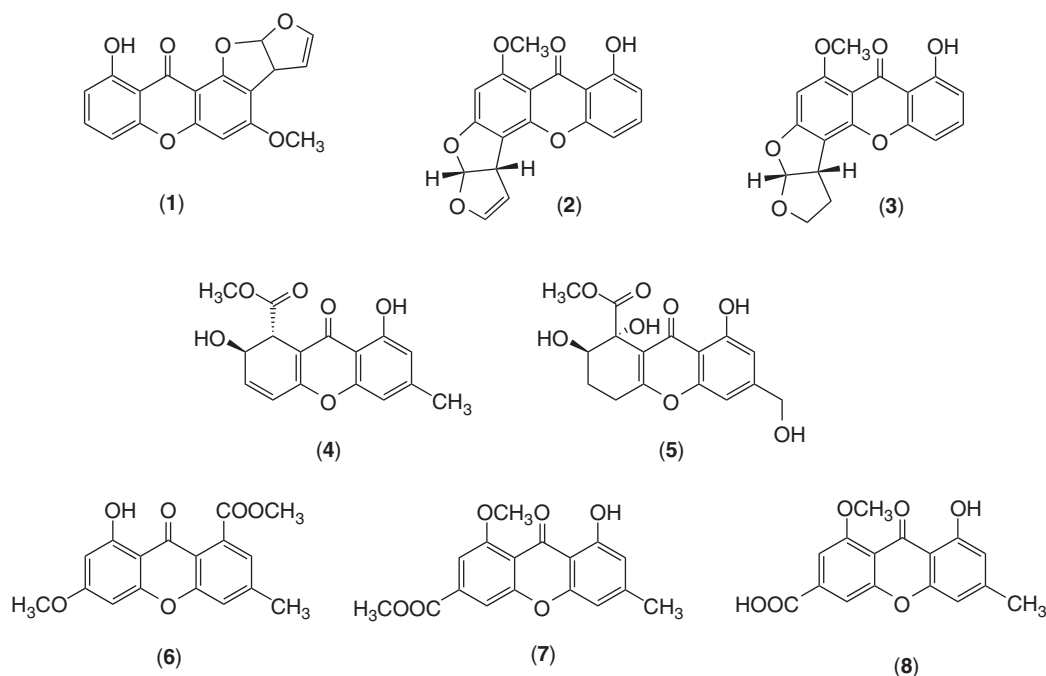
Figure 2 Xanthones of marine-derived fungi from the genus *Aspergillus*.

values in the range of 1.22-4.61 µg mL^{-1}. Considering the only structural difference between sterigmatocystin (**2**) and dihydrosterigmatocystin (**3**), it is believed that the allylic ether moiety (a double bond at C-2′) could play an important role in the cytotoxicity of sterigmatocystin (**2**) (Lee *et al*., 2010). Trisuwan *et al*. reported isolation of two new hydrogenated xanthone derivatives, aspergillusones A (**4**) and B (**5**) (Figure 2), along with other known compounds, from the ethyl acetate extract of the culture broth of *Aspergillus sydowii* PSU-F154, isolated from a gorgonian sea fan of the genus *Annella* (Subergorgiidae), and neither of the isolated compounds showed antioxidant activity in the DPPH assay (Trisuwan *et al*., 2011). Sun *et al*., in their ongoing program to discover new bioactive compounds from algicolous fungi, reported isolation of three new xanthone derivatives, yicathin A (**6**), yicathin B (**7**), and yicathin C (**8**), from the endophytic fungus *Aspergillus wentii*, obtained from the inner tissue of the marine red alga *Gymnogongrus flabelliformis* Harvey (Phyllophoraceae). The structures of the three xanthones were unambiguously established by nuclear magnetic resonance (NMR) and mass spectroscopic methods, as well as by quantum chemical calculations (Sun *et al*., 2013). Xanthones **6–8** were evaluated for their antibacterial (against *Escherichia coli* and *Staphylococcus aureus*) and antifungal (against phytopathogens *Colletotrichum lagenarium* and *Fusarium oxysporum*) activities, using standard agar diffusion test at 10 µg/disk. Yicathin B (**7**) was active against *E. coli* (inhibition diameter 9 mm) and yicathin C (**8**) was found to be able to inhibit *E. coli* (12.0 mm), *S. aureus* (7.5 mm), and *C. lagenarium* (11.0 mm). In addition, all the three xanthones were shown to exhibit weak brine shrimp (*Artemia salina*) toxicity with IC$_{50}$'s of 0.20, 0.22, and 0.30 µmol mL^{-1}, respectively (Sun *et al*., 2013).

2.1.2 From the Genus Penicillium (Trichocomaceae)

Some *Penicillium* spp. are a rich source of secondary metabolites, and a number of xanthones have been reported from the members of this genus. Shao *et al*. described isolation of two new xanthones, 8-(methoxycarbonyl)-1-hydroxy-9-oxo-9*H*-xanthene-3-carboxylic acid (**9**) and dimethyl 8-methoxy-9-oxo-9*H*-xanthene-1, 6-dicarboxylate (**10**) along with the known xanthone methyl 8-hydroxy-6-methyl-9-oxo-9*H*-xanthene-1-carboxylate (**11**) (Figure 3), from the culture broth of the mangrove endophytic fungus *Penicillium* sp. (ZZF 32#), isolated from the bark of *Acanthus ilicifolius* Linn. (Acanthaceae) collected from the South China Sea. Interestingly, the crude extract of the culture broth and the column fractions containing these xanthones were found to exhibit significant cytotoxicity against KB (human epidermoid carcinoma of the nasopharynx) and KBv200 (multidrug-resistant human epidermoid carcinoma of the nasopharynx) cells with IC$_{50}$ values of 1.5 and 2.5 µg mL^{-1}, respectively (cisplatin was used as the positive control with IC$_{50}$ value against KB and KBv200 of 0.56 and 0.78 µg mL^{-1}, respectively), whereas neither of the isolated compounds was active. Thus, the bioactivity demonstrated by the crude extract and its column fractions was not due to the individual xanthone constituents (IC$_{50}$ > 50 µg mL^{-1}). Furthermore, antifungal activity assay revealed that xanthone **10** exhibited modest activity against *F. oxysporum* f. sp. *cubense* with the minimal inhibitory concentration (MIC) value of 12.5 µg mL^{-1} (Shao *et al*., 2008; Rateb and Ebel, 2011; Xu, 2011). Another endophytic fungus *Penicillium* sp., isolated from *Melia azedarach* L. (Meliaceae) growing in a Chinese mangrove habitat, also afforded 7-hydroxyjanthinone (**12**) (Figure 3), which was found to be devoid of cytotoxic activity (Rateb and Ebel, 2011). Recently, Khamthong *et al*. (2012a,b) reported isolation of three xanthone derivatives: sydowinin A (**13**), pinselin (**14**), and conioxanthone A (**15**) (Figure 3), from the fungus *Penicillium citrinum* PSU-F51, isolated from the gorgonian sea fan (*Annella* sp.). These xanthones were found to be inactive in the antifungal assay against *Microsporum gypseum* at a concentration of 200 µg mL^{-1}. Another endophytic fungus, *Penicillium sacculum*, isolated from the halophyte *Atriplex* sp. (Amaranthaceae), was recently found to afford a new xanthone whose structure was established as 1-hydroxy-3-methoxy-6-sulfo-8-methylxanthone (**16**) (Figure 3) (Liu *et al*., 2012a,b). During the study to evaluate the methods of isolation and growth of marine-derived fungal strains in artificial media for the production of secondary metabolites, Kossuga *et al*. (2012) have isolated norliquexanthone or 1,3,6-trihydroxy-8-methyl-9*H*-xanthen-9-one (**17**) from *Penicillium raistrikii*, which was isolated from
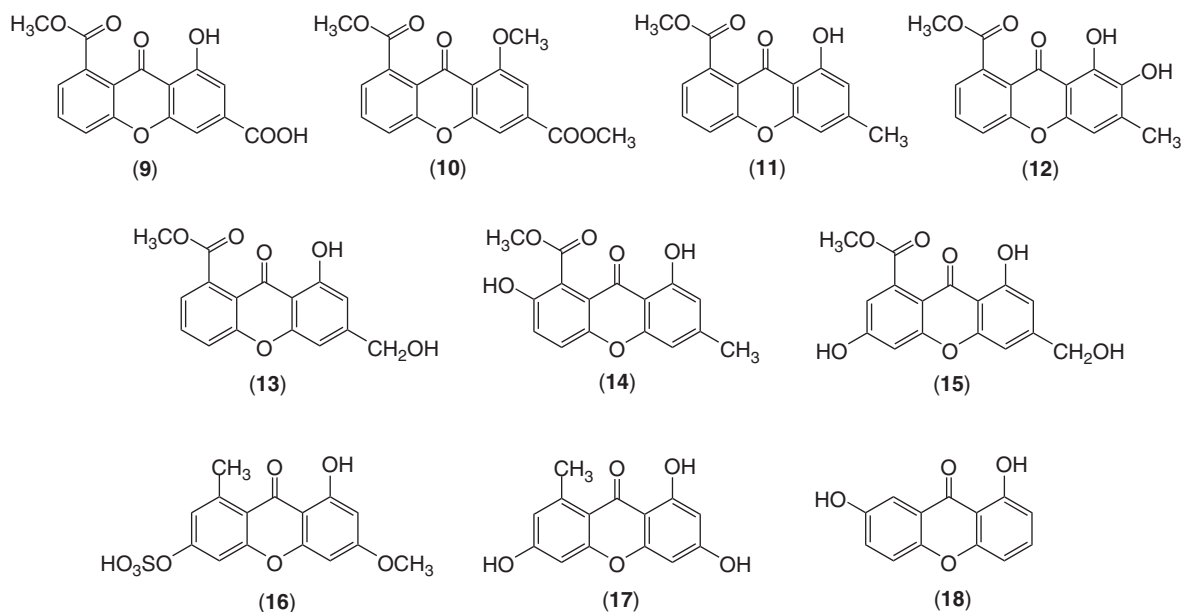
Figure 3   Xanthones of marine-derived fungi from the genus *Penicillium*.

the marine sponge *Axinella* cf. *corrugata* (Axinellidae) (Figure 3). Huang *et al*. reported isolation of 1,7-dihydroxy-9*H*-xanthen-9-one (**18**), from the mangrove endophytic fungus *Penicillium* sp. ZH16, isolated from the leaves of the mangrove tree *Avicennia* (Acanthaceae) from the South China Sea coast (Figure 3) (Huang *et al*., 2012).

2.1.3   From the Genus Monodictys (Dematiaceae)

The uncommon marine-derived fungus *Monodictys putredinis*, isolated from an unidentified green alga collected in Tenerife, Spain, was investigated for its secondary metabolites after cultivation on a solid biomalt medium (Pontius *et al*., 2008a,b). The separation steps of vacuum liquid chromatography (VLC) and high performance liquid chromatography (HPLC) resulted in isolation of two novel dimeric chromanones, monodictyochromone A (**19**) and monodictyochromone B (**20**) (Figure 4), whose structure consists of two unusually modified xanthone-derived subunits, being the only difference between the two compounds the site of connection of subunits I and II. Compounds **19** and **20** were also examined for their cancer chemopreventive potential

and were found to inhibit cytochrome P450 1A activity with IC$_{50}$ values of 5.3 and 7.5 µ*M*, respectively. In addition, both compounds were shown to display moderate activity as inducers of NAD(P)H: quinone reductase (QR) in cultured mouse Hepa 1c1c7 cells, with CD values (concentration required to double the specific activity of QR) of 22.1 and 24.8 µM, respectively. Furthermore, compound **19** was slightly less potent than compound **20** as aromatase inhibitor, showing IC$_{50}$ values of 24.4 and 16.5 µM, respectively (Pontius *et al*., 2008a,b). Krick *et al*. also reported isolation of, besides a benzophenone, four monomeric xanthones, monodictysin A–C (**21–23**) and monodictyxanthone (**24**) (Figure 4), from the extract of the same fungus. These compounds were subsequently tested in a series of *in vitro* bioassays relevant for the inhibition of carcinogenesis *in vivo* in order to evaluate their cancer chemopreventive potential (Krick *et al*., 2007; Rateb and Ebel, 2011). Xanthone **23** was shown to inhibit cytochrome P450 1A activity with an IC$_{50}$ value of 3.0 µ*M*. Besides, xanthones **21–23** were also identified as inhibitors of the Cyp1A isoenzyme, which is involved in the metabolic conversion of procarcinogens into carcinogens. Xanthones **22** and **23** displayed moderate activity as inducers of NAD(P)H:QR, a carcinogen-detoxifying
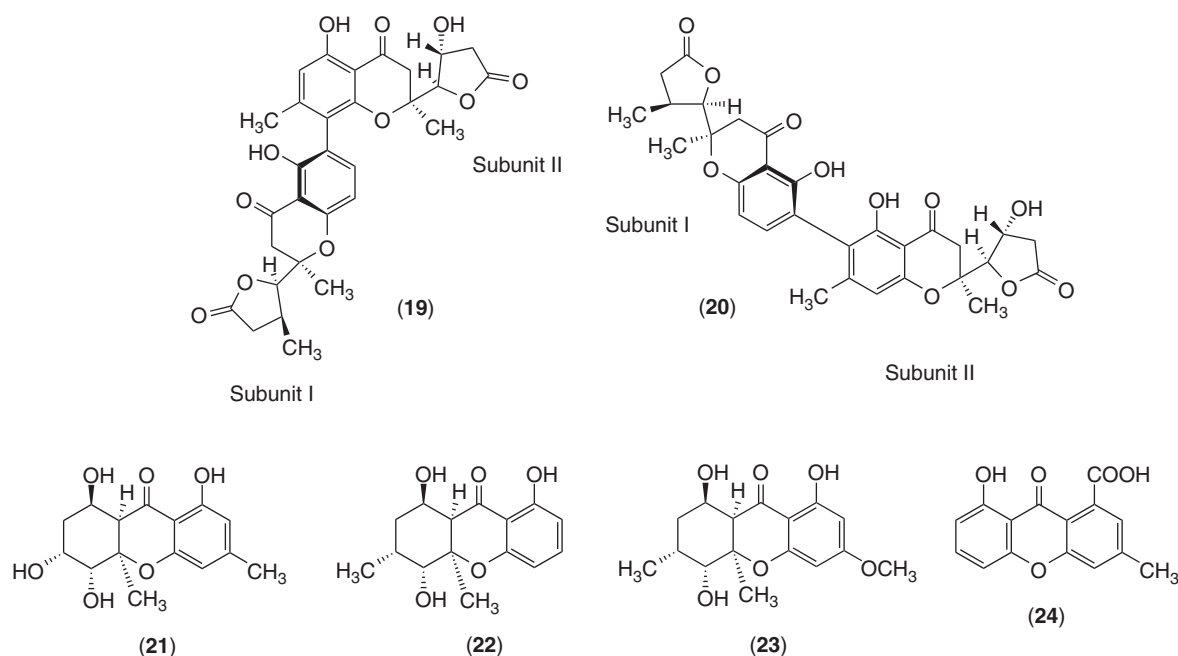
Figure 4   Xanthones of marine-derived fungi from the genus *Monodictys*.

enzyme, in cultured mouse Hepa 1c1c7 cells, with CD values of 12.0 and 12.8 μM, respectively. Compound **23** was a weak inhibitor of aromatase activity essential for the biosynthesis of estrogens. Although the xanthone-carboxylic acid **24** showed a dose-dependent Cyp1A activity inhibition with an IC_{50} value of $34.8 \pm 7.4 \mu M$, it did not induce QR activity. Overall, the substitution pattern of the core structure was found to strongly influence the biological effects of these compounds (Krick *et al.*, 2007).

2.1.4   From the Genus Emericella (Trichocomaceae)

The genus *Emericella* is one of the *Aspergillus* anamorphs, and the members of this genus are found to biosynthesize a remarkable diversity of secondary metabolites with motivating biological properties and, thus, representing potential leads for the developing of new pharmaceutical agents (Figueroa *et al.*, 2009). Thus, within the scope of a program aiming at the discovery of novel calmodulin (CaM)-inhibitors, useful as pesticide or drug leads, Figueroa *et al.*

reported isolation and the CaM inhibitor properties of two new xanthones: 15-chlorotajixanthone hydrate (**25**) and 14-methoxytajixanthone (**26**), together with the known shamixanthone (**27**) and tajixanthone hydrate (**28**) (Figure 5) from *Emericella* sp. strain 25379, isolated from the surface of a coral species collected on the Mexican Pacific coast (Figueroa *et al.*, 2009). The effect of xanthones **25–28** on CaM was initially assessed with the calmodulin-sensitive cAMP phosphodiesterase (PDE1) assay, which is commonly used to detect CaM antagonists; a human recombinant-CaM was used as the activator. The results showed that the activation of PDE1 was inhibited in the presence of **26** and **28** in a concentration-dependent manner. The effect of xanthones **26** ($IC_{50} = 5.54 \pm 1.28 \mu M$) and **28** ($IC_{50} = 5.62 \pm 1.25 \mu M$) was comparable to that of chlorpromazine (CPZ; $IC_{50} = 7.26 \pm 1.60 \mu M$), a well-known CaM inhibitor used as a positive control. A kinetic analysis using different amounts of CaM in the presence of different concentrations of **26** and **28** indicated that both xanthones acted as competitive antagonists of CaM, thus interfering with the formation of the CaM–PDE1 active complex. The estimated K_i (inhibition constant) values
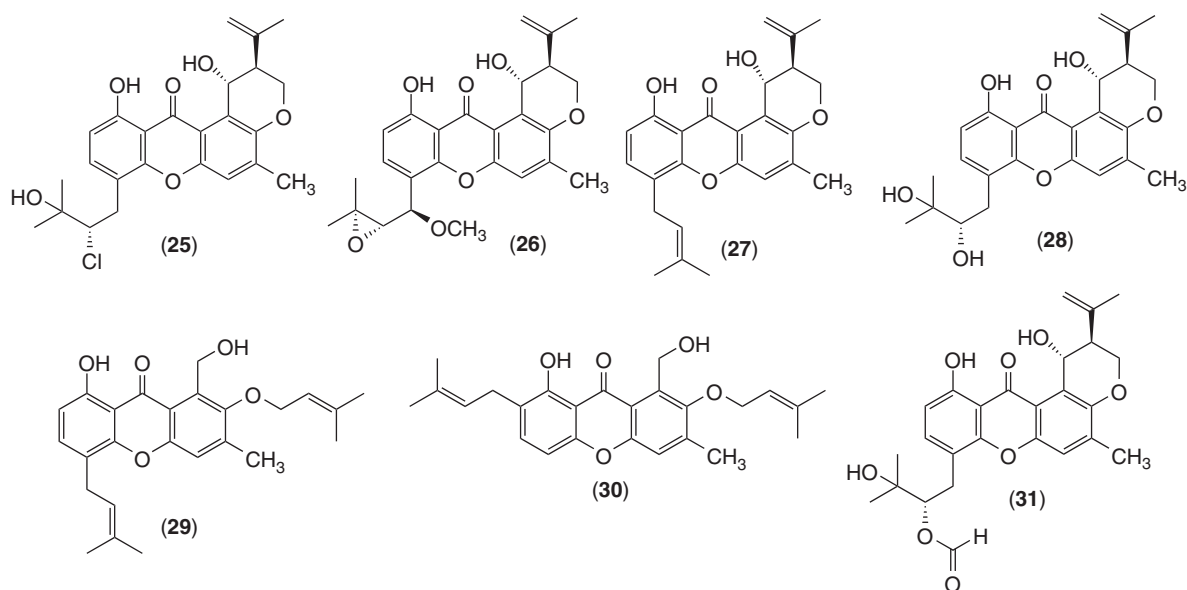
Figure 5 Xanthones of marine-derived fungi from the genus *Emericella*.

were 25.38 ± 2.26 and $13.92 \pm 2.29\ \mu M$, respectively. AutoDock predictions suggested that these compounds interact with the protein at the same binding site of trifluoropiperazine (TFP), a recognized CaM inhibitor. The CaM antagonist effect of **28** might be related with its mild cytotoxic action and other pharmacological properties yet to be discovered (Figueroa *et al.*, 2009). The results of their investigation led the authors to conclude that *Emericella* sp. contained novel type of competitive CaM inhibitors. The xanthones sterigmatocystin (**2**), shamixanthone (**27**), emericellin (**29**) (Figure 5), and the biosynthetically related benzoquinone derivatives were also isolated from the culture of *Emericella nidulans* var. *acristata*, an endophyte isolated from a Mediterranean green alga collected around Sardinia (Kralj *et al.*, 2006). The effects of the crude extract and the pure compounds (except for sterigmatocystin) on tumor growth *in vitro* were investigated in a survival and proliferation assay using a panel of 36 human tumor cell lines representing 11 different tumor types. Antitumor activity was defined as test/control value smaller than 50% compared to the untreated control cells. The crude extract effected antitumor activity in all 36 cell lines (100%) at $50\ \mu g\ mL^{-1}$, in 31 of the 36 cell lines (86%) at $5\ \mu g\ mL^{-1}$, and in 2 of 36 (6%) cell lines at $0.5\ \mu g\ mL^{-1}$. This is indicative of a selective and concentration-dependent antitumor activity of the extract and one or more of its components. However, shamixanthone (**27**) and emericellin (**29**) showed either only marginal or no antitumor activity *in vitro* (Kralj *et al.*, 2006). Bringmann *et al.* also reported isolation of shamixanthone (**27**) and isoemericellin (**30**) (Figure 5), along with other metabolites from *Emericella variecolor*, derived from the marine sponge *Haliclona valliculata* (Chalinidae). Malmstrøm *et al.* isolated, besides other metabolites, shamixanthone (**27**), tajixanthone hydrate (**28**), and varixanthone (**31**) (Figure 5), from the marine-derived strain M75-2 of *E. variecolor*. Varixanthone (**31**) was found to show no cytotoxic activity, at $1\ \mu g\ mL^{-1}$, against three tumor cell lines (P388, mouse lymphoma; A549, human lung carcinoma; and HT29, human colon carcinoma). However, it displayed antimicrobial activity against Gram-positive and Gram-negative bacteria, and it was also found to be active against *E. coli*, *Proteus* sp., *Bacillus subtillis*, and *S. aureus*, showing a MIC of $12.5\ \mu g\ mL^{-1}$ in all these cases but showed lower potency against *Enterococcus faecalis* (MIC = $50\ \mu g\ mL^{-1}$) (Malmstrøm *et al.*, 2002; Thomas, Kavlekar, and LokaBharathi, 2010; Bhatnagar and Kim, 2012; Masters and Bräse, 2012).
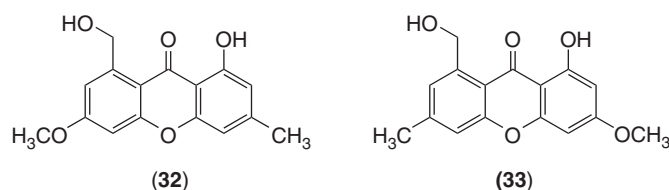
Figure 6   Xanthones of marine-derived fungi from the genus *Phoma*.
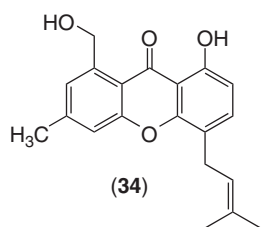
Figure 7   Xanthones of marine-derived fungi from the genus *Paecilomyces*.

2.1.5   *From the Genus Phoma (Didymellaceae)*

Pan *et al*. described isolation of two new xanthones: 1-hydroxy-8-(hydroxymethyl)-6-methoxy-3-methyl-9*H*-xanthen-9-one (**32**) and 1-hydroxy-8-(hydroxymethyl)-3-methoxy-6-methyl-9*H*-xanthen-9-one (**33**) (Figure 6), from a mangrove endophytic fungus *Phoma* sp. SK3RW1M., isolated from the roots of the mangrove tree *Avicennia marina* (Forsk.) Vierh., collected in Shankou mangrove in China (Pan *et al*., 2010; Rateb and Ebel, 2011). This was the first report on xanthone derivatives isolated as secondary metabolites from *Phoma* species [19]. Xanthones **32** and **33** were evaluated for their cytotoxic activity against KB and KBv200 cells, and preliminary results indicated that they were inactive (Pan *et al*., 2010; Rateb and Ebel, 2011).

2.1.6   *From the Genus Paecilomyces (Trichocomaceae)*

In the course of an ongoing search for natural potent antitumor agents from marine mangrove fungi, Wen *et al*. have found that the extract of the fungus *Paecilomyces* sp., a metatrophic fungus Tree1-7 collected from the bark of a mangrove from the Taiwan Strait, exhibited good cytotoxicity. Chromatographic purification of the methanol extract of the mycelia led to isolation of a new xanthone, paeciloxanthone (**34**) (Figure 7), together with emodin and chrysophanol (Wen *et al*., 2008; Rateb and Ebel, 2011; Xu, 2011; Simpson, 2012). Paeciloxanthone (**34**) was found to exhibit an *in vitro* cytotoxicity against hepG2 (IC$_{50}$ = 1.08 μg mL^{-1}), acetylcholine esterase (AChE) inhibitory (IC$_{50}$ = 2.25 μg mL^{-1}) activity, and antimicrobial activity against *Curvularia lunata* (walker) Boedijn, *E. Coli*., and *Candida albicans*, affording inhibitory zones of 6, 12, and 10 mm, respectively (Wen *et al*., 2008; Rateb and Ebel, 2011; Xu, 2011).

2.1.7   *From the Genus Chaetomium (Chaetomiaceae)*

Investigations of the marine-derived fungus *Chaetomium* sp., by Pontius *et al*., led to the isolation of the new natural products chaetoxanthones A, B, and C (**35–37**) (Figure 8). Chaetoxanthones A (**35**) and B (**36**) are dioxane/tetrahydropyran substituted xanthones whose structures are rarely found in natural products, whereas chaetoxanthone B (**37**) is a tetrahydropyran-substituted chlorinated xanthone. Chaetoxanthones A, B, and C (**35–37**) were tested in a series of *in vitro* bioassays for their antiprotozoal and cytotoxic activities. Chaetoxanthone B (**36**) exhibited selective antiprotozoal activity toward *Plasmodium falciparum* with an IC$_{50}$ value of 0.5 μg mL^{-1} (reference drug chloroquine, IC$_{50}$ = 0.08 μg mL^{-1}) and no cytotoxic effects toward L6-cells (IC$_{50}$ > 90 μg mL^{-1}) and 35 tumor cell lines (mean IC$_{50}$ > 10 μg mL^{-1}) (Pontius *et al*., 2008a,b). These results let Pontius *et al*. to suggest that the xanthone scaffold was a suitable pharmacophore for antiplasmodial activity, and the nature and position of substituents clearly influence the biological
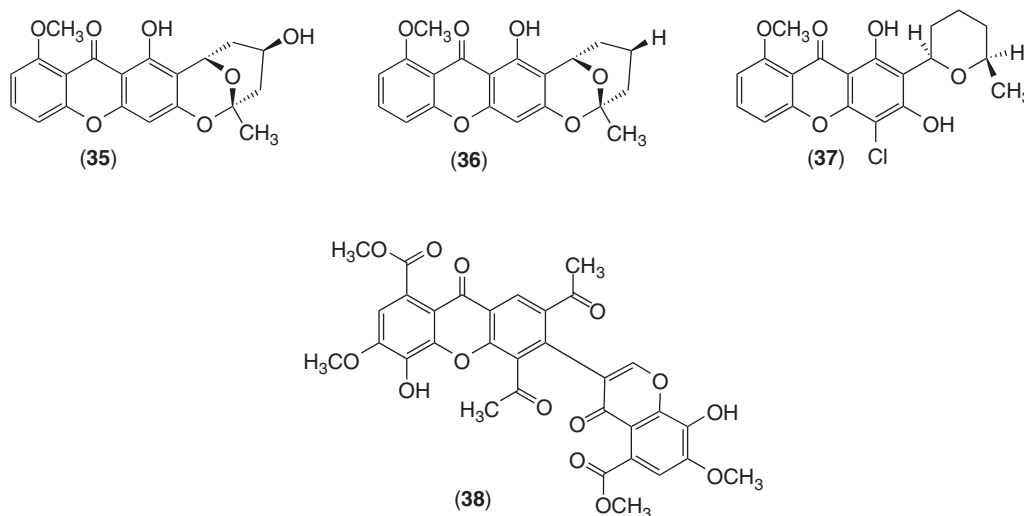
Figure 8 Xanthones of marine-derived fungi from the genus *Chaetomium*.

effect (Pontius *et al*., 2008a,b). Although chaetoxanthone B (**36**) did not reach the high antiplasmodial potency of some reported polyhydroxyxanthones or alkylamino-substituted xanthones with IC_{50} values below $0.1 \, \mu g \, mL^{-1}$, it was more active than prenylated, nonnitrogenous xanthone derivatives. Consequently, they concluded that the heterocyclic substitution in chaetoxanthone B (**36**) did not cause a remarkable enhancement in activity. Interestingly, chaetoxanthone A (**35**), with an additional hydroxyl group at C-3′ compared to chaetoxanthone B (**36**), had a much weaker activity toward *P. falciparum* ($IC_{50} = 3.5 \, \mu g \, mL^{-1}$) (Pontius *et al*., 2008a,b). Conversely, chaetoxanthone C (**37**) was found to be moderately active against *Trypanosoma cruzi*, the causative pathogen of Chagas disease, with an IC_{50} value of $1.5 \, \mu g \, mL^{-1}$ (reference drug benznidazole, $IC_{50} = 0.3 \, \mu g \, mL^{-1}$) without having considerable cytotoxic effects on L6-cells ($IC_{50} = 46.7 \, \mu g \, mL^{-1}$). The antiprotozoal activity of chaetoxanthone C (**37**) is comparable to that reported for a series of synthetic chlorinated xanthones carrying aminoalkyl side chains. However, these compounds were tested only toward *P. falciparum*, and the best candidate showed an *in vitro* IC_{50} value of $1.4 \, \mu g \, mL^{-1}$. Thus, the *in vitro* antiprotozoal activity of xanthones was further substantiated by this study. However, it can be inferred that substitution with a dioxane/tetrahydropyran moiety on the xanthone nucleus does not improve antiprotozoal activity

(Pontius *et al*., 2008a,b). Another xanthone isolated from the culture broth of the marine-derived fungus *Chaetomium* sp., obtained from an undisclosed marine alga, was chaetocyclinone C (**38**) (Figure 8). This xanthone was found to be active toward selected phytopathogenic fungi but was not cytotoxic (Rateb and Ebel, 2011).

2.1.8 From the Genus Wardomyces (Microascaceae)

The culture of the marine fungal isolate *Wardomyces anomalus* Brooks and Hansford, isolated from the green alga *Enteromorpha* sp. (Ulvaceae) collected in the Baltic Sea, afforded two new xanthone derivatives, 2,3,6,8-tetrahydroxy-1-methylxanthone (**39**) and 2,3,4,6,8-pentahydroxy-1-methylxanthone (**40**), in addition to the known xanthone 3,6,8-trihydroxy-1-methylxanthone (**41**) (Figure 9) (Abdel-Lateff *et al*., 2003; Masters and Bräse, 2012). Xanthone **39** was found to have significant DPPH radical scavenging effects (94.7%, at $25.0 \, \mu g \, mL^{-1}$) besides its capacity to inhibit peroxidation of linolenic acid (17.0%, at $7.4 \, \mu g \, mL^{-1}$). Furthermore, the total extract and xanthones **39** and **41** were shown to be inhibitors of TK p56[lck] tyrosine kinase but only minor antimicrobial activity was observed for xanthones **39–41** in agar diffusion assay (Abdel-Lateff *et al*., 2003).
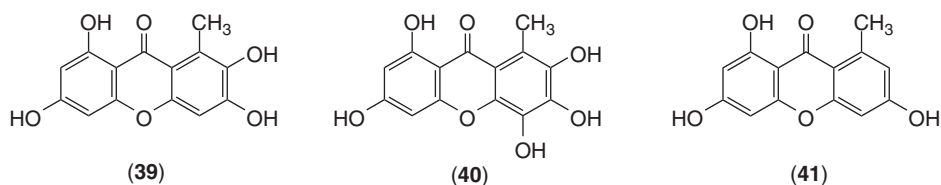
Figure 9 Xanthones of marine-derived fungi from the genus *Wardomyces*.

2.1.9 From Other Mangrove Endophytic Fungi

Marine fungi, especially mangrove endophytic fungi, have proved to be an abundant source for novel natural compounds. In the search for new metabolites from marine mangrove endophytic fungi, Huang *et al.* have reported isolation of two new xanthone derivatives, 1,7-dihydroxy-2-methoxy-3-(3-methylbut-2-enyl)-9*H*-xanthen-9-one (**42**) and 1-hydroxy-4,7-dimethoxy-6-(3-oxobutyl)-9*H*-xanthen-9-one (**43**) (Figure 10), from the fungus *Phomopsis* sp. (No. ZH19) (Diaporthaceae), isolated from the leaves of the mangrove tree *A. marina* collected from the South China Sea coast (Huang *et al.*, 2010b). The same authors reported also the isolation of a new xanthone derivative, 3,5,8-trihydroxy-2,2-dimethyl-3,4,4-trihydro-2*H*, 6*H*-pyrano[3,2-*b*]xanthen-6-one (**44**), together with the known 5,8-dihydroxy-2,2-dimethyl-2*H*,6*H*-pyrano[3,2-*b*]xanthen-6-one (**45**) (Figure 10), from an unidentified species of the endophytic fungus ZSU-H16 from the mangrove tree *Avicennia* from the South China Sea coast (Huang *et al.*, 2010a). Using the MTT cytotoxicity assay method, it was found that both xanthones **42** and **43** were able to inhibit the growth of KB cells (human epidermoid carcinoma of the nasopharynx) with IC_{50} values of 20 and 35 $\mu mol\,mL^{-1}$, and $KB_V 200$ cells with IC_{50} values of 30 and 41 $\mu mol\,mL^{-1}$, respectively, whereas xanthone **44** exhibited weak cytotoxicity against KB and KB_V 200 cells having IC_{50} values greater than 50 $\mu g\,mL^{-1}$ (Huang *et al.*, 2010b). During a screening for novel structures from marine-derived mangrove endophytic fungus from the South China Sea, Zhu and Lin have isolated sterigmatocystin (**2**), dihydrosterigmatocystin (**3**), and secosterigmatocystin (**46**) from the unidentified species of the fungal isolate 1850 from a leaf of *Kandelia candel* (L.) Druce (Rhizophoraceae) from an estuarine mangrove in Hong Kong. A preliminary bioassay showed

that sterigmatocystin (**2**) exhibited weak cytotoxic activity against tumor cell lines Bel-7402 (human hepatoma cell line) and NCI-H460 (human non small cell lung cancer cell line) with IC_{50} values of 96.53 and 72.52 $\mu g\,mL^{-1}$, respectively. All the three xanthones did not exhibit significant inhibitory activity against human DNA topoisomerase type (hTOP), showing IC_{50} values greater than 100 $\mu g\,mL^{-1}$ (Zhu and Lin, 2007). Sterigmatocystin (**2**) was also isolated together with 5-methoxysterigmatocystin (**47**) (Figure 10) from another unidentified species of the mangrove endophytic fugal strain ZSUH-36, obtained from the mangrove tree *A. ilicifolius* (Shao *et al.*, 2007; Xu, 2011).

Secalonic acid D (SAD) (**48**) (Figure 10), one of the most prominent mycotoxins first characterized as a metabolite of *Penicillium oxalicum* in 1970, was found to induce cleft palate via inhibiting G1/S-phase-specific CDK2 activity. SAD was later reisolated from several sources, including the endophytic fungus *Paecilomyces* sp. (tree 1–7) and a mangrove-associated unidentified strain No. ZSU44. It was also considered as an acutely toxic and teratogenic fungal metabolite. SAD (**48**) also displayed extraordinarily significant cytotoxicity and induced apoptosis in leukemia cells HL60 and K562 with IC_{50} values of 0.38 and 0.43 $\mu mol\,L^{-1}$, respectively. Its inhibitory effect on human topoisomerase I was assessed with IC_{50} at 0.16 $\mu mol\,mL^{-1}$. The study of the mechanism of action showed that SAD (**48**) caused cell cycle arrest at G1, maintained via the GSK-3β/β-catenin/c-Myc pathway (Xu, 2011). Hong also reported isolation of SAD (**48**) from the fermentation broth of the marine lichen-derived fungus *Gliocladium* sp. T31 (Hypocreaceae), collected from marine sediments in South Pole, in addition to its ability to inhibit mammalian DNA topoisomerase I via *in vitro* plasmid supercoil relaxation assay and EMSA (Hong, 2011). SAD (**48**) was found to exhibit a considerable inhibition on DNA topoisomerase I in a dose-dependent manner with the MIC
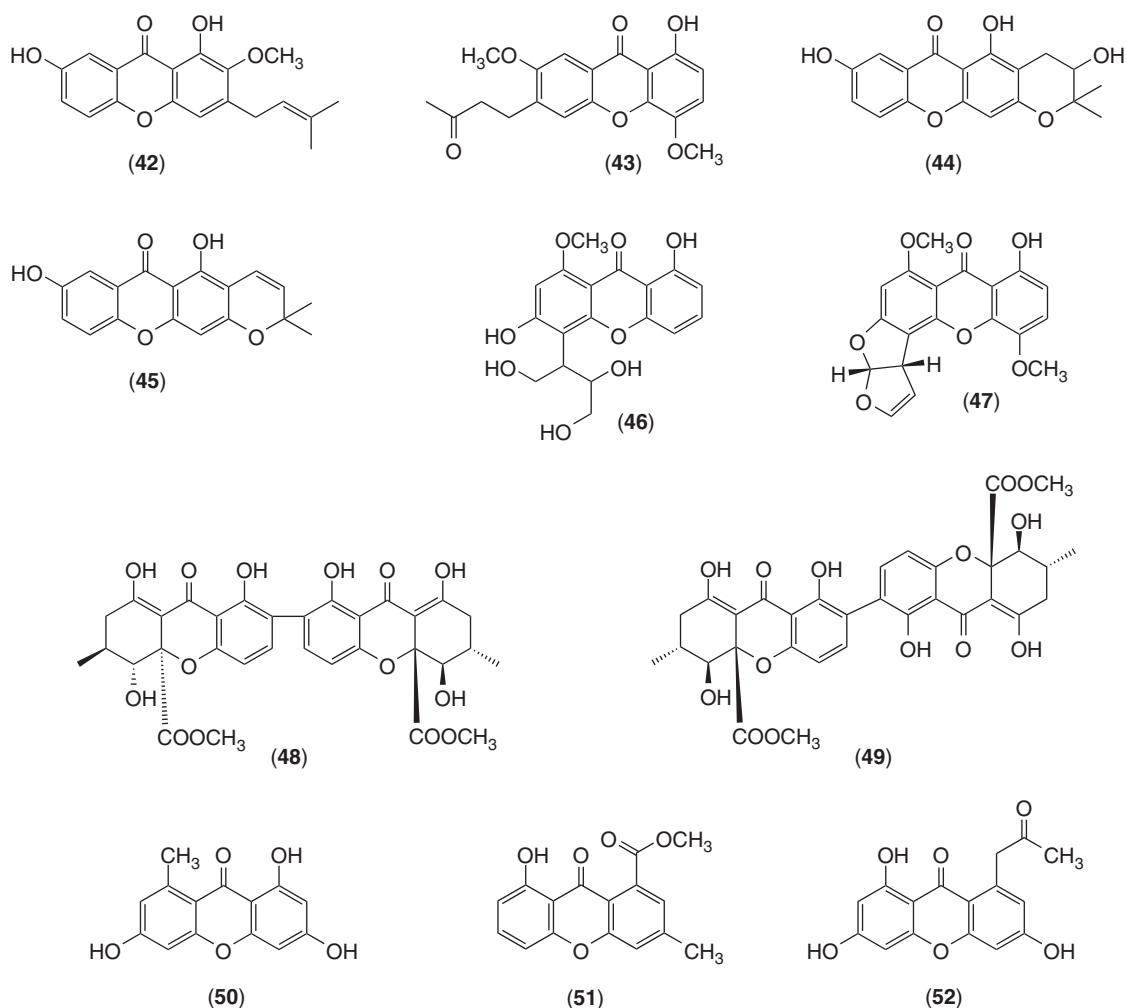
Figure 10 Xanthones from other mangrove endophytic fungi.

of 0.4 μ*M*. Unlike the prototypic DNA topoisomerase I poison camptothecin (CPT), SAD (**48**) inhibited the binding of topoisomerase I to DNA but did not induce the formation of topoisomerase I–DNA covalent complexes. This was the first study reporting the inhibitory effect of SAD (**48**) on mammalian DNA topoisomerase I. The ability of SAD (**48**) to inhibit the proliferation of several tumor cells and to interfere with DNA topoisomerase I suggested its potential as an anticancer candidate (Hong, 2011).

Liu *et al.* reported isolation of a xanthone dimer skyrin (**49**) and norlichexanthone (**50**) (Figure 10) from the extract of the endophytic fungus

Talaromyces sp. ZH-154, from the mangrove tree *K. candel*. Both xanthones showed significant antimicrobial activity against *Pseudomonas aeruginosa* (Xu, 2011). Li *et al.* isolated a new xanthone derivative, 8-hydroxy-3-methyl-9-oxo-9*H*-xanthene-1-carboxylic acid methyl ether (**51**) (Figure 10), from the coculture broth of two mangrove fungi (strain No. K38 and E33) collected from the South China Sea coast. This xanthone was evaluated for its antifungal activity against five representative fungi: *Gloeasporium musae*, *Blumeria graminearum*, *F. oxysporum*, *Peronophthora cichoralearum*, and *Collectotrichum glocosporioides*) using the disk assay

method. Xanthone **51** was found to exhibit broad inhibitory activity against these microorganisms, especially *G. musae* and *P. cichoralearum* (Li *et al.*, 2011).

Another fungal genus that produces many types of bioactive metabolites is *Trichoderma* (Hypocreaceae). Khamthong *et al.*, in their effort to search for biologically active compounds from a marine-derived fungi, reported isolation of trichodermaxanthone (**52**) from the broth extract of the marine-derived fungus *Trichoderma aureoviride* PSU-F95 (Figure 10) (Khamthong *et al.*, 2012a,b).

2.2 From Marine-Derived Actinomycetes

2.2.1 From the Genus Actinomadura (Thermomonosporaceae)

Rodríguez *et al.* reported isolation of the new natural polycyclic xanthone IB-00208 (**53**) (Figure 11) from the fermentation broth of the actinomycete *Actinomadura* sp., isolated from the northern coast of Spain. IB-00208 (**53**) showed a potent cytotoxic activity against several tumor cell lines of both human: A-549 (human lung adenocarcinoma), HT-29 (human colon adenocarcinoma), and SK-MEL-28 (human melanoma); and murine: P-388 (murine leukemia). It also showed a good antibacterial activity against Gram-positive organisms (*S. aureus*, *B. subtilis*, *Micrococcus luteus*), but poor activity against Gram-negative bacteria (*E. coli*, *Klebsiella pneumonia*, *P. aeruginosa*) (Malet-Cascón *et al.*, 2003; Rodríguez *et al.*, 2003; Masters and Bräse, 2012).
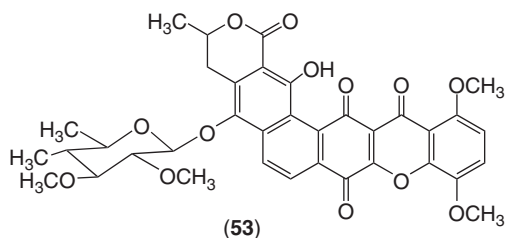
Figure 11 The polycyclic xanthone IB-00208 (**53**) from *Actinomadura* sp.

2.2.2 From the Genus Streptomyces (Streptomycetaceae)

With the aim of discovering novel antibacterial natural products from marine bacteria, Liu *et al.* isolated four new polycyclic antibiotics: citreamicin θ A (**54**), citreamicin θ B (**55**), citreaglycon A (**56**), and dehydrocitreaglycon A (**57**) (Figure 12), from marine-derived *Streptomyces caelestis*, isolated from the coastal water of the Red Sea near Jeddah.

All four compounds were found to display antibacterial activity against *Staphylococcus haemolyticus*, *S. aureus*, and *B. subtillis*. Citreamicin θ A (**54**), citreamicin θ B (**55**), and citreaglycon A (**56**) also exhibited low MIC values of 0.25, 0.25, and 8.0 μg mL^{-1}, respectively, against methicillin-resistant *Staphylococcus aureus* (MRSA) ATCC 43300. Moreover, citreamicin θ A (**54**) and citreamicin θ B (**55**) were found to exhibit significant cytotoxic activity against HeLa cells (human cervical cancer) with IC$_{50}$ values of 0.055 and 0.072 μg mL^{-1}, respectively (Liu *et al.*, 2012a,b).

3 EXTRACTION, ISOLATION, AND PURIFICATION OF XANTHONES FROM MARINE-DERIVED MICROORGANISMS

3.1 Extraction of Xanthones from the Culture Media

The majority of xanthones have been isolated from the marine organisms-associated fungi and few of them were reported from the marine actinomycetes (Rodríguez *et al.*, 2003; Liu *et al.*, 2012a,b). Consequently, the methods of extraction of these compounds depended on the type of the culture media. When the fungi were cultured in solid media, the fungal biomass and media were first homogenized and extracted with organic solvents such as acetone or ethyl acetate. The group of König has used the solid media such as malt-yeast agar (MYA) medium (Pontius *et al.*, 2008a,b) and biomalt medium (Krick *et al.*, 2007; Kralj *et al.*, 2006; Abdel-Lateff *et al.*, 2003) to culture the algicolous fungi to produce the bioactive secondary metabolites. The fungal biomass and the media were then homogenized using Ultra-Turrax apparatus, and the mixture was extracted exhaustively with ethyl acetate to produce crude extracts. Ueda *et al.* used the solid medium containing brown
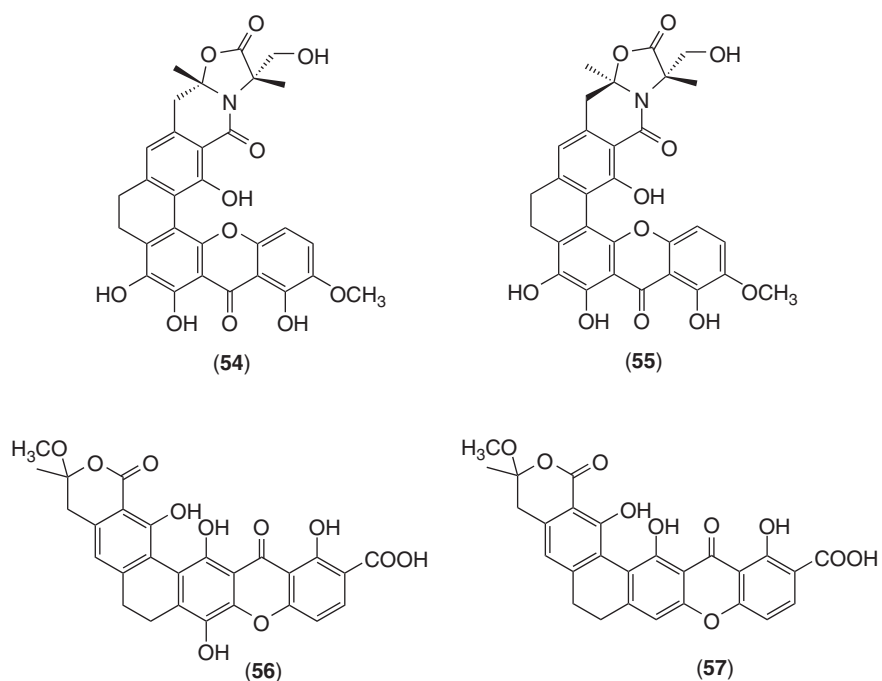
Figure 12 Xanthones from the marine-derived actinomycetes *Streptomyces caelestis*.

rice, bacto-yeast extract, sodium tartrate, potassium hydrogen phosphate, and water, for a static culture of the marine sponge-derived fungus *Tritirachium* sp. SpB081112MEf2. The solid culture was then extracted with 80% aqueous acetone, and after concentration *in vacuo*, the aqueous concentrate was extracted with ethyl acetate to give a crude ethyl acetate extract (Ueda, Takagi, and Shin-ya, 2010). Malmstrøm *et al.* used the yeast extract sucrose (YES) medium to culture the marine-derived strain of the fungus *E. variecolor* to investigate its secondary metabolites (Malmstrøm *et al.*, 2002). Then, mycelium and agar were harvested and extracted with a mixture of ethyl acetate, chloroform, and methanol (3 : 2 : 1) containing 1% of formic acid in a Stomacher bag. After evaporation of the solvent, the dried extract was split into two fractions by partition between aqueous methanol solution and hexane to give the hexane and methanol crude extracts.

On the other hand, when the fungi were cultured in liquid media, the mycelia were first separated from the culture broth by filtration before extraction by organic solvents. Bringmann *et al.* used the WSA liquid medium to culture the fungus *E. variecolor* and isolated from the marine sponge *H. valliculata*

(Bringmann *et al.*, 2003). After separation from the culture medium, the fungus was extracted exhaustively with a 1 : 1 mixture of dichloromethane and methanol, whereas the medium was extracted three times with ethyl acetate. Both extracts were then dried *in vacuo* and partitioned between aqueous methanol and petroleum ether. The methanol fraction of the medium extract was desalted by partitioning between water and ethyl acetate to give the ethyl acetate crude extract. Zhu and Li used the liquid medium containing glucose, peptone, yeast extract, and sodium chloride, to grow the marine-derived mangrove endophytic fungus (stain 1850). After incubation at 30 °C for 35 to 40 days, the mycelia were separated from the culture broth by filtration through cheese cloth. The mycelia were dried by air and extracted with methanol to give crude mycelial extract. The culture filtrate was then concentrated below 50 °C and extracted by shaking with an equal volume of ethyl acetate to give crude broth extract (Zhu and Lin, 2007). Shao *et al.* also used the same type of medium (glucose, peptone, yeast extract, and sodium chloride) and extraction process to isolate two new xanthone derivatives from the endophytic fungus *Penicillium* sp. (ZZF32#), isolated from the

bark of the mangrove tree *A. ilicifolius* from the South China Sea (Shao *et al*., 2008). Wen *et al*. used the same type of liquid medium to culture the fungus *Paecilomyces* sp. (Tree1–7), isolated from mangrove saprophytic bark from the Taiwan Strait. After 30 days of incubation at 28°C, the culture was filtered through cheese cloth to separate mycelia from the culture broth. After air drying, the mycelia were dipped in methanol, and the filtrate was concentrated *in vacuo* below 55°C and extracted five times by shaking with equal volume of ethyl acetate (Wen *et al*., 2008). Figueroa *et al*. cultured the fungus *Emericella* sp. strain 25379, isolated from the surface of a coral collected at Marietas Islands (Mexico), in a liquid medium composed of Czapek concentrate, potassium hydrogen phosphate, powdered yeast extract, and sucrose. The culture broth and the mycelia were then separated by filtration through cheese cloth. Both the culture broth and the mycelia were extracted with dichloromethane (Figueroa *et al*., 2009). The Chinese groups have also used similar liquid media and methods of extraction to study xanthone derivatives from the marine-derived fungi. Huang *et al*. used the liquid GPY medium (glucose, peptone, yeast extract, and sodium chloride) as a culture medium for the endophytic fungi, isolated from the mangrove tree *A. marina* from South China Sea. After incubation for 30 days at room temperature, the cultures were separated into mycelia and filtrate by filtration through the cheese cloth (Huang *et al*., 2010a, 2010b). The filtrate was then concentrated below 50 °C and extracted by shaking with equal volume of ethyl acetate. Collection and evaporation of ethyl acetate *in vacuo* yielded the crude extract. Pan *et al*. used not only the same liquid medium (GPY) to culture the mangrove endophytic fungus, *Phoma* sp. SK3RW1M, but also a similar method for extraction of xanthones from the culture medium (Pan *et al*., 2010). Li *et al*. also used the liquid GPY medium to coculture the marine fungi (strain Nos. E33 and K 38) and the same method of extraction of xanthones from the filtrate (Li *et al*., 2011). Trisuwan *et al*. cultured the marine-derived fungus *A. sydowii* PSU-F154, isolated from a sea fan *Annella* sp., collected from the coastal area of Southern Thailand in the potato dextrose broth at room temperature for 4 weeks. After filtration, the filtrate and mycelia were extracted with ethyl acetate to afford broth and mycelial ethyl acetate crude extracts (Trisuwan *et al*., 2011). Using the same methods of culture and extraction, Khamthong *et al*. were able to isolate several

xanthones from the culture of the marine-derived fungus *P. citrinum* PSU-F51, isolated from the gorgonian sea fan *Annella* sp., collected from the Similan islands in Southern Thailand (Khamthong *et al*., 2012a,b). Liu *et al*. cultivated the fungus *P. sacculum*, separated from the halophyte *Atriplex* sp., by shaking at 150 rpm and at 24°C in flasks containing the liquid medium composed of potato extract, peptone, yeast extract, glucose, and seawater. After 17 days, the fermented broth was filtered through cheese cloth and separated into supernatant and the mycelia. The mycelia were extracted with acetone to afford the crude extract (Liu *et al*., 2012a,b).

Besides the marine-derived fungi, only few marine-derived actinomycetes were found to produce xanthones. Rodríguez *et al*. cultured *Actinomadura* sp. (strain BL-42-PO13-046), an actinomycete isolated from the Northern coast of Spain in a shake flasks containing liquid medium composed of glucose, tryptone, calcium carbonate, sodium chloride, monobasic potassium phosphate, and distilled water at pH 7. After completion of cultivation, whole harvested broth was filtered with diatomaceous earth. The mycelial cake was then extracted with a 2 : 1 : 1 mixture of chloroform, methanol, and water. After filtration, the organic layer was concentrated under reduced pressure to give a brownish oily residue (Rodríguez *et al*., 2003). Recently, Liu *et al*. used a liquid medium containing starch, peptone, yeast extract, and sea salt to culture *S. caelestis*, a marine-derived actinomycete, collected from the coastal water of the Red Sea. The culture was maintained at 23°C for 5 days after which was filtered with eight layers of cheese cloth. The broth was then extracted with ethyl acetate, and the mycelia were extracted with acetone and methanol (1 : 2 v/v). The extracts from the broth and the mycelia were combined and partitioned between water and hexane. The resulting aqueous residue was further extracted with ethyl acetate to give the crude ethyl acetate extract (Liu *et al*., 2012a,b).

More often than not, the broth and mycelial organic extracts are directly applied on column chromatography for fractionation procedure. However, when the crude extracts are to be directly separated by HPLC, they should be cleaned up or fractionated by solid-phase extraction (SPE). SPE is a very popular technique currently available for rapid and selective sample preparation. The versatility of SPE allows use of this technique for many purposes, such as

purification, trace enrichment, desalting, derivatization, and class fractionation. The principle of SPE involves partitioning between a liquid (sample matrix or solvent with analytes) and a solid (sorbent) phase. SPE uses the same type of stationary phases as in liquid chromatography columns. Reversed phase (RP) involves a polar or moderately polar sample matrix (mobile phase) and a nonpolar stationary phase. The analyte of interest is typically mid- to nonpolar. Normal phase (NP) involves a polar analyte, a mid- to nonpolar matrix (e.g., acetone, chlorinated solvents, and hexane) and a polar stationary phase (Żwir-Ferenc and Biziuk, 2006). Abdel-Lateff *et al.* used a Bakerbond RP octadecyl C-18 SPE and a gradient elution from water to methanol to fractionate the ethyl acetate extract of the algicolous marine fungus *W. anomalus*, before purification by RPs HPLC to isolate 2,3,6,8-tetrahydroxy-1-methylxanthone (**39**), 2,3,4,6,8-pentahydroxy-1-methylxanthone (**40**), and 3,6,8-trihydroxy-1-methylxanthone (**41**) (Abdel-Lateff *et al.*, 2003).

3.2 Isolation and Purification of Xanthones from Crude Extracts

The first step of isolation of secondary metabolites involves fractionation of the crude extract to allow the compounds of similar polarities to be eluted together. Several types of column chromatography are commonly used to fulfill this purpose. The factors determining the method of choice include sample capacity and cost. Low pressure liquid chromatography (LPLC) such as column chromatography is still a major tool for fractionation of the crude fungal extracts to isolate xanthones. In this method, the mobile phase flows down through the stationary phase at an atmospheric pressure without any additional forces either by vacuum or pressure. Several types of stationary phases can be used in LPLC; however, silica gel is the most popular as it is available in relatively low cost and it can be applied to a wide range of compound classes. Malmstrøm *et al.* used column chromatography of silica gel to isolate shamixanthone (**27**) from the less polar fraction and varixanthone (**31**) and tajixanthone hydrate (**28**) from the most polar fraction of the crude extract of *E. variecolor* (Malmstrøm *et al.*, 2002). Zhu and Li also used silica gel column with gradient elution from petroleum ether to ethyl acetate as a fractionation process to isolate the xanthones sterigmatocystin

(**2**), dihydrosterigmatocystin (**3**), and secosterigmatocystin (**46**) from the ethyl acetate crude extract of a marine-derived mangrove endophytic fungus (Zhu and Lin, 2007). Similarly, Shao *et al.*, reported isolation of 8-(methoxycarbonyl)-1-hydroxy-9-oxo-9H-xanthene-3-carboxylic acid (**9**), dimethyl 8-methoxy-6-methyl-9-oxo-9H-xanthene-1-carboxylate (**10**), and methyl 8-hydroxy-6-methyl-9-oxo-9H-xanthene-1-carboxylate (**11**), using silica gel column chromatography, with gradient elution from petroleum to ethyl acetate, for fractionation of the ethyl acetate crude extract of the fungus *Penicillium* sp. (ZZF32#), followed by preparative separation on silica gel thin layer chromatography (TLC) (Shao *et al.*, 2008). Wen *et al.* also used silica gel column chromatography, with gradient elution from petroleum ether to ethyl acetate, to isolate paeciloxanthone (**34**), emodin, and chrysophanol, from the crude ethyl acetate extract of the marine mangrove fungus *Paecilomyces* sp. (Tree1-7) (Wen *et al.*, 2008). In their search for CaM inhibitors, Figueroa *et al.* have first fractionated the dichloromethane crude extract of the fungus *Emericella* sp. by silica gel open column with hexane-dichloromethane-ethyl acetate gradient, followed by further fractionation by another silica gel column or Sephadex LH-20 column (eluted with methanol) before final purification by crystallization, preparative TLC of silica gel or by RP-HPLC, to isolate four xanthones: 15-chlorotajixanthone hydrate (**25**), 14-methoxytajixanthone (**26**), shamixanthone (**27**), and tajixanthone hydrate (**28**) (Figueroa *et al.*, 2009). Huang *et al.* (2010b) reported isolation of 1,7-dihydroxy-2-methoxy-3-(3-methylbut-2-enyl)-9H-xanthen-9-one (**42**) and 1-hydroxy-4,7-dimethoxy-6-(3-oxobutyl)-9H-xanthen-9-one (**43**) from the mangrove endophytic fungus No. ZH19 and 3,5,8-trihydroxy-2,2-dimethyl-3,4,4-trihydro-2H,6H-pyrano[3,2-b]xanthen-6-one (**44**) and 5,8-dihydroxy-2,2-dimethyl-2H,6H-pyrano[3,2-b]xanthen-6-one (**45**) from the mangrove endophytic fungus No. ZSU-H16 (Huang *et al.*, 2010a), using column chromatography of silica gel and gradient elution from petroleum ether to ethyl acetate to fractionate the ethyl acetate crude extracts before purification by preparative TLC of silica gel. Pan *et al.* and Li *et al.* used the same method of fractionation of crude extracts to isolate various xanthones from the mangrove endophytic fungus *Phoma* sp. and from the coculture broth of two marine fungi, respectively (Pan *et al.*, 2010; Li *et al.*, 2011). Liu *et al.* also used the silica gel column chromatography with

gradient elution from chloroform to methanol for the initial fractionation of the crude mycelial acetone extract of the marine-derived fungus *P. sacculum*. The obtained fractions were further fractionated by Sephadex- LH 20 and then purified by HPLC, to give 1-hydroxy-3-methoxy-6-sulfo-8-methylxanthone (**16**) (Liu *et al*., 2012a,b).

Another useful stationary phase is the hydroxyl-propylated dextran gel Sephadex LH-20®, which has been widely used in separation of natural products, including xanthones. Owing to its relatively high cost, this gel is not normally used for fractionation of the crude extracts but instead for purification of fractions first obtained from the silica gel column chromatography. However, Trisuwan *et al*. and Khamthong *et al*. used the column chromatography of Sephadex LH-20, with methanol as a mobile phase, to fractionate crude extracts of the fungi *P. citrinum* and *A. sydowii* (Trisuwan *et al*., 2011; Khamthong *et al*., 2012a, 2012b).

Medium pressure liquid chromatography (MPLC) is also commonly used for preliminary fractionation of crude fungal extracts because of its lower cost and higher throughput. Ueda *et al*. used NP-MPLC, with a stepwise solvent system of *n*-hexane–ethyl acetate (3 : 1) and chloroform–methanol (49 : 1), for fractionation of the ethyl acetate crude extract of the marine sponge-derived fungus *Tritirachium* sp., before purification of the new xanthones by NP-TLC and RP-HPLC (Ueda, Takagi, and Shin-ya, 2010). Lee *et al*., in their investigation on the bioactive metabolites from the sponge-derived fungus *A. versicolor*, used a stepped-gradient (50% to 100% methanol elution) RP-MPLC for the preliminary fractionation of the crude ethyl acetate extract, followed by purification of the obtained fractions by a RP-HPLC to yield the xanthones sterigmatocystin (**2**) and dihydrosterigmatocystin (**3**), besides anthraquinone derivatives (Lee *et al*., 2010).

Vacuum liquid chromatography (VLC) is another widely used method for fractionation of crude extracts. VLC is a column chromatography to which the vacuum is applied to increase flow rate and, thus, speed up the fractionation procedure. Another advantage of VLC is that application of the extract and eluents is easily achieved owing to the open end of the column. Besides, both NP and RP silica gel can be used as stationary phase. Pontius *et al*. used NP-VLC (Merck silica gel 60, 63–200 μm, and petroleum ether–acetone–methanol gradient elution) to fractionate the crude ethyl acetate extract of the

fungal culture of *Chaetomium* sp. and the RP-VLC with methanol–water gradient to further fractionate the fractions obtained from the first column, before final purification of chaetoxanthones (**35–37**) by RP(-18)-HPLC (Pontius *et al*., 2008a, 2008b). Krick *et al*. also used VLC of silica gel 60 (63–200 μm) and a gradient elution from petroleum ether to ethyl acetate, acetone, and methanol to fractionate the ethyl acetate crude extract of the culture of the marine algicolous fungus *M. putredinis*, before purification with NP- and RP-HPLC to obtain monodictysins A, B, and C (**21–23**), monodictyxanthone (**24**), and monodictyphenone (Krick *et al*., 2007). In the same way, Kralj *et al*. used the VLC of silica gel 60 of 63–200 μm particle size and eluting with a gradient of dichloromethane–ethyl acetate and methanol for preliminary fractionation of the crude ethyl acetate extract of another algicolous fungus *E. nidulans* var. *acristata*, followed by further fractionation of the obtained fractions with Sephadex LH-20 column, to isolate several secondary metabolites including prenylated xanthones [shamixanthone (**27**) and emericellin (**29**)] (Kralj *et al*., 2006). Another type of column chromatography used for preliminary fractionation of crude extracts and coarsely separated fractions is flash chromatography (FC). In this method, nitrogen or compressed air are applied on top of the column to force the mobile phase to flush through the tightly packed stationary phase. Similar to VLC, both NP and RP silica gel can be used as stationary phase for FC. Besides being a rapid method of fractionation, the increasing peak resolution is normally obtained as it can use the stationary phase of the smaller particle size (e.g., 40 μm silica gel). Using vacuum flash chromatography (VFC) eluted with a stepwise gradient of ethyl acetate–methanol, followed by purification with LPLC of silica gel and eluted with mixtures of chloroform–methanol, Rodríguez *et al*. have isolated a cytotoxic polycyclic xanthone from the crude extract of a marine-derived *Actinomadura* (Rodríguez *et al*., 2003). Liu *et al*. used a RP C-18 FC and eluted with a gradient of water to methanol, for fractionation of the ethyl acetate fraction of the marine-derived actinomycetes *S. caelestis*. The eluted fractions from the FC were then further fractionated by a LPLC of Sephadex LH-20, using methanol as eluent before final purification by RP semipreparative HPLC, to give the antibacterial xanthone derivatives citreamicin θ A (**54**), citreamicin θ B (**55**), citreaglycon A (**56**), and dehydrocitreaglycon A (**57**) (Liu *et al*., 2012a,b).

After fractionation of crude extracts by either of the methods described above (LPLC, MPLC, VLC, and FC), these coarsely purified fractions can be further fractionated by column chromatography of either silica gel or Sephadex LH-20, to yield less complex fractions, which will be ultimately purified by crystallization, preparative TLC or HPLC, depending on the amount and nature of the target compounds in the fractions. Because many xanthones are crystalline or solid, crystallization in an appropriate mixture of organic solvents can be achieved very easily. This method is not only less costly but also yields a final product with higher degree of purity. On the contrary, if the xanthone does not exist in solid form, preparative TLC of silica gel is a method of choice as it is rapid and less expensive. As most xanthones can be detected by a ultraviolet (UV) lamp at 254 nm, it is easy to isolate them from the matrix on the TLC plate, and consequently, many marine-derived xanthones were purified by this method (Zhu and Lin, 2007; Shao *et al.*, 2008; Figueroa *et al.*, 2009; Huang *et al.*, 2010b; Ueda, Takagi, and Shin-ya, 2010; Khamthong *et al.*, 2012a,b). However, it should bear in mind that the impurities in the silica gel used for preparative TLC can compromise the purity of the final product, especially when the desirable compounds exist in a very small concentration. In this case, HPLC should be considered as a preferred method for purification and both NP- and RP-HPLC have been widely used to purify the marine-derived xanthones (Abdel-Lateff *et al.*, 2003; Bringmann *et al.*, 2003; Krick *et al.*, 2007; Pontius *et al.*, 2008a,b; Lee *et al.*, 2010; Ueda, Takagi, and Shin-ya, 2010; Liu *et al.*, 2012a,b).

4 STRUCTURE ELUCIDATION OF XANTHONES FROM MARINE-DERIVED MICROORGANISMS

In relation to xanthones isolated from terrestrial organisms (Vieira and Kijjoa, 2005), the number of xanthones isolated from marine organisms is comparatively small and almost all of them were isolated from marine organisms-associated or endophytic mangrove fungi. Both simple oxygenated and prenylated xanthones have been reported from the marine sources. Similar to their terrestrial counterparts, the structure elucidation of the marine-derived xanthones is based, mainly, on 1D (¹H, ¹³C NMR, and DEPT experiments) and 2D (COSY, HSQC,

HMBC, and NOESY) NMR spectral analysis. Silva and Pinto have summarized the use of the NMR techniques for structure assignment of xanthones and the typical proton and carbon chemical shifts of simple and prenylated xanthones (Silva and Pinto, 2005).

Generally, with the ¹H and ¹³C NMR spectra, one can decide if the compound under study is a simple oxygenated or prenylated xanthone. For simple oxygenated xanthones, their ¹H NMR spectra normally exhibit the signals of aromatic protons about δ6.2–9.0, depending on the number, nature, and positions of the substituents on the benzene rings. The number of the aromatic protons exhibited in the ¹H NMR spectrum can also determine the degree of substitution. The multiplicity and coupling constants of the aromatic protons can be useful for determination of the substitution pattern on rings A and B, and this can be confirmed by cross peaks observed in the COSY spectrum. In addition, the chemical shift values of the hydroxyl protons can also be indicative of its position. The hydrogen bonding hydroxyl group in the *peri* position adjacent to the carbonyl carbon (C-9) normally appears as a singlet around δ12–14, whereas the hydroxyl group at other positions appears as broad singlet around δ9–10.

Considering the xanthone scaffold (Figure 1), it is obvious that the ¹³C NMR spectrum is very useful to determine if the compound under investigation is a xanthone derivative because the xanthone nucleus contains one carbonyl carbon (C-9) and twelve sp^2 hybridized carbons, two of which are oxygen bearing (C-4a and C-10a). Together with DEPT 90° experiment, it is possible to deduce the degree of substitution as this spectrum exhibits only the signals of the methine sp^2 carbons (CH). The ¹³C NMR chemical shift values are also helpful to establish the position of the substituents. The carbonyl carbon (C-9) of xanthones with a hydroxyl group in the *peri* position adjacent to it appears at δ80–185, whereas the carbonyl carbon without the hydroxyl group in this position normally resonates around δ174–175 (Kijjoa *et al.*, 2000a,b). However, it is not easy to predict the chemical shifts of the oxygen bearing quaternary sp^2 carbons of the xanthone nucleus (C-4a and C-10a), which are found to resonate at δ144–160, depending on the nature and position of substituents on each ring. On the contrary, the carbon chemical shift of the methoxyl substituent is very useful to determine its position. The sterically compressed

methoxyl groups, that is, when they are flanked by two oxygenated substituents or one oxygenated substituent and a carbonyl group, are found to have resonance at $\delta 61$–63; otherwise, they are found at $\delta 55$–57 (Kijjoa *et al*., 1998). Although the proton and carbon chemical shifts from the ^{1}H and ^{13}CNMR spectra were important in deducing the structures of myriad of xanthones, 2D NMR (COSY, HSQC, HMBC, and NOESY) have been used to elucidate unambiguously not only simple oxygenated and prenylated xanthones but also xanthones with complex structures. Through the diagonal cross peaks in the COSY spectrum, it is possible to observe the proton coupling systems which, in turn, allow us to identify the substitution pattern on the benzene rings of the xanthone scaffold. On the other hand, the HSQC cross peaks enable us to pinpoint the proton directly bonded to a particular carbon. Therefore, a combination of COSY and HSQC spectra could be used to determine the structures of fragments of molecule composed entirely of protonated carbons; however, the different fragments could not be assembled into a full structure when these fragments are separated by nonprotonated carbons or heteroatoms (Breton and Reynolds, 2013). Thus, the HMBC spectrum, whose cross peaks allow us to observe the connectivity of the proton and the carbons separated by two or three bonds, is fundamental to solve this problem. The cross peaks between the aromatic, hydroxyl, methoxyl protons and the aromatic and carbonyl carbons in the substituted xanthones can be used to establish the position of the hydroxyl and methoxyl groups on the ring. In addition, the HMBC cross peaks can be also useful to determine the position of the prenyl and modified prenyl substituents (Figueroa *et al*., 2009). NOESY experiment, which based on the nuclear Overhauser enhancement between the protons in close proximity, is sometimes very important to determine the positions of the substituents on the xanthone nucleus (Figueroa *et al*., 2009), in addition to the relative configuration of the stereogenic carbons of the substituents of the xanthone derivatives. Liu *et al*. used the NOESY cross peaks between the methyl protons on C-3 (CH$_3$-21) and C-18 methylene protons to determine the relative configuration of C-3 and C-19 of citreamicin θ B (**55**) (Liu *et al*., 2012a,b). In order to determine the absolute configuration of the stereogenic carbons of the substituents of xanthone derivatives by NMR, a modified Mosher method has been used (Ohtani *et al*., 1991). This method is

based on the differences between the proton chemical shifts in the (*R*)-MPA and in the (*S*)-MPA esters. Pontius *et al*. applied this method to determine the absolute configuration of C-3′ of the tetrahydropyran moiety of chaetoxanthone A (**35**) (Pontius *et al*., 2008a,b), whereas Figueroa *et al*. applied the advanced Mosher's methodology to determine the absolute configuration at C-20 and C-25 of 14-methoxytajixanthone (**26**) (Figueroa *et al*., 2009). Besides the NMR methods, the positive and negative Cotton effects observed in the circular dichroism (CD) spectrum is very useful to prove the configuration of the stereogenic carbons. Liu *et al*. used the CD curve to observe the opposite Cotton effects of citreamicin θ A (**54**) and citreamicin θ B (**55**), which led to the conclusion that they were diastereoisomers (Liu *et al*., 2012a,b). On the other hand, Pontius *et al*., by comparison of the CD curves of chaetoxanthone A (**35**) and chaetoxanthone B (**36**), have concluded that the latter was isolated as a racemate (Pontius *et al*., 2008a,b).

Electron impact (EI) and chemical ionization (CI) low resolution mass spectra (LRMS) had been widely used to determine the molecular weight of xanthone derivatives. The limitation of these techniques is that they do not provide an accurate mass of a molecule and, hence, its molecular formula. On the contrary, high resolution mass spectrum (HRMS) gives an accurate molecular mass and, thus, provides the molecular formula of the compounds. Consequently, HRMS has proved to be vital for structure elucidation of all classes of secondary metabolites, and therefore there is no exception for xanthone derivatives, especially those isolated from marine organisms as they sometimes contain chlorine, bromine, nitrogen, and sulfur atoms (Figueroa *et al*., 2009; Liu *et al*., 2012a,b). From the molecular formula obtained from HRMS, one can determine the degree of unsaturation of the compound, and because the xanthone nucleus has 10° of unsaturation, the total degree of unsaturation of the molecule can allow us to determine the nature of the substituents, for example, cyclic/noncyclic or saturated/unsaturated. There are a few HRMS techniques used to determine the molecular mass/molecular formula of small molecules such as xanthones, depending mainly on the type of the ion source. HR-EIMS, which uses a hard source (incident energy 70 eV) for ionization, gives the *m/z* value of a molecular ion [M$^+$]. Although electron impact causes extensive fragmentation of a molecule, it normally gives a strong

molecular ion peak [M⁺] for simple xanthones (Kijjoa *et al.*, 2000a,b). On the other hand, fast atom bombardment (FAB), which uses high energy argon or xenon atoms to bombard molecules in the ion source, has been used to determine the molecular mass of unstable large and small molecules. ⁺FAB-HRMS, which gives a strong $[M + H]^+$ ion, has been used to determine the molecular formula of many xanthone derivatives. Recently, electrospray ionization (ESI) technique, which uses soft ionization method to obtain stable molecular ion, was developed. HR-ESIMS is now widely used to determine an accurate mass of the $[M + H^+]$ ion and molecular formula of large and small thermally fragile molecules. Liu *et al.* used HR-ESIMS to determine the molecular formula of polycyclic xanthones isolated from the marine-derived actinomycetes *S. caelestis* (Liu *et al.*, 2012a,b).

Even though infrared (IR) and UV spectra are not currently the first choice for structure elucidation of secondary metabolites, including xanthones, they can also give some valuable structural information. Xanthone derivatives exhibit characteristic absorptions of the conjugated carbonyl (C=O stretching) at 1700–1720 cm⁻¹ or the hydrogen-bonded carbonyl (with the hydroxyl group on C-1 or C-8) at 1650–1660 cm⁻¹ and aromatic at 3000–3100 cm⁻¹ (C–H stretching) and 1450–1600 cm⁻¹ (C=C stretching) in the IR spectra. Xanthones with a hydroxyl substituent also exhibit absorption band characteristic of the hydroxyl group at 3300–3500 cm⁻¹ (O–H stretching). Xanthones also exhibit characteristic absorptions of extended conjugated aromatic system in the UV spectrum at 200–400 nm, and the patterns of absorption are dependent on the number and positions of the hydroxyl and methoxyl groups, but in general, there is a strong absorption above 340 nm (Dean, 1963). Kijjoa *et al.* used the effects of shift reagents on the UV absorption maxima to determine the position of the hydroxyl substituents (Kijjoa *et al.*, 2000a,b). Similar to flavonoids, xanthones containing hydroxyl group adjacent to the carbonyl group, that is, on C-1 or C-8, form acid stable complexes with AlCl₃. Furthermore, Kijjoa *et al.*, after being unable to distinguish between the structures of 3,8-dihydroxy-1,2,4-trimethoxyxanthone and 2,8-dihydroxy-1,3,4-trimethoxyxanthone by analysis of HMBC correlations, used the UV spectrum with shift reagent to solve the structural problem. Because the UV spectrum of the compound exhibited a pronounced bathochromic shift on addition of sodium acetate solution, which is characteristic of a phenolic hydroxyl group *para* to the carbonyl carbon, they have concluded that the structure of the compound was 3,8-dihydroxy-1,2,4-trimethoxyxanthone (Kijjoa *et al.*, 2000a).

Another important technique for structure elucidation is X-ray crystallography. This technique is useful not only to elucidate conclusively the structure of the compounds but also to determine the absolute configuration of the stereogenic carbons. Because the xanthone scaffold does not have stereogenic carbon, the single crystal X-ray diffraction is not widely used to elucidate the structure of xanthone derivatives. However, this technique can be valuable for xanthone derivatives with complex structure, especially those having stereogenic carbons. Kijjoa *et al.* obtained the X-ray crystal structure of 7-hydroxy-1,2,3,8-tetramethoxyxanthone, which confirmed the structure elucidated by NMR spectral analysis (Kijjoa *et al.*, 1998). Gales and Damas have reviewed the X-ray diffraction data of 47 xanthone derivatives (Gales and Damas, 2005). Krick *et al.* obtained the X-ray crystal structure of the xanthone derivative monodictysin A (**21**), which allowed them to confirm its relative configuration found on the basis of NMR data but could not determine its absolute configuration reliably (Krick *et al.*, 2007). Shao *et al.* was able to obtain the X-ray crystal structure of methyl 8-hydroxy-6-methyl-9-oxo-9*H*-xanthene-1-carboxylate (**11**), which was used to confirm its structure elucidated by NMR methods (Shao *et al.*, 2008). Pan *et al.* also used the X-ray crystal structure to confirm the structure of 1-hydroxy-8-(hydroxymethyl)-3-methoxy-6-methyl-9*H*-xanthen-9-one (**33**), isolated from a mangrove endophytic fungus *Phoma* sp. (Pan *et al.*, 2010).

5   CONCLUSION

Xanthones are a class of secondary metabolites, with interesting biological and pharmacological activities, normally found in higher plants and fungi; however, they are not frequently mentioned as marine natural products. This was due to the fact that a vast majority of marine natural products have been traditionally derived from marine macroorganisms such as invertebrates and algae. Only recently, attention has been focused on the marine microorganisms as sources

of secondary metabolites with unique structure and interesting biological and pharmacological activities. For this reason, marine-derived fungi have emerged as a potential source of marine natural products as they are capable of producing a variety of interesting secondary metabolites. Because marine fungi are normally associated with other marine macroorganisms, it is also believed that they are true producers of the compounds isolated thereof. Similar to their terrestrial counterparts, marine-derived fungi also produce xanthones. Although many common xanthones have been isolated from both marine-derived and terrestrial fungi, some xanthones with complex structure and rare substituents have been found only in the marine-derived fungi and actinomycetes.

Therefore, the chemical diversity allied with a myriad of biological and pharmacological activities of marine-derived xanthones can make the research in this field more challenging. Furthermore, this class of compounds could provide the core scaffolds for the development of new drugs. While the marine world offers an extremely rich resource for novel compounds, it also represents a large field of research that requires inputs from various scientific areas to bring the marine chemical diversity up to its therapeutic potential.

ACKNOWLEDGMENTS

REFERENCES

Abdel-Lateff, A., Klemke, C., König, G. M., et al. (2003) *J. Nat. Prod.*, **66**, 706–708.

Baker, D. D., Chu, M., Oza, U., et al. (2007) *Nat. Prod. Rep.*, **24**, 1225–1244.

Bhatnagar, I. and Kim, S.-K. (2012) *Environ. Toxicol. Pharmacol.*, **34**, 631–643.

Blunt, J. W., Copp, B. R., Hu, W.-P., et al. (2009) *Nat. Prod. Rep.*, **26**, 170–244.

Breton, R. C. and Reynolds, W. F. (2013) *Nat. Prod. Rep.*, **30**(4), 501–524.

Bringmann, G., Lang, G., Steffens, S., et al. (2003) *Phytochemistry*, **63**, 437–443.

Bugni, T. S. and Ireland, C. M. (2004) *Nat. Prod. Rep.*, **21**, 143–163.

Dean, F. M. (1963) *Naturally Occurring Oxygen Ring Compounds*, Butterworths, London, p. 266.

El-Seedi, H. R., El-Ghorab, D. M. H., El-Barbary, M. A., et al. (2009) *Curr. Med. Chem.*, **16**(20), 2581–2626.

Figueroa, M., González, M. C., Rodríguez-Sotres, R., et al. (2009) *Bioorg. Med. Chem.*, **17**, 2167–2174.

Gales, L., and Damas, M. (2005) *Curr. Med. Chem.*, **12**, 2499–2515.

Glaser, K. B., and Mayer, A. M. S. (2009) *Biochem. Pharmacol.*, **78**, 440–448.

Hong, R. (2011) *Pharm. Biol.*, **49**(8), 796–799.

Huang, Z., Yang, R., Guo, Z., et al. (2010a) *Chem. Nat. Compd.*, **46**(3), 348–351.

Huang, Z., Yang, R., Yin, X., et al. (2010b) *Magn. Reson. Chem.*, **48**, 80–82.

Huang, Z., Yang, J., Cai, X., et al. (2012) *Nat. Prod. Res.*, **26**(14), 1291–1295.

Jimeno, J., Faircloth, G., Sousa-Faro, J. M. F., et al. (2004) *Mar. Drugs*, **2**, 14–29.

Khamthong, N., Rukachaisirikul, V., Phongpaichit, S., et al. (2012a) *Tetrahedron*, **68**, 8245–8250.

Khamthong, N., Rukachaisirikul, V., Tadpetch, K., et al. (2012b) *Arch. Pharm. Res.*, **35**(3), 461–468.

Kijjoa, A., and Sawangwong, P. (2004) *Mar. Drugs*, **2**, 73–82.

Kijjoa, A., Gonzalez, M. J., Pinto, M. M. M., et al. (1998) *Phytochemistry*, **49**, 2159–2162.

Kijjoa, A., Gonzalez, M. J., Afonso, C. M., et al. (2000a) *Phytochemistry*, **53**, 1021–1024.

Kijjoa, A., Gonzalez, M. J., Pinto, M. M. M., et al. (2000b) *Phytochemistry*, **55**, 833–836.

König, G. M., Kehraus, S., Seibert, S. F., et al. (2006) *Chem. Bio. Chem.*, **7**, 229–238.

Kossuga, M. H., Romminger, S., Xavier, C., et al. (2012) *Braz. J. Pharmacogn.*, **22**(2), 257–267.

Kosta, S., Jain, R., and Tiwari, A. (2008) *Pharmacologyonline*, **1**, 1–3.

Kralj, A., Kehraus, S., Krick, A., et al. (2006) *J. Nat. Prod.*, **69**, 995–1000.

Krick, A., Kehraus, S., Gerhäuser, C., et al. (2007) *J. Nat. Prod.*, **70**, 353–360.

Lee, Y. M., Li, H., Hong, J., et al. (2010) *Arch. Pharm. Res.*, **33**(2), 231–235.

Li, C., Zhang, J., Shao, C., et al. (2011) *Chem. Nat. Compd.*, **47**(3), 382–384.

Liu, L.-L., Xu, Y., Han, Z., et al. (2012a) *Mar. Drugs*, **10**, 2571–2583.

Liu, T., Zhang, L., Li, Z., et al. (2012b) *Chem. Nat. Compd.*, **48**(5), 771–773.

Malet-Cascón, L., Romero, F., Espliego-Vázquez, F., *et al*. (2003) *J. Antibiot.*, **56**(3), 219–225.

Malmstrøm, J., Christophersen, C., Barrero, A. F., *et al*. (2002) *J. Nat. Prod.*, **65**, 364–367.

Masters, K.-S., and Bräse, S. (2012) *Chem. Rev.*, **112**, 3717–3776.

Ohtani, I., Kusumi, T., Kashman, Y., *et al*. (1991) *J. Am. Chem. Soc.*, **113**, 4092–4096.

Pan, J.-H., Deng, J.-J., Chen, Y.-G., *et al*. (2010) *Helv. Chim. Acta*, **93**, 1369–1374.

Pinto, M. M. M., and Castanheiro, R. A. P. (2009) Natural prenylated xanthones: chemistry and biological activities, in *Natural Products: Chemistry, Biochemistry and Pharmacology*, ed. G. Brahmachari, Narosa Publishing House PVT. LTD, Nova Deli, India, Chap. 17, pp. 520–676.

Pinto, M. M. M., Sousa, M. E., and Nascimento, M. S. J. (2005) *Curr. Med. Chem.*, **12**, 2517–2538.

Pontius, A., Krick, A., Kehraus, S., *et al*. (2008a) *J. Nat. Prod.*, **71**, 1579–1584.

Pontius, A., Krick, A., Mesry, R., *et al*. (2008b) *J. Nat. Prod.*, **71**(11), 1793–1799.

Pouli, N., and Marakos, P. (2009) *Anti Cancer Agents Med. Chem.*, **9**(1), 71–98.

Rateb, M. E., and Ebel, R. (2011) *Nat. Prod. Rep.*, **28**, 290–344.

Rodríguez, J. C., Puentes, J. L. F., Baz, J. P., *et al*. (2003) *J. Antibiot.*, **56**(3), 318–321.

Saleem, M., Ali, M. S., Hussain, S., *et al*. (2007) *Nat. Prod. Rep.*, **24**, 1142–1152.

Shao, C., She, Z., Guo, Z., *et al*. (2007) *Magn. Reson. Chem.*, **45**, 434–438.

Shao, C., Wang, C., Wei, M., *et al*. (2008) *Magn. Reson. Chem.*, **46**, 1066–1069.

Silva, A. M. S., and Pinto, D. C. G. A. (2005) *Curr. Med. Chem.*, **12**, 2481–2497.

Simpson, T. J. (2012) *Chem. Bio. Chem.*, **13**, 1680–1688.

Sun, R.-R., Miao, F.-P., Zhang, J., *et al*. (2013) *Magn. Reson. Chem.*, **51**, 65–68.

Thomas, T. R. A., Kavlekar, D. P., and LokaBharathi, P. A. (2010) *Mar. Drugs*, **8**, 1417–1468.

Trisuwan, K., Rukachaisirikul, V., Kaewpet, M., *et al*. (2011) *J. Nat. Prod.*, **74**, 1663–1667.

Ueda, J.-Y., Takagi, M., and Shin-ya, K. (2010) *J. Antibiot.*, **63**, 615–618.

Vieira, L. M. M., and Kijjoa, A. (2005) *Curr. Med. Chem.*, **12**(21), 2413–2446.

Wen, L., Lin, Y.-C., She, Z.-G., *et al*. (2008) *J. Asian Nat. Prod. Res.*, **10**(2), 133–137.

Wu, Z.-J., Ouyang, M.-A., and Tan, Q.-W. (2009) *Pest Manag. Sci.*, **65**, 60–65.

Xu, J. (2011) *Curr. Med. Chem.*, **18**, 5224–5266.

Zhu, F., and Lin, Y. (2007) *Chem. Nat. Compd.*, **43**(2), 132–135.

Żwir-Ferenc, A., and Biziuk, M. (2006) *Polish J. Environ. Stud.*, **15**(5), 677–690.

Sesquiterpenes and Other Terpenoids

Eirini Kouloura, Job Tchoumtchoua, Maria Halabalaki and Alexios-Leandros Skaltsounis

Division of Pharmacognosy and Natural Products Chemistry, Department of Pharmacy, University of Athens, Athens, Greece

1   INTRODUCTION

The terpenoids constitute the most numerous and structurally varied family of plant natural products (>40,000 metabolites) (Roberts, 2007). The terms *terpenoid* and *isoprenoid* broadly refer to any compound derived from repetitive fusion of branched five carbon units based on isopentane skeleton and are frequently used interchangeably. The word terpene originates from turpentine (resin of pine trees), which derives (via French and Latin) from the Greek word terebinthine ($\tau\epsilon\rho\epsilon\beta\iota\nu\theta\acute{\iota}\nu\eta$). Terebinthine, the resin from terebinth tree, was originally distilled (Barnhart, 1995) and used for the isolation of the first terpenoids in the 1850s (Croteau, Kutchan, and Lewis, 2000). The term *terpene* normally refers only to a hydrocarbon molecule, whereas *terpenoid* refers to a terpene that has been modified, such as oxygenated, hydrogenated, and dehydrogenated derivatives (Zwenger and Basu, 2008).

The majority of terpenoids are derived from the plant kingdom; however, some representatives have also been reported as metabolites of insects and marine organisms. All living organisms naturally produce terpenoids as part of primary metabolism; however, most of them are produced via secondary metabolism (Withers and Keasling, 2007). These terpenoids are synthetized to facilitate certain essential physiological functions; however, the biological properties and roles of the majority are still not widely explored. Terpenoids and their derivatives play key roles in all aspects of life as constituents of membranes, regulators of gene expression, defense compounds, vitamins, mating pheromones, growth regulators (phytohormones), components of signal transduction pathways, and constituents of photosynthetic pigments (Gershenzon and Dudareva, 2007).

Taking advantage of these characteristics, humans have used terpenoid components for several pharmaceutical applications. Among them, the lactonic sesquiterpene artemisinin from *Artemisia annua* L. (Asteraceae) has become one of the most widely used antimalarial drugs in the world (Martin *et al.*, 2003). Similarly, the anticancer diterpene paclitaxel (Taxol®), originally isolated from the bark of Pacific yew, *Taxus brevifolia* Nutt. (Taxaceae), has been used for many years as a chemotherapeutic agent (Goodman and Walsh, 2001). Moreover, terpenoids have become commercially useful products as flavoring essences, solvents, and polymers. Besides, agriculture has also shown an increasing interest in terpenoids using them as agrochemicals (pyrethrins and azadirachtin) (Elakovich Stella, 1988).

The very first terpenoids isolated from turpentine were named monoterpenes even though they were later found to possess two isoprene units (Croteau, Kutchan, and Lewis, 2000). Another largely studied subgroup is sesquiterpenes (i.e., one and a half terpene), which possess three isoprene units and 15 carbon atoms. Sesquiterpenoids are characterized by

a great chemical and biological interest and are widely distributed throughout the plant kingdom. They have been isolated broadly as constituents of essential oils located in the leaves, fruits, and seeds and from resins or oils located in special channels in the stem or root barks (Bakkali *et al*., 2008). Similar to terpenoids, sesquiterpenoids also exhibit important physicochemical role in plants and biological properties, which have been utilized by human societies. Moreover, sesquiterpene lactones constitute an important group of these compounds, providing a big number of metabolites with a wide structural diversity (around 3000 reported different structures). Besides, sesquiterpene lactones are of significant chemotaxonomic importance and possess considerable biological activities (Mahesh *et al*., 2010; Rodriguez, Towers, and Mitchell, 1976). In general, sesquiterpenes possess a vast range of pharmacological activities, such as antioxidant, cytotoxic, anti-inflammatory, and antimicrobial activities. For instance, T-cadinol is a calcium antagonist, possibly interacting with the dihydropyridine binding sites on the calcium channels (Claeson, Zygmunt, and Högestätt, 1991); α-humulene, given either orally or by aerosol, exhibited marked anti-inflammatory properties in a murine model of airways allergic inflammation, an effect that seemed to be mediated via reduction of inflammatory mediators, adhesion molecule expression, and transcription factors activation (Rogerio *et al*., 2009); a mixture of furanodiene-6-one and methoxyfuranoguaia-9-ene-8-one showed antibacterial and antifungal activity against standard pathogenic strains of *Escherichia coli*, *Staphylococcus aureus*, *Pseudomonas aeruginosa* and *Candida albicans*; these compounds also had local anaesthetic activity, blocking the inward sodium current of excitable mammalian membranes (Dolara *et al*., 2000).

The importance of sesquiterpenoids and other terpenoids in medicine, industry, and agriculture has led to numerous studies related to their detection, isolation, and identification. At the same time, the "small terpenoids phytochemistry" as an emerging area in natural products science was introduced. Interestingly, particular physicochemical features characterizing this group complicate significantly their study and the discovery of new ones. Owing to their apolar nature, particular consideration has to be undertaken for their extraction from the starting material, their analytical manipulation, and their targeted isolation.

Moreover, stereochemical and configuration aspects obscure significantly their analysis and structural elucidation, necessitating the incorporation of specific techniques and combined approaches. In addition, many representatives of sesquiterpenoids and other terpenoids are relatively labile, and isomerization or rearrangement phenomena within the molecules are often observed, resulting in an additional factor of difficulty during their study.

Nevertheless, recent advances in phytochemical analysis and chemistry and that in the available instrumentation resulted in new analytical techniques and approaches, facilitating considerably the study of sesquiterpenoids and other terpenoids.

2 CLASSIFICATION, DISTRIBUTION, AND BIOSYNTHESIS OF TERPENOIDS IN PLANT KINGDOM

2.1 Classification and Distribution

Terpenoids are broadly classified according to the number of carbons and subsequently isoprene units they contain. In general, isoprene units are linked to each other by head-to-tail coupling, forming regular terpenoids; however, in some cases, other coupling mechanisms take place, resulting in irregular terpenoid formation (Ruzicka, 1953) (Figure 1). Terpenoids range in size from the five-carbon hemiterpenes, built up from a single isoprene unit, to giant molecules comprising thousands of isoprene units such as natural rubber (Bhat, Nagasampagi, and Sivakumar, 2005a). Specifically, terpenoids that are formed from two, three, and four isoprene units are named as monoterpenes, sesquiterpenes and diterpenes, respectively (Table 1).

Along with this classification, terpenoids are characterized from presence or not and the numbers of rings provided on the main structure. Specifically, sesquiterpenoids, according to the number of rings, may be classified to acyclic (e.g. farnesene), monocyclic (e.g. germacrene), bicyclic (e.g. caryophyllene), tricyclic (e.g. cedrol), tetracyclic (e.g. gossypol), and pentacyclic (e.g. siccanochromene) derivatives (Cordell, 1976). Following specific structural characteristics, sesquiterpenoids could also be classified to the widely acknowledged groups of sesquiterpene hydrocarbons (e.g. farnesene and humulene),
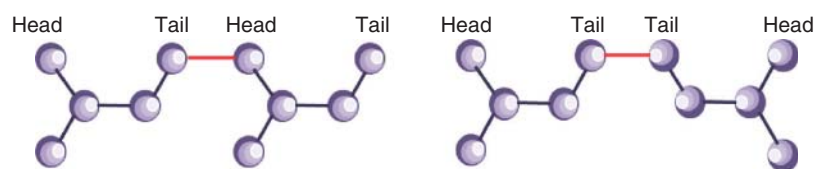
Figure 1 Regular (head-to-tail) and irregular (tail-to-tail) arrangement of isoprene units.

Table 1 Classification of sesquiterpenoids according to their carbon atom number and isoprene units.

Class of terpenoids	No. carbon atoms	No. isoprene units
Hemiterpenoids	C5	1
Monoterpenoids	C10	2
Sesquiterpenoids	C15	3
Diterpenoids	C20	4
Sesterterpenoids	C25	5
Triterpenoids	C30	6
Natural rubbers	>C40	>8

oxygenated hydroxyl and carbonyl derivatives, aldehydes (e.g. farnesal and lepidozenal), alcohols (e.g. δ-elemanol and β-germacrenol), esters (e.g. torilin and ejaponines), and sesquiterpene lactones (e.g. artemisitene and artegallin) (Table 2).

As far as their distribution is concerned, sesquiterpenoids are widely spread in the plant kingdom. Among the most widely reported plant families for the occurrence of these compounds is Asteraceae, characterized by a significant high abundance of sesquiterpenoids and sesquiterpene lactones (Seaman, 1982; Zdero and Bohlmann, 1990). However, sesquiterpenoids are also present in Apiaceae (Chizzola, 2010; Drew *et al.*, 2009), Acanthaceae (Facundo, Pinto, and Rezende, 2005; Moronkola, Atewolara-Odule, and Olubomehin, 2009), Anacardiaceae (Ono *et al.*, 2008), Aristolochiaceae (Silva-Brandao, Solferini, and Trigo, 2006), Cannabaceae (Wang *et al.*, 2008), Cupressaceae (Grantham and Douglas, 1980; Yatagai, Sato, and Takahashi, 1985), Euphorbiaceae (De *et al.*, 2009; Juergens *et al.*, 2006), Geraniaceae (Demarne, Viljoen, and Van, 1993; Lalli *et al.*, 2006), Hepatidae (Bedoya, Abad, and Bermejo, 2008), Lamiaceae (Ali *et al.*, 2007; Zhou *et al.*, 2011), Lauraceae (Demarne, Viljoen, and Van, 1993; Lalli *et al.*, 2006), Magnoliaceae (Zhong *et al.*, 2006), Myrtaceae (Nakamura *et al.*; Pereira *et al.*), Pinaceae (Khan, Dubovenko, and Pentegova, 1983; Kula, Masarweh, and Gora, 1996), Rutaceae (Guy *et al.*, 2001; Pala-Paul *et al.*,

2009), Valerianaceae (Itokawa *et al.*, 1993), Zingiberaceae (Dong *et al.*, 2013; Wabo, Tane, and Connolly, 2006), and other families (Bedoya, Abad, and Bermejo, 2008).
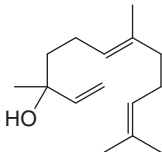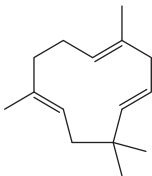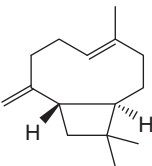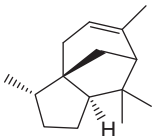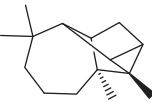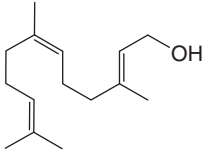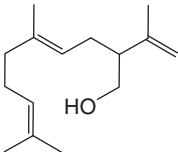
2.2 Biosynthesis

All terpenoids are synthesized through the condensation of isoprene units (2-methyl-l,9-butadiene). In 1914, Wallach formulated the *isoprene rule*, which comprises the basic principle for the biosynthesis of most terpenoids. According to this rule, most terpenoids could be synthesized by the repetitive fusion of isoprene units. This concept was further developed and enriched in 1953 by Ruzicka when the *biogenetic isoprene rule* was published. This suggestion for the biosynthesis of terpenoids is mainly based on biochemical aspects involving electrophilic elongations, cyclizations, and rearrangements without emphasizing the precise chemical nature of biological precursors (Ruzicka, 1953).
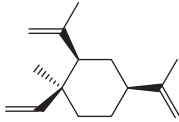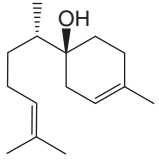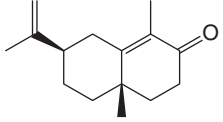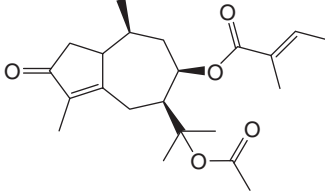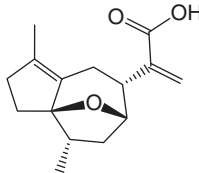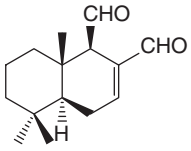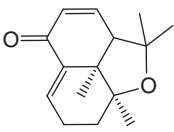
Metabolically speaking, the biosynthesis of terpenoids can be divided into three stages: formation of the isoprene unit, condensation of the isoprene units and chain elongation, and cyclization and modification of linear terpenoid precursors.

An essential biosynthetic step for the formation of terpenoids is the biosynthesis of isopentenyl diphosphate (IPP) and dimethylallyl diphosphate

Table 2   Classification of sesquiterpenoids according to their structure characteristics.
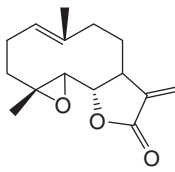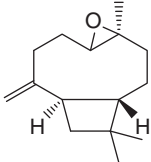
Type of sesquiterpenoids	Name	Example	
According to the number of the rings	Acyclic		Nerodiol
	Monocyclic		α-Humulene
	Bicyclic		β-Caryophyllene
	Tricyclic		α-Cedrene
	Tetracyclic		Longicyclene
According to the fusion of the isopren units	Regular		Farnesol
	Irregular		Sesquilavandulol

Table 2 *(Continued)*

Type of sesquiterpenoids	Name	Example	
According to the functional groups/structure	Hydrocarbons		β-Elemene
	Alcohol		β-Bisabolol
	Ketone		α-Cypenone
	Esters		Torilin
	Acids		Lacitemzine
	Aldehydes		Polygodial
	Ethers		Nardonoxide

(Continued overleaf)

Table 2   *(Continued)*

Type of sesquiterpenoids	Name	Example
	Oxides	Caryophyllene oxide
	Lactones	Parthenolide

(DMAPP) derivatives. These precursor isoprene units may be derived from two pathways, the *mevalonate* (MVA) pathway, which was until lately known as the only biosynthetic pathway, and a recently discovered nonstandard MVA pathway involving the phosphorylation of isopentenyl phosphate, known as *1-deoxy-*D-*xylulose-5-phosphate* (*DXP*) pathway (Figure 2). The MVA pathway is present in all eukaryotes and some Gram-positive prokaryotes, whereas the DXP pathway is generally found in prokaryotes. Particularly, in plants, IPP is synthesized in the plastids through DXP pathway, whereas, in cytosol, through MVA pathway. The DXP pathway produces both the isomers IPP and DMAPP, whereas in the MVA pathway, only IPP is generated and further isomerized to DMAPP by the action of isopentenyl diphosphate isomerase (*idi*) (Bouvier, Rahier, and Camara, 2005; Kuzuyama and Seto, 2003; Withers and Keasling, 2007).

The generation of the higher order terpenoid building blocks [geranyl pyrophosphate (GPP), farnesyl pyrophosphate (FPP), and geranylgeranyl pyrophosphate (GGPP)] is the next step of terpenoids biosynthesis. These precursors result from the action of the prenyltransferases GPP synthase, FPP synthase, and GGPP synthase, respectively. All three enzymes can use IPP and DMAPP for the elongation steps. In addition, FPP and GGPP synthases can extend GPP and FPP to form the respective C15 and C20 products (Mahmoud and Croteau, 2002). In order to build up the *regular* terpenes, isoprenyl units are connected by the sequential head-to-tail addition, whereas the *irregular* terpenes are produced by non–head-to-tail joining of the two building units or modifications with carbon–carbon bond rearrangements, resulting in a complex tracing of the original pattern of isoprene units.

Finally, the formation of the different monoterpenes, sesquiterpenes, and diterpenes from their acyclic precursors is performed through the action of the terpenoid synthases, also called as *cyclases*, because their chemical involvement results in cyclic derivatives. Moreover, the cyclization and further reactions such as oxidations, reductions, and isomerizations are responsible for the final conversion of the acyclic branch point intermediates to the diverse key skeletons of the monoterpene (GPP), sesquiterpene (FPP), and diterpene (GGPP) derivatives (Figure 3).

Thus, sesquiterpenoids are formed from complex rearrangement reactions of carbocations derived from FPP under the control of a class of sesquiterpene synthases or cyclases. As a result, a diverse array of structurally and stereochemically complex organic molecules are derived from the simple acyclic and achiral FPP (Cane, 1990). In Figure 4, some representative key skeletons are presented, illustrating the wide chemical variety of this group.

A special category of sesquiterpenes are sesquiterpene lactones that are formed from head-to-tail condensation of three isoprene units and subsequent
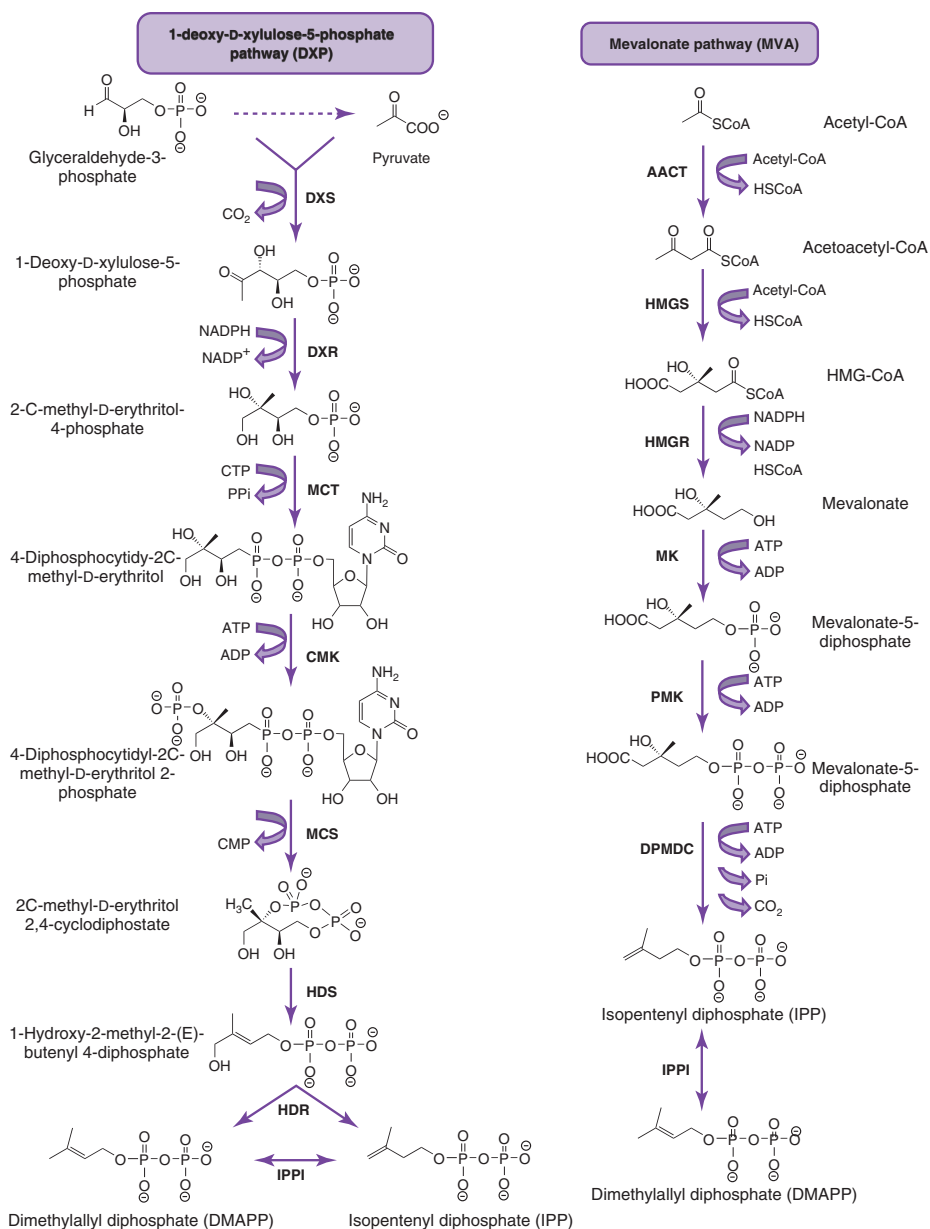
Figure 2 Cytosolic MVA and plastidial DXP pathways of IPP and DMAPP synthesis in plants. DXP pathway – enzyme abbreviations: DXS, 1-deoxy-D-xylulose-5-phosphate synthase; DXR, 1-deoxy-D-xylulose-5-phosphate reductoisomerase; MCT, 2-C-methyl-D-erythritol-4-phosphate cytidyl transferase; CMK, 4-(cytidine 5′-diphospho)-2-C-methyl-D-erythritol kinase; MCS, 2-C-methyl-D-erythritol-2,4-cyclodiphosphate synthase; HDS, 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate synthase; HDR, 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate reductase; IPPI, IPP-DMAPP isomerase. Mevalonate (MVA) pathway – enzyme abbreviations: AACT, acetoacetyl-CoA thiolase; HMGS, 3-hydroxy-3-methylglutaryl-CoA synthase; HMGR, 3-hydroxy-3-methylglutaryl-CoA reductase; MK, mevalonate kinase; PMK, 5-phosphomevalonate kinase; DPMDC, 5-diphosphomevalonate decarboxylase; IPPI, isopentenyl diphosphate-dimethylallyl diphosphate isomerase. Substrate abbreviations: DMAPP, dimethylallyl diphosphate; HMGCoA, 3-hydroxy-3-methylglutaryl-CoA; IPP, isopentenyl diphosphate.
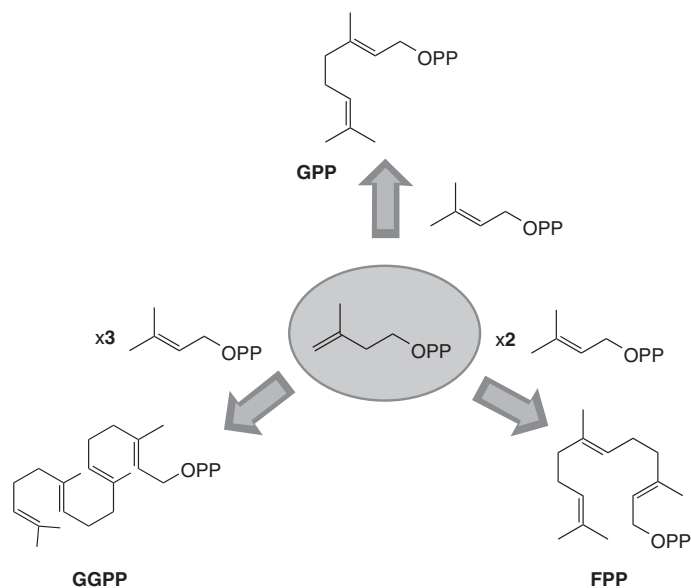
Figure 3 Biosynthetic precursors of monoterpenoids, sesquiterpenoids, and diterpenoids.
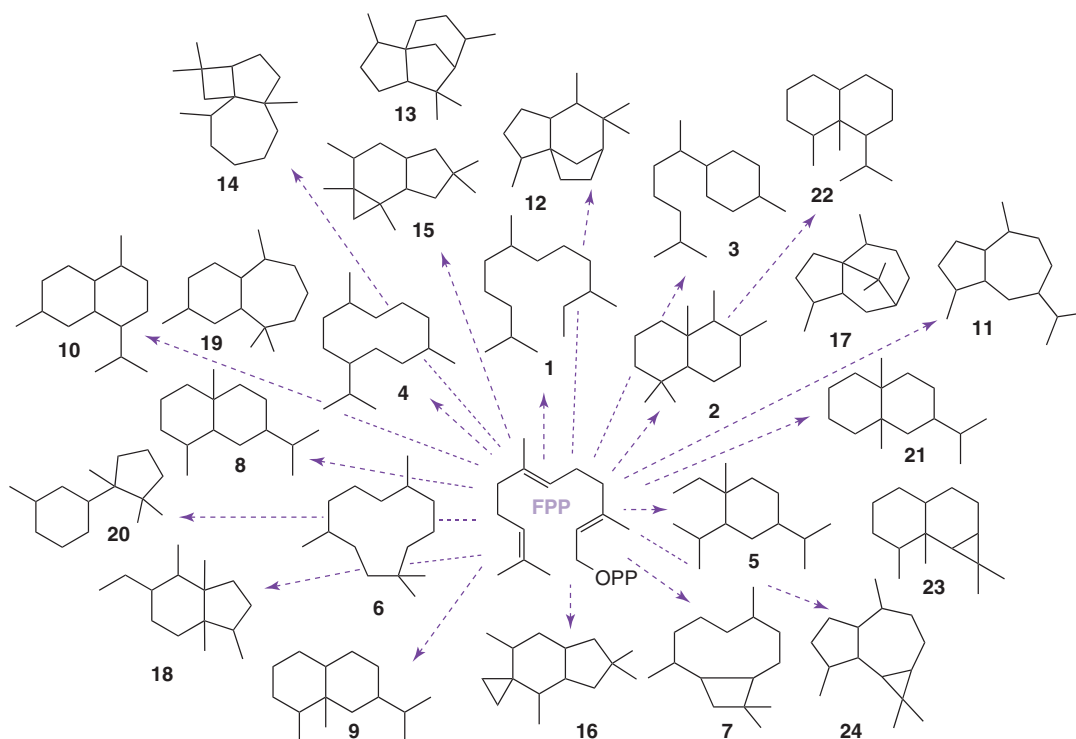
Figure 4 Various skeletal types of sesquiterpenoids: 1, farnesane; 2, drimane; 3, bisabolane; 4, germacrane; 5, elemane; 6, humulane; 7, caryophyllane; 8, eudesmane; 9, eremophilane; 10, cadinane; 11, guaiane; 12, zizaane; 13, cedrane; 14, panasinsane; 15, marasmane; 16, illudane; 17, patchoulane; 18, pinguisane; 19, himachalane; 20, herbertane; 21, valerane; 22, nardosinane; 23, aristolane; and 24, gurjunane.

cyclization and oxidative transformation to produce *cis*- or *trans*-fused lactones (Chaturvedi, 2011). The formation of the lactone ring constitutes an important common feature of the sesquiterpene lactones containing, in many cases, an α-methylene group. Among other modifications, the incorporation of hydroxyls or esterified hydroxyls and epoxide rings are common (Picman, 1986).

3 EXTRACTION

Extraction is a crucial first step in the analysis of medicinal plants, as it is necessary to obtain the desired chemical components from the herbal materials for further separation and characterization. Despite the number of pharmacopoeial references using conventional techniques, the utilization of modern, reproducible extraction techniques provides significant advantages such as reduction in organic solvent consumption and sample degradation, improvement in extraction efficiency, and ease of automation. The extraction of medicinal plants is likely to play an important role in the overall effort of efficacy of natural products analysis. Therefore, the development of simple and effective extraction methods is of great interest in the recent years. As a result, apart from the traditional extraction techniques such as distillation or maceration, many new ones have been developed, including ultrasound-assisted extraction (UAE), pressurized liquid extraction (PLE), supercritical fluid extraction (SFE), and microwave-assisted extraction (MAE) (Camel, 2001; Kaufmann and Christen, 2002; Luque de Castro, Jiménez-Carmona, and Fernández-Pérez, 1999; Ormeno, Goldstein, and Niinemets, 2011; Wang and Weller, 2006) (see **Extraction Methodologies: General Introduction**).

3.1 Distillation

3.1.1 *Hydrodistillation and Steam distillation*

Hydrodistillation (HD) is one of the most traditional methods used for the extraction of terpenoids. This technique is mainly applied for the production of essential oils from different plant materials. HD comprises a simple apparatus consisting of a still containing the plant material distributed in the water and a heater. This direct contact of plant material and water is the main characteristic of this method. Heating is performed in the presence of water, water-vapor pressure increases, and so does the vapor pressure of the essential oil. The practical advantages of this technique are the low cost, the ease to construct and handle, and the fact that enables field operation. However, this technique has been controversial for subsequent determination of the oil chemical composition because of the possible transformation of certain components such as esters and aldehydes by heat, steam, and pH. Moreover, HD is a time-consuming technique requiring several hours (Handa, 2008). In order to reduce the length of the process, to minimize the alteration of the natural constituents by possible oxidation, to decrease the loss of the most polar compounds and save energy, steam distillation (SD) could be applied as an alternative, which involves forcing the steam through the plant material. SD seems to be more suitable than HD for the extraction of rosemary *Rosmarinus officinalis* L. (Lamiaceae) essential oils, because it gave better yield and better oil composition (Boutekedjiret *et al*., 2003).

SD is widely used in rural areas applying similar equipment to HD. The main difference is that the plant material is in indirect contact with the boiling water as it is suspended in a perforated grid or plate above the water, reducing the capacity of the still but affording a better quality of oil. An optional approach is the direct SD. Here, the steam generated outside the still in a stand-alone steam generator generally referred to as a boiler. The real advantage of this satellite steam generation is that the amount of steam can be readily controlled; hence, plant material is heated no higher than 100 °C and, consequently, it should not undergo thermal degradation. Even though much higher capital expenditure is needed to build such a facility, SD is the most widely accepted process for the production of essential oils in large scale. For instance, SD has been used for the extraction of oils rich in sesquiterpene compounds from five *Desmos* species (Annonaceae) such as α-pinene, β-elemene, β-caryophyllene, germacrene D, bicyclogermacrene, and α-humulene (Dai *et al*., 2012). *Sclerocarya birrea* (A. Rich.) Hochst (Anacardiaceae) was also extracted using SD, and a sesquiterpene-rich oil of around 96% was obtained (Kpoviessi *et al*., 2011). Finally, *Senecio rufinervis* D.C (Asteraceae) extracted by SD was found to contain mainly germacrene D, followed by β-pinene,

β-caryophyllene, and β-longipinene, and to exhibit a significant analgesic activity (Mishra *et al.*, 2010).

3.1.2 Soxhlet Solvent Extraction

Soxhlet extraction, one of the oldest solvent extraction methodologies, is a leaching (liquid–solid) technique, typically pursued at or near atmospheric pressure and based on the choice of solvent, the use of elevated temperatures, and/or agitation. Normally, the sample is placed inside a thimble holder and attached onto a flask containing the extraction solvent. During operation, the solvent is heated to reflux and returns into the flask through the condenser. This procedure is repeated until the extraction is achieved (Luque de Castro and García-Ayuso, 1998). Soxhlet has been a standard technique during more than one century, and at present, it is the main reference to which the performance of other leaching methods is compared and used in a variety of pharmacopoeial monographs. Soxhlet extraction is a general and well-established technique, which surpasses in performance other conventional extraction techniques (Wang and Weller, 2006). A fundamental limitation of this method mostly toward modern extraction methods is the toxic solvent residue often present in the extract when toxic solvents (e.g. petroleum ether and benzene) are used for the extraction procedure. Besides, Soxhlet extraction is time consuming, necessitating large amounts of solvents and cannot be applied for the extraction of thermolabile compounds (Luque de Castro and García-Ayuso, 1998).

By optimizing the extraction conditions of alkamides from *Echinacea purpurea* (L.) Moench (Asteraceae) roots, a comparison study of conventional versus modern extraction methods was performed. Regarding sesquiterpenes composition of the obtained extracts, using Soxhlet extraction with *n*-hexane for 48 h as reference extraction method was observed to be the most effective method compared with UAE and MAE methods, despite the time consumption of Soxhlet extraction (Hudaib *et al.*, 2003). Zhou *et al.*, when comparing two conventional extraction methods, bottle-stirring and Soxhlet extraction, in order to obtain the maximum yield of parthenolide from feverfew (*Tanacetum parthenium* L., Asteraceae) concluded that both conventional methods had more or less the same yield using five different solvents. In case of water percentage addition to bottle-stirring extraction method, better results

were observed comparing with the water-free extraction procedure (Zhou, Kou, and Stevenson, 1999).

3.2 Maceration and Stirring Solvent Extraction

Maceration is also a widely applied extraction process, however, not that efficient for obtaining essential oils. Actually, it is a simple process where the ground plant material is immersed in a container with a single or a mixture of organic solvents. Specifically, for the extraction of small terpenoids, a variety of nonpolar solvents is applied such as diethylether, pentane, or hexane. The efficiency depends on the correct choice of solvents and temperature and the use of agitation choice in order to increase the solubility of compounds and to improve mass transfer (Singh, 2008).

In comparison to other techniques, the conventional maceration is time consuming as it needs generally more than a day to yield extracts. For sesquiterpenes, nonpolar solvent are often required. For instance, dichloromethane was used to obtained sesquiterpenes myrrhterpenoids from *Commiphora myrrha* (Nees) Engl. (Burseraceae) (Xu *et al.*, 2012). However, polar solvents such as methanol are sometimes utilized for total extraction, and then sesquiterpenes are obtained by suspension of the total extract in water and partition with ethyl acetate. This procedure was applied for example for the extraction of neuroprotective sesquiterpenes from *Petasites japonicus* (Sieb. et Zucc.) Maxim (Asteraceae) (Wang *et al.*, 2013).

Maceration has been used for the extraction of sesquiterpene lactones from various plants with a wide range of biological activities. Among others, some examples of sesquiterpene lactones with immunotoxic effects have been isolated from *Capersium rosulatum* Miquel (Astaraceae) (Moon and Zee, 2011), with anticancer properties from *Vernonia amygdalina* Del. (Asteraceae) (Luo *et al.*, 2011), with antinociceptive and antipyretic activities from *Centaurea solstitialis* L. ssp. *solstitialis* and *C. depressa* M. Bieb (Asteraceae) (Tabanca *et al.*, 2005).

3.3 Ultrasonic-Assisted Extraction (UAE) or Sonification

The UAE is a relatively novel technique that presents the beneficial effect of involving the use of ultrasound, which increases the permeability of cells and

produces microcavitation. After interaction with the plant material, ultrasound waves alter their physical and chemical properties, and their cavitational effect facilitates the penetration of the solvent, solubilization, and release of target compounds and enhances the mass transport by disrupting plant cells walls (Shirsath, Sonawane, and Gogate, 2012). It is a very promising technique as it improves the extraction efficiency and rate and allows solvents to be saved. Many thermally unstable compounds may degrade during other extraction processes such as thermal extraction and/or distillation (Luque de and Capote, 2007). UAE is performed at lower temperatures, diminishing the danger of thermal degradation as demonstrated during the extraction of essential oils from fresh garlic cloves (*Allium sativum* L., Amaryllidaceae) (Kimbaris *et al.*, 2006). In summary, the parameters to consider for the optimization of this technique are ultrasonic power, frequency, temperature, solvents, solvent–sample ratio, particle size, and structure of the target compounds (Esclapez *et al.*, 2011).

For instance, flaked caraway seeds (*Carum carvi* L., Apiaceae) were extracted at atmospheric pressure using Soxhlet, conventional, and UAE techniques. Paying attention to the two main compounds of the plant extract, carvone and limonene, the authors found that carvone yield and plant extract quality were better by UAE compared to those given by conventional methodology and the extraction rate of carvone and limonene were more rapid by UAE (Chemat *et al.*, 2004).

In order to potentially replace the conventional destructive extraction process of menthol from peppermint plants, *Mentha x piperata* L. (Lamiaceae), Shotipruk *et al*. studied the feasibility of using ultrasound to extract menthol from biologically viable peppermint plants. This study showed that plants ultrasonicated for 1 h at 22 °C in a standard 40 kHz ultrasonic bath released approximately 17.8 µg of menthol per gram of leaf tissue (2% of total product). The amount of menthol release increased with the time of treatment and was greatly affected by the temperature of the ultrasonic bath water. An increase from 2% to 12% of total product was observed when the temperature was increased from 22 to 39 °C. When the temperature effects were isolated, the mechanism of the product release was found to be that of cavitation. The treated plants remained viable

and were ready for the subsequent ultrasound extraction after approximately 4 days of recuperation. However, the amount of product released was reduced in subsequent extractions. This study has shown the possibility of using an online ultrasonic, nondestructive extraction method to continuously release intracellular plant metabolites from the plants while maintaining the plant's viability (Shotipruk, Kaufman, and Wang, 2001). HD and UAE have also been compared and evaluated regarding the capacity to isolate essential oil from fresh page mandarin flowers (*Citrus* hybrids, Ruraceae). The results of this study showed that essential oil obtained by UAE contained more oxygenated compounds and fewer monoterpenes components than those isolated by HD, enhancing then the idea of using this technique as a valuable tool for the recovery of aroma compounds from plants (Darjazi, 2011).

UAE method was associated to vacuum distillation for the extraction of flavor compounds from different varieties of *Mentha spicata* L. (Lamiaceae) using 70% ethanol during 5, 10, and 15 min and compared with essential oil obtained by HD. The gas chromatography-mass spectrometry (GC-MS) analyses of the samples showed that UAE associated with vacuum distillation provided extracts with higher flavoring strength because of the increased concentration of oxygenated compounds (mainly sesquiterpenes and monoterpenes). Therefore, UAE in combination with vacuum distillation in comparison with HD offered good advantages in terms of yield, selectivity, stability, and quality of flavor compounds extracts from *M. spicata* and is presented as a promising technique to recover aroma compounds from plants (Da Porto and Decorti, 2009).

In conclusion, it has become more and more common to use UAE for the extraction of natural products as it requires overall less energy and it is easy to handle. The conditions of the process have a strong influence on the extraction yield and may be first optimized prior to any process. The scaling up of this technique requires the establishment of a valid method for the characterization of the power applied and the design of a well-defined reactor. The research in scale-up, the design-improved sonoreactors, and the combination with other techniques seem to significantly enhance the development and commercialization of this technique (Esclapez *et al.*, 2011).

3.4 Pressurized Liquid Extraction (PLE)

Compared with traditional solvent extraction techniques, PLE presents many advantages in terms of automation, reproducibility, significant solvent reduction, and sample preparation time (Kaufmann and Christen, 2002). Besides, PLE is broadly recognized as a green extraction technique owing to the low organic solvent consumption (Mustafa and Turner, 2011). PLE is a solid–liquid extraction process performed by pumping the solvent into the extraction vessel containing the sample at elevated temperature and pressure. The pressurized solvent at high temperatures accelerates the extraction by increasing the solubility of the analytes in the solvent and by exalting the desorption rate of the analyte from the sample matrix (Mottaleb and Sarker, 2012). The degradation of thermolabile compounds constitutes a limitation of this technique (Wang and Weller, 2006). In order to achieve satisfactory extraction efficiency during PLE, it is important to optimize different parameters, including the extraction solvent used, the extraction temperature, the number of static cycles, and the static time (Ormeno, Goldstein, and Niinemets, 2011) (see **New Trends in Extraction of Natural Products: Microwave-Assisted Extraction and Pressurized Liquid Extraction**).

According to comparative studies concerning terpenoid extraction methods, PLE exhibited the highest extraction yield in the shortest extraction time and characterized by significant reproducibility. Specifically, PLE was proposed as the technique of choice for the extraction of monoterpenes and sesquiterpenes from *Juglans regia* L. (Juglandaceae). According to this study, the efficiency of PLE extraction proved to be superior to SD, Soxhlet extraction, UAE, and the solvent extraction in shaker (Fojtova, Lojkova, and Kuban, 2008). Five sample preparation methods (SD, extraction in the Soxhlet apparatus, SFE, solid-phase microextraction, and PLE) used for the isolation of aroma-active components from *Thymus vulgaris* L. (Lamiaceae) were compared; PLE and SFE (supercritical fluid extraction) were proved the most effective extraction methods for sesquiterpenoid components of the essential oil of *T. vulgaris*. According to this study, PLE resulted in the extraction of nonvolatile compounds together with the essential oil, which may be a reluctant parameter for the selection of this extraction method (Dawidowicz *et al*., 2008). Tam *et al*., after the comparison of three extraction methods (HD, PLE,

and SFE) and the optimization of the best among them, proposed a PLE extraction method with methanol as solvent at 140 °C and 1000 psi for the highest extraction efficiency of eight sesquiterpenes: α-copaene, cyperene, β-selinene, selina-4,11-diene, aristol-9-en-8-one, aristol-9-en-3-one, α-cyperone, and β-cyperone, from *Cyperus rotundus* L. (Cyperaceae) (Tam *et al*., 2007).

PLE technique has been applied to extract 11 sesquiterpenoids from *Curcuma* species (Zingiberaceae). Yang *et al*. developed a PLE and GC-MS method for identification and quantitative determination/estimation of germacrene D, curzerene, γ-elemene, furanodienone, curcumol, isocurcumenol, furanodiene, germacrone, curdione, curcumenol, and neocurdione. Different extraction parameters were studied, and the optimized conditions included methanol as solvent at 120 °C and 1500 psi, static extraction time of 5 min, and one static cycle (Yang *et al*., 2005).

Owing to the growing interest for the development of high throughput methods for standardized extracts, PLE can be applied at the scientific and industrial fields for applications involving the screening of large amount of different extracts (Benthin, Danz, and Hamburger, 1999).

3.4.1 Subcritical Water Extraction (SWE)

Subcritical water extraction or superheated water extraction (SWE) or pressurized hot water extraction (PHWE) is a specific application of pressurized solvent extraction using water as extraction solvent. PHWE, as solventless extraction technique, is an environmentally friendly, simple, and inexpensive technique (Mustafa and Turner, 2011). Subcritical water conditions are obtained at temperatures varying between the boiling point of water (100 °C) and the critical temperature of water (374 °C). The pressure is adjusted in order to maintain water in the liquid state. The main interest in the use of water as extracting agent is related to its physicochemical properties that change dramatically by increasing the temperature. Thus, under these conditions, water has possesses properties resembling to certain organic solvents and has been suggested to extract medium polar compounds. If thermally unstable compounds are to be extracted and high extraction temperatures cannot be applied, it is advisable to use a modifier, such as

ethanol, so that the extractions can be performed at lower temperatures and the degradation of thermally labile compounds could be avoided. The parameters that may affect the extraction efficiency in PHWE include temperature, extraction time, and addition of organic solvent or surfactants (Ong, Cheong, and Goh, 2006).

Applications of this technique involve the extraction of volatile compounds such as essential oils from plant materials. Thus, many studies have addressed the comparison of this technique with traditional extraction methods such as HD and Soxhlet extraction. The extraction of volatile essential oil from *Origanum onites* L. (Lamiaceae) with PHWE, SD, and Soxhlet extraction resulted in comparable yields of essential oil. Thus, PHWE is considered preferable as it provides advantages over the conventional techniques in velocity and cost and as an environmental friendly method. Interestingly, different temperatures within the liquid state limits of the water led to different concentrations of polar and nonpolar compounds, advantage that may be used for the selective extraction of different polarity compounds (Ozel and Kaymaz, 2004).

SWE at 50 bar and 150 °C was applied for the analysis of the volatile compounds from the fruits of *Amomum* spp. (Zingiberaceae); among them are sesquiterpenes such as copaene, β-elemene, caryophyllene, α-farnesene, α-caryophyllene, germacrene D, γ-elemene, α-cubebene, and β-sesquiphellandrene. The development of a PHWE–HS-SPME with GC-MS method for the analysis of these compounds provided the above-mentioned advantages comparing to the SD (Deng *et al.*, 2005).

In order to estimate the quality and the efficiency of the cedarwood oil's extraction method from *Juniperus virginiana* L. (Cupressaceae), PHWE was compared with SFE. The analysis of the three major sesquiterpene components of the cedarwood oil – cedrol, thujopsene, and cedrene – was performed by SFC. Cedrene seems to be the dehydration product of cedrol; thus, the optimization of both the extraction methods was processed, considering not only the overall oil yield but also the cedrol/cedrene ratio. Eller and Taylor concluded that, using temperatures over 100 °C during extraction, almost all the amount of cedrol was transformed into cedrene. Although the highest cedrol/cedrene ratio was obtained using PHWE at 50 °C under 750 psi,

the overall yield was fairly low comparing to that obtained by SFE (Eller and Taylor, 2004).

Aiming to optimize the conditions for the extraction of the essential oil from *Thymbra spicata* L. (Lamiaceae) four different temperatures were compared (100, 125, 150, and 175 °C), using a 30 min extraction method at a flow rate of 2 mL/min, 60 bar. When observing the three sesquiterpene constituents percentage of the essential oil of *T. spicata* – caryophyllene, spathulenol, and caryophyllene oxide – it was evident that, by altering the temperature, the percentage was increased while, regarding monoterpene constituents, an opposite behavior was observed (Ozel, Gogus, and Lewis, 2003). Moreover, the extraction of volatile essential oil from *Cuminum cyminum* L. (Apiaceae) at a high temperature of 150 °C by PHWE gave comparable yields with reference to Soxhlet extraction and SD (Yang *et al.*, 2007).

As mentioned above, PHWE, combing the advantages of PLE apparatus and the use of an environmental friendly solvent, the water, is tending to replace the traditional extraction methods for essential oils as it provides more automated and reproducible results. For instance, many articles in the literature support the use of PHWE as an alternative extraction method (Gamiz-Gracia and Luque de Castro, 2000; Ozel, Gogus, and Lewis, 2003; Tam *et al.*, 2007; Zhou *et al.*, 2007).

3.5 Supercritical Fluid Extraction (SFE)

SFE has been intensively used in the recent decades as an alternative to conventional solvent extraction of plant natural products. SFE, which possesses as general objective the reduction of organic solvent and the increase of sample throughput, presents many advantages as far as the cost, the speed of the experiment, the yield, and the quality of the obtained extracts are concerned (Huang *et al.*, 2007; Reverchon and De Marco, 2006; Zhannan *et al.*, 2009). Several solvents are used in SFE such as propane, ethane, hexane, pentane and butane, dimethyl ether, ammonia, and xenon and carbon dioxide (Pereira and Meireles, 2010). However, carbon dioxide (CO_2) seems to be the most attractive fluid for this technique, thanks to its properties regarding toxicity, flammability, and cost. Actually, CO_2 (critical conditions, $T_c = 30.9$ °C and $P_c = 73.8$ bar) is an inexpensive, safe, and environmentally friendly gas. Added to this, supercritical Sc-CO_2 is

highly diffusible, and its solvent's strength is easily tunable (Gonzalez-Coloma *et al.*, 2012). Moreover, CO_2 is gaseous at room temperature and pressure, simplifying significantly the recovery of analytes and providing solvent-free analytes, so that the expensive postprocessing of the extracts for elimination of solvents is avoided (Herrero *et al.*, 2010). The low polarity is the main drawback of this technique as it is a limitation for the extraction of polar compounds. To tackle this problem, organic solvents called polar modifiers or cosolvents can be added to change the polarity of the supercritical fluid and to increase its solvating power toward the analytes of interest (Casas *et al.*, 2007; Senseman and Ketchersid, 2000) (see **Supercritical Fluid Extraction**).

SFE has been proved to be a technique of choice for the extraction of sesquiterpenoids in comparison to conventional techniques; this being done mainly by optimizing the pressure and the temperature. For instance, the supercritical carbon dioxide extract of *Centipeda minima* (L.) A. Braun & Asch. (Asteraceae) was found to be enriched in sesquiterpenes lactones comparing to the conventional ethanol extract and lead to the isolation of six new sesquiterpene lactone with antiproliferative properties (Wu *et al.*, 2012). Similarly, an optimized SFE process lead to an extract richer in taginin C (an antiplasmodial sesquiterpene lactone) in comparison to Soxhlet and maceration extractions of *Tithonia diversifolia* (Hemsl.) A. Gray (Asteraceae) (Ziémons *et al.*, 2005). It has been demonstrated that the extraction with Sc-CO_2 provided a higher yield and a better quality of patchouli (*Pogostemon cablin* Benth., Lamiaceae) essential oil (richer in sesquiterpenoid patchoulol) than SD (Donelian *et al.*, 2009). Another comparative study has showed that the two main constituents of the leaves of the Yemeni *Schinus molle* L. (Anacardiaceae), the sesquiterpenes germacrene D and β-caryophyllene, are present in higher amount in the supercritical extract comparing to the hydrodistilled oil (Ali *et al.*, 2011); however, the total amount of the sesquiterpenes was higher for SFE than for HD, whereas the total amount of monoterpenes was similar in both cases. Most recently, volatile constituents of the aerial parts of *Capillipedium parviflorum* (R. Br.) Stapf (Poaceae) were studied by two different extraction procedures, HD and Sc-CO_2 extraction, and compared with headspace (HS) analysis. Monoterpenes were represented in low concentrations in HD oil and were totally absent in Sc-CO_2 and in HS analysis. Very little variability was observed

in the oxygenated monoterpenes in both Sc-CO_2 and HD oil, representing a low percentage of constituents. Sesquiterpene hydrocarbons were highly represented in Sc-CO_2 in comparison with HD and HS techniques, and oxygenated sesquiterpenes were represented in higher percentage in Sc-CO_2 than in HD oil and HS analysis (Saini *et al.*, 2012).

SFE has been applied for the extraction of sesquiterpenes and other terpenoids in many other plants such as lemongrass *Cymbopogon citratus* (DC.) Stapf (Poaceae) (Carlson *et al.*, 2001), marjoram *Majorana hortensis* Moench. (Lamiaceae) (Rodrigues *et al.*, 2002), *Thymbra spicata* L. (Lamiaceae) (Sonsuzer, Sahin, and Yilmaz, 2004), *Foeniculum vulgare* Mill. (Apiaceae) and *Thymus vulgaris* L. (Lamiaceae) (Diaz-Maroto *et al.*, 2005), *Helichrysum italicum* (Roth) G. Don (Asteraceae) (Ivanovic, Ristic, and Skala, 2011), *Calendula officinalis* L. (Asteraceae) (Kaškonienė *et al.*, 2011), and *Mentha spicata* L. (Lamiaceae) (Almeida, Mezzomo, and Ferreira, 2012). It is worth noting that essential oils have been the most extensively chemical group extracted with this method.

3.6 Microwave-Assisted Extraction (MAE) Techniques

MAE is a simple and modern leaching technique applied during the recent years for the extraction of various constituents from plant materials. Extraction assisted by microwaves is based on the fact that the energy is delivered rapidly to the total volume of the solvent and the solid plant matrix and results in their efficient and homogeneous heating. Subsequently, the heating of the solvent in contact with the sample accelerates the diffusion of the analytes from the sample matrix into the solvent (Hemalatha, Mandal, and Mohan, 2007). The effect of the microwave energy is strongly dependent on the nature of both solvent and solid matrix. Solvents possessing a high dielectric constant absorb microwave energy stronger, whereas the percentage of the water containing in the plant matrix enhance the extraction recoveries (Sparr Eskilsson and Bjorklund, 2000; Wang and Weller, 2006). The most important advantages of this method comparing with conventional extraction methods are the significant reduction in extraction time (usually to less than 30 min) and the significant decrease in organic solvent consumption,

resulting in high quality of extracts and mostly essential oils. Besides, MAE excels over the conventional methods in terms of cost and energy (Deng *et al.*, 2007). Two types of instruments are commercially available: the open and the close systems. Regarding the close systems, extraction is performed under controlled pressure and temperature, whereas extraction with open systems is performed under atmospheric pressure (Chan *et al.*, 2011; Kaufmann and Christen, 2002) (see **New Trends in Extraction of Natural Products: Microwave-Assisted Extraction and Pressurized Liquid Extraction**).

Microwaves are applied as a heating source during HD [microwave-assisted hydrodistillation (MAHD)], providing better results in terms of extraction time and slightly better results regarding extraction efficiency. This application of microwaves is widely used for the essential oil extraction. Golmakani and Rezaei, comparing the efficiency of MAHD and HD for the extraction of the essential oil from *Zataria multiflora* Boiss. (Lamiaceae), concluded that the overall yield of the essential oil was more or less the same, whereas the extraction time was significantly shorter when MAHD was applied. In addition, with the aid of a scanning electron microscope was observed that the leaves exposed to the microwaves were provoked a sudden eruption, resulting in the extraction of the essential oil in shorter time (Golmakani and Rezaei, 2008). Optimization of MAHD was performed for the isolation of rosemary essential oil (*R. officinalis* L., Lamiaceae). Using a fixed power of 900 W for 18 min, it was possible to obtain rosemary essential oil with a composition comparable to that reported in a monograph of the European Pharmacopoeia. In addition, rosemary essential oil by MAHD was compared with essential oils obtained by three different extraction techniques (SE, SFE, and HD) and characterized generally by a chemical profile very similar to the solvent-extracted and conventional hydrodistilled products. The supercritical fluid extract was characterized by the highest amounts of sesquiterpenes, probably because of the marked selectivity of CO_2 toward these compounds (Lo Presti *et al.*, 2005). Comparable results have been obtained from the analysis of volatile secondary metabolites from Colombian *Lippia alba* (Mill.) N.E. Brown (Verbenaceae) and *Xylopia aromatica* (Lamarck) Martius (Annonaceae) using HD, simultaneous distillation–solvent extraction (SDE), MAHD, and SFE. Among the four extracts, sesquiterpene compounds were in equal quantity when HD, MAHD, and SDE methods were applied, whereas SFE extract was consisted by higher amounts of sesquiterpenes (Stashenko, Jaramillo, and Martinez, 2004a, b).

Recently, microwave extraction was also performed at atmospheric pressure without any solvent or water addition. Solvent-free microwave extraction (SFME) is a combination of microwave heating and dry distillation technique where fresh plant materials are subjected to extraction in order to obtain mainly the containing essential oil. The green technique SFME is based on the heating of the *in situ* water within the plant material, resulting in the swelling of the glands and oleiferous receptacles and subsequently their burst. Thus, this process frees essential oil, which is evaporated by the *in situ* water of the plant material by azeotropic distillation (Lucchesi, Chemat, and Smadja, 2004a). Comparing three essential oils from different aromatic herbs obtained by SFME and conventional HD, Lucchesi *et al.* concluded that, although essential oils obtained by both methods were quantitatively the same, higher amounts of oxygenated compounds were observed in essential oils extracted by SFME. The presence of further oxygenated compounds in the SFME essential oils is probably due to the decrease of thermal and hydrolytic effects, compared with HD. The sesquiterpene constituents of the essential oils obtained by both extraction methods were qualitatively and quantitatively similar; however, in basil, essential oil (*Origanum basilicum* L., Lamiaceae) from SFME was poorer in terms of containing constituents (Lucchesi, Chemat, and Smadja, 2004b). In order to reduce further the extraction time, Wang *et al.* proposed the addition of an absorption solid medium (carbonyl iron powders) providing good microwave absorption capacity and good chemical stability. Improved SFME, when compared with conventional SFME, MAHD, and conventional HD for the extraction of essential oil from dried *C. cyminum* L. (Apiaceae) and *Zanthoxylum bungeanum* Maxim. (Rutaceae), exhibited the same percentage of sesquiterpene constituents, whereas the extraction time was significantly shorter (Wang *et al.*, 2006).

Appling the same concept of SFME, without the use of a solvent, microwave hydrodiffusion and gravity (MHG) technique have been developed for the extraction of essential oils from plant materials. This technique is based on a physical phenomenon, known as *hydrodiffusion*, which allows the extract to diffuse outside the plant material and drop by earth gravity

(Bousbia *et al.*, 2009a). MHG, similar to MAE technique, possesses many advantages in terms of extraction time, reproducibility, and environmental compatibility. The potential of the green MHG technique has been investigated for the isolation of essential oils from citrus peels and compared with conventional methods, HD and cold pressing (CP), giving comparable products (Bousbia *et al.*, 2009b). Similarly, the essential oils obtained by MHG and HD from *Mentha spicata* L., *M. pulegium* L., and *Rosmarinus officinalis* L. (Lamiaceae) were qualitatively and quantitatively the same (Bousbia *et al.*, 2009a; Vian *et al.*, 2008) (Figure 5).

4 ANALYSIS AND CHARACTERIZATION

4.1 Thin Layer Chromatography (TLC) and High Performance Thin Layer Chromatography (HPTLC)

Thin layer chromatography (TLC) is a simple, rapid, and versatile chromatographic method for the analysis of small terpenoids. Although, recently, a variety of more sensitive techniques have been developed for the analysis of these compounds, TLC remains a useful tool as a preliminary technique, including the initial screening and the quantitative evaluation of small terpenoids in plant extracts, providing valuable information for subsequent analyses (Kowalska, Sherma, and Waksmundzka-Hajnos, 2008) (see **Thin-layer Chromatography, with Chemical and Biological Detection Methods**). Moreover, most pharmacopoeial monographs of medicinal plants include TLC and, recently, HPTLC (Upton, 2010) as analytical techniques of reference for the quality control of herbal extracts based on chromatographic fingerprinting, a sequence of characteristic zones of substances (Widmer, Reich, and Ankli, 2008).

The separation of these compounds is performed mostly on silica or alumina stationary phase (Attaway, Barabas, and Wolford, 1965; Gocan, 2009), whereas the mobile phase is usually consisted of solvents of medium to low polarity according to the polarity of the analytes. Specifically, better resolution of different terpene derivatives was obtained by the impregnation of TLC layers with silver cations
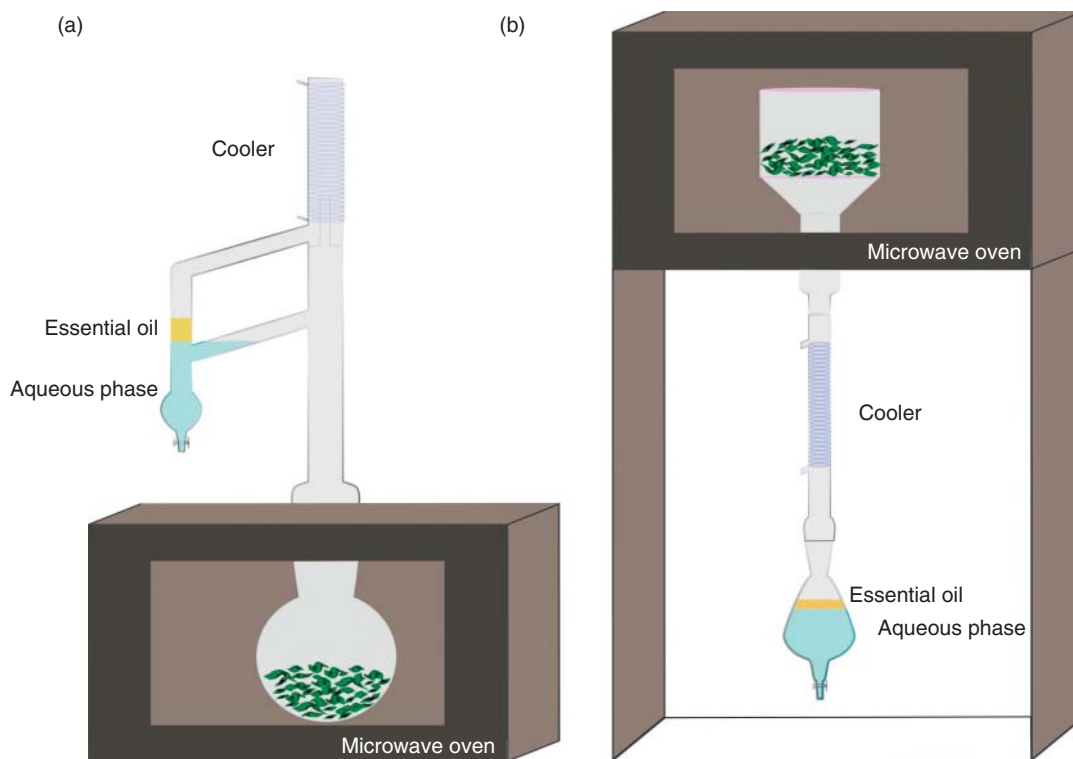
Figure 5   (a) Microwave-assisted hydrodistillation apparatus; (b) microwave hydrodiffusion and gravity apparatus.

($AgNO_3$). For example, the use of $AgNO_3$ enhanced the resolution and led to the successful isolation of sesquiterpene hydrocarbons (Lawrence, 1968) from *C. rotundus* L. (Cyperaceae) (Sonwa and König, 2001) and from different virgin olive oils (*Olea* sp., Oleaceae) (Bortolomeazzi *et al*., 2001).

The visualization of sesquiterpenoids is mainly carried out through a variety of visualization reagents providing different colors, whereas UV detection is only feasible when UV active groups are present (Attaway, Barabas, and Wolford, 1965; Glasl *et al*., 1999, 2001). Specific solution reagents have been reported for the detection of some categories, for instance, acidic reagents containing either vanillin or related benzaldehyde or benzoic acid derivatives are highly sensitive for sesquiterpene lactones (Picman *et al*., 1980), in addition to ethanolic solution of aluminum chloride (5%) (Villar *et al*., 1984). Furthermore, the colors of the revealed spots after the use of specific visualization reagents may indicate some structural features on the basic structure. For example, when analyzing different sesquiterpene lactones occurring in Asteraceae family, by means of a simple TLC method, it was possible to estimate several structural details within guaianolides, germacranolides, pseudoguaianolides, seco-pseudoguaianolides, and germacranes. The different colors revealed after the spray of TLC plates with anisaldehyde reagent were correlated with the diverse substitution patterns of the different sesquiterpene lactone structures (Nowak *et al*., 2011). Similarly, the color of the TLC spots after spraying with concentrated sulfuric acid was associated to the position of the ester group on the basic structure of guaianolides (Nowak, 1993).

In recent years, high performance thin layer chromatography (HPTLC), an advanced TLC technique, is increasingly being used as an analytical tool for the fast analysis of plant natural products. Providing a variety of advantages such as simplicity, rapidity, reproducibility, and cost effectiveness, HPTLC enables high speed separation and qualitative identification of natural products in complex mixtures. Therefore, HPTLC methods have been developed for the fingerprinting and profiling of plant extracts and for the quantification of marker compounds in the extracts in order to evaluate and standardize their quality (Reich and Widmer, 2009).

Quantitative determination by HPTLC is generally performed by scanning densitometry. Owing to the high degree of automation and robustness, baseline separation could be achieved, leading to reproducible results (Gocan and Cimpan, 2005; Shinde, Chavan, and Wakte, 2011). For instance, the quantitative determination of zerumbone, a pharmacologically active monocyclic sesquiterpenoid, in the rhizomes, roots, leaves, and flowers of *Zingiber zerumbet* Smith (Zingiberaceae) was performed through a validated HPTLC method. Aluminum foil-backed silica gel 60 F254 HPTLC plates were developed using a mobile phase of ethyl acetate–hexane (1.5 : 8.5), whereas UV detection was performed densitometrically at 250 nm (Rout, Mishra, and Sherma, 2009). Similarly, the quantitative determination of two major diterpene plant growth promoters in *Callicarpa macrophylla* Vahl (Lamiaceae), calliterpenone and calliterpenone monoacetate, was performed by both HPLC and HPTLC validated methods. Standard and extract samples were developed on normal-phase HPTLC plates using a mobile phase of ethyl acetate : hexane (55 : 45), and detection was performed after derivatization with vanillin–sulfuric acid reagent densitometrically, at 610 nm. HPTLC methods are considered suitable for rapid screening of plant samples, whereas HPLC methods provide better precision (Verma *et al.*, 2009). In order to evaluate the quality of different absinthes, the quantity of absinthin in various origin absinthes (*Artemisia absinthium* L., Asteraceae) was determined using an HPTLC method, on normal phase plates, developed with acetone/acetic acid/toluene/dichloromethane (10 : 10 : 30 : 50). After visualization with anhydride/sulfuric acid/ethanol (10 : 10 : 100) solution, detection of absinthin was performed densitometrically at 554 nm (Lachenmeier, 2007).

Apart from HPLC methods, artemisinin and other analogs have been quantitatively determined in *Artemisia annua* L. (Asteraceae) extracts (Bhandari *et al*., 2005; Widmer, Handloser, and Reich, 2007) and pharmaceutical products (Gabriels and Plaizier-Vercammen, 2003; Gabriels and Plaizier-Vercammen, 2004) by HPTLC using different stationary and mobile phases. For instance, Bhandari *et al*. proposed a reversed-phase (RP)-HPTLC method for the simultaneous quantification of artemisinin together with its precursors arteannuin-B and artemisinic acid using 0.2% TFA in water/ACN (35 : 65) as a mobile phase. The quantitative determination of the compounds of interest in *A. annua* extracts was performed after the derivatization with anisaldehyde reagent, at 426 nm (Bhandari *et al*., 2005).

4.2   Gas Chromatography (GC)

GC is a traditional chromatographic method that enables volatiles separation and identification although supplementary evidence is frequently required for reliability purposes, avoiding equivocated characterizations. Generally, regarding sesquiterpenes and monoterpenes, GC is commonly used for the analysis of their volatile fraction, which is the main constituent of essential oils (Table 3), whereas their nonvolatile fraction such as sesquiterpenes lactones, pyrethrins and iridoids, is principally analyzed by LC. As in any chromatographic separation, the main goal is to resolve all the compounds of interest in a minimum time, and a critical factor is the selection of a suitable analytical column (dimension and stationary phase type). Peak enlargement and broadening phenomena are reduced by applying adequate chromatographic parameters (Özek and Demirci, 2012).

In the early development of the research in the essential oil field, more attention has been given to the profiles of volatiles. However, this task was difficult owing to the complexity of these real-world samples. Nevertheless, the recent improvements in instrumental analytical chemistry and especially in the area of chromatography have been significantly beneficial towards this direction, and nowadays, the number of known constituents has increased considerably. Moreover, the requirements for high resolution and trace analysis are satisfied by modern column technology. For instance, inert, thermostable, and efficient open-tubular columns are available, along with associated selective detectors and injection methods, which allow on-column injection of liquid and thermally labile samples (d'Acampora Zellner *et al.*, 2010). Since the introduction of capillary columns in GC, packed columns are rarely used. The main advantage of packed columns is that they can support larger sample size ranges (10–20 µL), increasing, therefore, the dynamic range of the analysis even for trace-level components with no need of prior fractionation or concentration. However, their low permeability represents a significant drawback as it requires high pressure to significantly improve their low resolution; therefore, higher column flow rates are required. The chemical nature of the stationary phase and film thickness, in addition to the column length and internal diameter, are to be considered in the choice of capillary column in GC analysis of

volatiles. Generally, essential oil GC analyses are carried out on 25–50 m columns, with 0.20–0.32 mm internal diameters, and 0.25 µm stationary phase film thickness. Mainly capillary columns with dimethyl polysiloxane (methyl silicone) nonpolar and Carbowax 20M polar phases are used. Carbowax 20M phases include DB-Wax, BP-20, PEG 20M, and HP 20 phases, whereas methyl silicone phases include SE-30, SF-96, OV-1, OV-101, BP 1, CP-Sil 5CB, SP 2100, DB 1, DB 5, and HP 1 phases. Among these, DB 1 or DB 5 and CP-Sil 5 are the mostly used fused-silica capillary GC columns (Merfort, 2002).

On nonpolar columns, volatiles are generally separated according to their boiling point, whereas, for nonpolar compounds, the separation is polarity based. Therefore, volatile components with similar boiling points elute in a narrow retention time range on nonpolar columns. Aiming to overcome this limit, the analytical method can be modified by applying a slower oven temperature rate to widen the elution range of the oil or using a polar stationary phase, as oxygenated compounds are better retained compare to hydrocarbons. However, choosing different stationary phases may provide a little improvement, as resolution can be enhanced for a series of compounds but new coelutions could be generated (Marriott, Shellie, and Cornwell, 2001).

The most commonly used GC detectors are flame ionization detector (FID), thermal connectivity detector (TCD), electron capture detector (ECD), nitrogen/phosphorus detector (NPD), photoionization detector (PID), atomic emission detector (AED), mass selective detector (MSD), flame photometric detector (FPD), and chemiluminescence detector (CD) (Özek and Demirci, 2012). Because these detectors do not provide structural information for the analyzed molecules and retention data, more precisely retention indices are used as primary criterion for peak assignment. Moreover, the need to express gas chromatographic retention data in a standardized system has been recognized, and retention index values proved to be a valuable method (Zellner *et al.*, 2008). The *retention index* (*I*) in respect to the homologous series of *n*-alkanes, proposed by Kováts, is a useful tool for the preliminary identification of volatiles components (Kováts, 1958). The drawback of the Kováts retention index is that the GC is supposed to operate in isothermal conditions, limiting the analysis to a narrow range of boiling points as the early peaks may emerge rapidly and can overlap, whereas higher boiling compounds emerge

Table 3 GC analysis of various sesquiterpenoids and other terpenoids.

Main categories	Plant material	GC column dimensions	Temperature program	Detector conditions	Purpose	References
Sesquiterpenes and monoterpenes	*Ageratina adenophora* (Spreng.) King & H. Rob. (aerial part, essential oil)/Asteraceae	50 m × 0.25 mm × 0.25 μm DB-1	80 to 225 °C, 4 °C/min	FID (300 °C)	Identification and quantification	Palá-Paúl *et al.* (2002)
		50 m × 0.25 mm × 0.25 μm SE-30 and BP-5	70 to 240 °C, 4 °C/min	MD 800 MS, 70 eV, scan mode (35–350U)	Identification	
46 compounds (mostly monoterpenes and sesquiterpenes)	*Cymbopogon flexuosus* (Nees ex Steud.) J.F. Watson (leaves, essential oils),	2 × (50 m × 0.22 mm × 0.25 μm) BP-1 and BP-20	60 to 220 °C, 2 °C/min (20 min)	2 FID (250 °C)	Characterization	Ottavioli *et al.* (2009)
	Cymbopogon tortilis (J. Presl) A. Camus (leaves, essential oils)/Poaceae	60 m × 0.22 mm × 0.25 μm Rtx-1	60 to 230 °C, 2 °C/min (45 min)	Quadrupole EIMS, 70 eV (35–350 Da)		
Monoterpenes, sesquiterpenes and C$_{12}$ sesquiterpenes, and phenyl propanoids	*Pimpinella aurea* D.C., *Pimpinella corymbosa* Boiss., *Pimpinella peregrina* L., *Pimpinella puberula* (DC.) Boiss. (fruits, stems and leaves and roots/essential oils)/Apiaceae	60 m × 0.25 mm × 0.25 μm HP-Innowax FSC	60 °C (10 min), 4 °C/min, 220 °C (10 min), 1 °C/min, 240 °C	FID (250 °C)	Quantitative comparison	Tabanca *et al.* (2005)
				GC-MS, 70 eV, Scan range (*m/z* 35–425)		
Monoterpenes and sesquiterpenes	*Valeriana officinalis* L. (roots)/ Valerianaceae	30 m × 0.25 mm × 0.25 μm DB-5 and 60 m × 0.25 mm × 0.25 μm DB-Wax	40 °C (5 min), 5 °C/min, 280 °C	FID (290 °C)	Comparative study of extraction methods	Safaralie, Fatemi, and Sefidkon (2008)
		30 m × 0.25 mm × 0.25 μm DB-5	60 to 260 °C, 3 °C/min	Varian 3400 GC-MS, 70 eV, Scan range (40–400U) transfer line temperature at 280 °C		

(*Continued overleaf*)

Table 3 *(continued)*

Main categories	Plant material	GC column dimensions	Temperature program	Detector conditions	Purpose	References
Monoterpenes, sesquiterpenes, and diterpenes	Anthemis species (14 essential oil samples)/Asteraceae	30 m × 0.32 mm × 0.25 μm Supelcowax-10	75 to 250 °C, 2.5 °C/min (20 min)	FID (230 °C)	Qualitative and quantitative analyses and comparison	Saroglou et al. (2006)
		30 m × 0.25 mm × 0.25 μm HP-5	60 °C (5 min), 4 °C/min, 280 °C	EIMS, 70 eV		
		30 m × 0.25 mm × 0.50 μm HP-Innowax	60 to 260 °C, 3 °C/min			
Sesquiterpenes and monoterpenes	Matricaria chamomilla L. (essential oil)/Asteraceae	5 m × 0.1 mm × 0.1 μm	40 °C (0.1 min), 30 °C/min, 95 °C; 35 °C/min, 155 °C; and 200 °C/min, 280 °C (0.5 min); total analysis: 5 min	High frequency FID (300 Hz at 250 °C); compressed air: 350 mL/min; hydrogen: 35 mL/min; nitrogen: 30 mL/min	Analysis and quantification	Heuskin et al. (2009)
	Nepeta cataria L. (essential oil)/Lamiaceae					
25 sesquiterpenes and 10 monoterpenes	Matricaria chamomilla L. (essential oil)/Asteraceae	30 m × 0.32 mm × 0.25 μm	40 °C (5 min), 5 °C/min, 230 °C and 30 °C/min, 280 °C (5 min)	EIMS, 70 eV, full-scan (35–350 amu)	Separation and identification	Heuskin et al. (2009)
8 sesquiterpenes and 11 monoterpenes	Nepeta cataria L. (essential oil)/Lamiaceae					
Bilobalide and 6 ginkgolides	Ginkgo biloba L./Ginkgoaceae	10 m × 0.53 mm × 0.88 μm, HP-1	250 °C (2 min), 5 °C/min, 300 °C (1 min); total analysis: 13 min	Compressed air: 400 mL/min; hydrogen: 45 mL/min; nitrogen: 25 mL/min	Determination and quantification	Sayadi et al. (2010)
Monoterpenes and sesquiterpenes	Juglans regia L., Juglans nigra L., Juglans sieboldiana var. cordiformis Lam. (leaves)/Juglandaceae	30 m × 0.25 mm × 0.25 μm HP-5MS	45 °C (2 min), 8 °C/min, 300 °C (15 min)	SIM mode	Qualitative and quantitative analyses and comparison	Fojtová, Lojková, and Kubáň (2008)

Analytes	Sample/Family	Column	Temperature program	Detector	Objective	Reference
Monoterpenes, sesquiterpenes, and derivatives	*Vetiveria zizanioides* (L.) Nash (root, thai vetiver oil)/Poaceae	Set 1: 1D (30 m × 0.25 mm × 0.25 µm) BPX5; 2D (0.8 m × 0.1 mm × 0.1 µm) BP20	120 to 180 °C, 2 °C/min (10 min)	GC × GC-FID (100 Hz) GC × GC-qMS (20 Hz), mass range (m/z 40–240 Hz)	Separation, comparison and relative quantification of different extraction methods, identification and comparison of different extraction methods	Pripdeevech, Wongpornchai, and Marriott (2010)
		Set 2: 1D (30 m × 0.25 mm × 0.25 µm) Solgel Wax; 2D (1 m × 0.15 mm × 0.25 µm) BP1	120 to 260 °C, 2 °C/min (10 min)			
		Set 3: 1D (20 m × 0.25 mm × 0.25 µm) BPX5; 2D (1 m × 0.1 mm × 0.1 µm) BP20	120 to 180 °C, 2 °C/min (10 min)			
Monoterpenes and sesquiterpenes	*Daucus carota* L./Apiaceae	50 m × 0.25 mm × 0.2 µm CP-Wax 52CB	31 °C (1 min), 1.5 °C/min, 80 °C; 1 °C/min, 125 °C; 18 °C/min, 190 °C (10 min)	FID	Determine the volatile composition and aroma active components of carrots stored under different temperature conditions	Kjeldsen, Christensen, and Edelenbos (2003)
		50 m × 0.25 mm × 0.25 µm b-cyclodextrin Chrompack	32 °C (1 min), 2 °C/min, 92 °C (16 min); 2 °C/min, 145 °C (10 min); 18 °C/min, 180 °C (10 min)	GC-MS mass range (m/z 39–350) and GC-MS/MS		
Mostly monoterpenoids	*Myrtus communis* L. (berries from 10 localities of Corsica, France/Myrtaceae	2 × (60 m × 0.22 mm × 0.25 µm) Rtx-1 and Rtx-Wax	60 to 230 °C, 2 °C/min (35 min)	Dual FID	Identification and quantification of volatile components	Barboni et al. (2010)
		2 × (60 m × 0.22 mm × 0.25 µm) Rtx-1 and Rtx-Wax	60 to 230 °C, 2 °C/min (35 min)	EI-qMS, 70 eV, mass range (m/z 35–350 Da)		

(Continued overleaf)

Table 3 (*Continued*)

Main categories	Plant material	GC column dimensions	Temperature program	Detector conditions	Purpose	References
Mostly monoterpenoids	Mexican *Bursera* (volatile fractions from the bark of 8 species)/ Burseraceae	25 m × 0.20 mm × 0.33 µm HP-2 Ultra	40 °C (5 min), 8 °C/min, 305 °C (10 min)	EIMS, 70 eV	Identification	Zúñiga *et al.* (2005)
Oxygenated monoterpenes	*Pistacia lentiscus* var. *chia* L. (mastic water extracts obtained from mastic oil)/Anacardiaceae	30 m, DB-5 ———— b-Dex-sm chiral column	60 °C (5 min), 3 °C/min, 280 °C (15 min)	EIMS	Chemical characterization	Paraschos *et al.* (2011)
Sesquiterpenes (α-pinene, β-elemene, curcumol, germacrone, curdione)	Rhizomes Curcumae & Radix Curcumae/ Zingiberaceae	30 m × 0.25 mm × 0.25 µm HP-5MS	50 °C, 20 °C/min, 150 °C; 2 °C/min, 180 °C; 20 °C/min, 200 °C (3 min)	SIM mode, 70 eV	Quantitative analysis and comparison of essential oils from both the plants	Lu *et al.* (2011)

as flat and often immeasurable peaks (Castello, Moretti, and Vezzani, 2009). When separation and identification of the components of a complex mixture cannot be obtained in isothermal conditions due to their extended boiling point range, temperature programming is necessary; at low initial temperature, the early peaks are well resolved and the high boiling compounds eluted in a shorter time as sharp peaks. In this case, the total analysis time is considerably reduced. The mostly used programmed temperature retention index is the one proposed by Van den Dool and Kratz (Hoigné, Windmer, and Gäumann, 1963; van Den Dool and Kratz, 1963). In the literature, the *Kovats index* (*KI*) is also called *retention index* (*I*), whereas the index obtained from Van den Dool and Kratz equation is called *retention index* (*I*), *linear retention index* (*LRI*), or *programmed-temperature retention index* (*PTRI* or I_T). The use of the retention indices is of great importance for the comparison of retention data obtained in different laboratories using various experimental conditions as they are nearly independent to the parameters of the GC analysis.

A clear example of the evolution of the GC separation is the study of the essential oil from rue (*Ruta graveolens* L., Rutaceae), a medicinal and aromatic plant (Kubeczka, 1981). Bruno, in 1961, succeeded, for the first time, to analyze this essential oil and separate eight constituents using GC (Bruno, 1961). A couple of years later, in 1964, the same oil was analyzed in an improved manner using a Perkin Elmer GC equipped with a 2 m packed column and a TCD operated under isothermal conditions, yielding 20 separated constituents. After the introduction of temperature programming of the column oven, an additional improvement of the separation of this oil was achieved, yielding approximately 80 constituents. The last noteworthy improvements were a result of the development of high resolution capillary columns and the sensitive FID detectors. Using a 50 m glass capillary with 0.25 mm I.D., the rue oil could be separated into approximately 150 constituents, in 1981 (Kubeczka, 2010).

4.2.1 High Speed Gas Chromatography (HSGC)

High speed separations have represented a major development in high speed gas chromatography (HSGC) the recent years. The various concepts for these techniques have been nicely reviewed by Cramers *et al*. (1999). Using high speed separations, the analysis time could be considerably reduced, whereas the sample throughput is increased. Specifically, in order to reduce the analysis time without loosing separation efficiency, the HSGC technique should be performed by manipulating a number of analysis parameters, such as column length, column I.D., stationary phase, film thickness, carrier gas, linear velocity, oven temperature, and ramp rate. Different terms are used in the literature to categorize the capillary gas chromatographic analysis. These classifications are based either on GC analysis time or on column internal diameter types. In conventional GC analysis, the separation is performed in more than half an hour or 0.25 mm I.D. columns; fast GC analysis in minute range or 0.10–0.18 mm I.D. columns; very fast GC analysis in seconds; and ultrafast GC in subseconds or columns of 0.05 mm I.D. or even less.

Bicchi *et al*. have reported a comparison of fast and conventional GC analysis of juniper essential oil (*Juniperus communis* L., Cupressaceae), showing that the two GC profiles are fully superimposable while analysis time decreases from about 60 min to about 20 min. The analysis conditions were the following: (a) conventional GC-MS; column, MEGA 5 (25 m × 0.25 mm I.D. × 0.25 μm); temperature program, from 40 °C (1 min) to 250 °C (5 min) at 3 °C/min; (b) fast GC-MS: column, MEGA 5 (10 m × 0.10 mm I.D. × 0.10 μm); temperature program, from 40 °C (1.46 min) to 250 °C (1.46 min) at 10.25 °C/min. The 21 following targeted compounds were clearly identified in both fast and conventional GC analysis cases: monoterpenes α-thujene, α-pinene, β-pinene, sabinene, myrcene, α-terpinene, *p*-cymene, limonene, β-phellandrene, terpinolene and terpinen-4-ol and sesquiterpenes, α-cubebene, α-copaene, β-elemene, β-caryophyllene, germacrene D, δ-cadinene, germacrene B, germacran-D-4-ol, spathulenol and α-cadinene (Bicchi, Cagliero, and Rubiolo, 2011).

Another comparative study between fast and conventional GC was performed by Mondello *et al*. on citrus essential oils, a very complex mixtures that can be divided into two fractions: a 90–95% volatile fraction (monoterpenes and sesquiterpenes hydrocarbons and their oxygenated derivatives along with aliphatic aldehydes, alcohols, and esters) and a 1–10% nonvolatile one (hydrocarbons, fatty

acids, sterols, carotenoids, waxes, coumarins, psoralens, and flavonoids) (Mondello *et al*., 2003). Five cold-pressed citrus essential oils, bergamot, bitter orange, sweet orange, mandarin, and lemon, were analyzed. For all citrus oils analyzed, the fast GC analysis technique using a 10 m × 0.1 mm I.D. × 0.1 µm film thickness column, with a temperature program 50–250 °C at 14 °C/min presented almost the same separation performance with the conventional GC analysis using 30 m × 0.25 mm I.D. × 0.25 µm film thickness with a temperature program of 50–250 °C (10 min) at 3 °C/min with a speed gain factor of about 5. Quantitative data also showed good reproducibility, and these results demonstrated the effectiveness of fast GC applications through the use of narrow bore capillary columns, in the separation of very complex matrixes.

4.2.2 *Chiral Gas Chromatography*

It is well known that terpenoids and specially mono-, sesqui-, and diterpenes possess generally chiral centers and consequently many enantiomers, possibly exhibiting different biological activities or may influence the sensory impression. For instance, the (+)-enantiomer of nootkatone exhibits the bitter taste of grapefruit and the typical grapefruit flavor, which is 2200 times more intense than that of the (−)-enantiomer (Haring *et al*., 1972). Only the (+)-enantiomer of α-vetivone from vetiver essential oil is responsible for the specific woody, floral, balsamic fragrance. The investigation of a chiral compound in terms of stereochemistry is, therefore, of great importance (König *et al*., 1994). Therefore, the introduction of enantioselective capillary columns with high separation efficiency was a prominent development in GC analysis area (Gil-Av, Feibush, and Charles-Sigler, 1966).

An optimum performance was reached after the development of Chirasil-Val, a methyl-polysiloxane phase containing about 6% branched aliphatic side chains with L-valine in diamide linkage, and similar polymeric chiral diamide stationary phases (Frank, Nicholson, and Bayer, 1977), which possessed high thermal stability. Separation of enantiomers by hydrogen-bonding chiral stationary phases usually requires derivatization of the analytes (e.g., formation of carbamates, cyclic carbonates, or oximes) in order to increase its volatility and/or to introduce

suitable functions for additional hydrogen bonding and to improve the detection of trace amounts of enantiomers (Adlard, 1987). Another approach in chiral separation is based on the diastereomeric association between chiral molecules and chiral transition metal complexes, which was firstly described by Schurig (Schurig, 1977). In this approach, hydrogen bonding interaction is not essential for chiral recognition, and a number of compounds can be separated. However, the nonsufficient thermal stability of the applied metal complexes was a serious drawback for this approach.

An important breakthrough in the GC separation of enantiomers was the use of modified cyclodextrins as chiral stationary phases in high resolution capillary GC. The use of cyclodextrin derivatives that act enantioselectively by partial intrusion of enantiomers into the cyclodextrin cavity was described in 1988 (König, Lutz, and Wenz, 1988b). Another crucial step in chiral GC was the development of non polar acetylated cyclodextrins (Kubeczka, 2010). König *et al*. reported their first results in 1988 with per-*O*-pentylated and selectively 3-*O*-acylated-2,6-di-*O*-pentylated α-, β-, and γ-cyclodextrins, which are highly stable, soluble in nonpolar solvents, and possess a high enantioselectivity toward many chiral compounds. Until now, a number of additional cyclodextrin derivatives with improved thermal stability have been synthesized, allowing the resolution of a large number of enantiomeric pairs from various classes including chiral monoterpene hydrocarbons and sesquiterpenes (König *et al*., 1988a).

Enantioselective GC in all its different forms found a wide variety of applications (Table 4). Biochemical studies into the mechanisms of terpenoid biosynthesis (Bohlmann, Meyer-Gauen, and Croteau, 1998; Wagschal *et al*., 1994), discovery of genetic markers such as monoterpenes of *Larix* species (Pinaceae) (Holm and Hiltunen, 1997), and studies into the biogenesis of the essential oils of *A. annua* L. (Asteraceae) (Holm, Vuorela, and Hiltunen, 1997b) reflect other diverse uses of enantiomeric separations. Additional investigations on sesquiterpenes were presented by König *et al*. (König *et al*., 1994). Moreover, Bicchi *et al*., after a study with columns, concluded that particular chiral separations using certain cyclodextrine derivatives (CDD) preferentially resolved certain enantiomers. For instance, a 2,3-di-*O*-ethyl-6-*O*-*tert*-butyldimethylsilyl-β-CD on polymethylphenylsiloxane (PS086) phase allowed

Table 4 Various applications of enantioselective GC.

Plant material	Family	Application	References
Coffea arabica L.	Rubiaceae	Estimation of enantiomeric ratio of epoxygeraniols	Emura *et al.* (1997)
Mentha species	Lamiaceae	Enantiomeric distribution of piperitone in essential oils	Ravid, Putievsky, and Katzir (1994)
Calamintha incana (sm.) Heldr.	Lamiaceae		
Artemisia judaica L.	Asteraceae		
Pelargonium graveolens L'Hér.	Geraniaceae	Enantiomeric composition of citronellol in essential oils	Ravid *et al.* (1992)
Eucalyptus citriodora (Hook.) K.D. Hill & L.A.S.Johnson	Myrtaceae		
Mentha citrata Ehrh.	Lamiaceae		
Rosmarinus officinalis L.	Lamiaceae	Enantiomeric distribution of verbenone in essential oils	Ravid *et al.* (1997a)
Ocimum species and basil oils	Lamiaceae	Enantiomeric composition of linalool in essential oils	Ravid *et al.* (1997b)
Angelica archangelica L.	Apiaceae	Enantiomeric composition of chiral monoterpene hydrocarbons	Holm *et al.* (1997a)
Pelargonium graveolens L'Hér.	Geraniaceae	Enantiomeric distribution of *cis*- and *trans*-rose oxide ketones in geranium oils of different origin	Wüst *et al.* (1997)
Pelargonium hybrid	Geraniaceae	Measuring the chirality of geographically and seasonally different geranium oils	Doimo, Fletcher, and D'Arcy (1999)
Oils from *Citrus* species	Rutaceae	Chirality evaluation and authenticity profile of neroli and petitgrain oil	Juchelka *et al.* (1996)
Lavandula species	Lamiaceae	Quality assessment of essential oil	Kreis and Mosandl (1992)
Melaleuca alternifolia (Maiden & Betche) Cheel and *Melaleuca linariifolia* Smith	Myrtaceae	Origin of (+)-δ-cadinene and cubenols in the essential oils	Cornwell *et al.* (2000)
Araucaria araucana (Molina) K. Koch, *Araucaria heterophylla* (Salisb.) Franco, and *Araucaria bidwillii* Hook	Araucariaceae	Enantiomeric composition of sesquiterpene and diterpene hydrocarbons in foliage oils	Pietsch and König (2000)
Citrus x limon (L.) Burman	Rutaceae	Applications of cyclodextrin derivatives in GC enantioseparations of essential oil components	Bicchi *et al.* (1997b)
Iris pallida Lam.	Iridaceae		
Lavandula angustifolia P. Miller	Lamiaceae		
Mentha x piperita L.	Lamiaceae		
Rosa x damascena P. Miller	Rosaceae		
Pinus sylvestris L.	Pinaceae	The enantiomeric composition of monoterpene hydrocarbons of pine needle oils	Hiltunen and Laakso (1995)
Juniperus communis L.	Cupressaceae		
Abies sachalinensis (Fr. Schmidt) Mast, *Abies mayriana* Miy et Kudo	Pinaceae	Enantiomeric composition of monoterpene hydrocarbons as chemotaxonomic markers of conifer needle essential oils	Holm, Laakso, and Hiltunen (1994)
Picea abies (L.) H. Karst.	Pinaceae	Enantiomeric composition of the seven major chiral monoterpenes	Persson *et al.* (1996)
Agathis spp. and *Araucaria* spp.	Araucariaceae	Enantiomeric composition of the monoterpenes, α-pinene, β-pinene, camphene, and limonene in resins	Wang *et al.* (1997)

Table 4 (*Continued*)

Plant material	Family	Application	References
Oils from *Citrus* species	Rutaceae	Characteristic authenticity profile	Mosandl and Juchelka (1997)
Cymbopogon winterianus Jowitt	Poaceae	Enantiomeric distribution of limonene, linalool, citronellal, and β-citronellol	Lorenzo *et al.* (2000)
Cryptomeria japonica (L.f.) D. Don	Cupressaceae	Isolation and identification of chiral diterpenes from essential oils	Pietsch and König (1997)
Araucaria araucana (Molina) K. Koch	Araucariaceae		
Thuja occidentalis L.	Cupressaceae		
Podocarpus spicatus Poepp.	Podocarpaceae		
Pinus pseudostrobus Lindl.	Pinaceae		
Lophocolea bidentata (L.) Dumort.	Geocalycaceae		
Pellia epiphylla (L.) Corda	Pelliaceae		
Buddleja tucumanensis Griseb.	Scrophulariaceae	Enantiomeric distribution of α-pinene, sabinene, β-pinene, limonene, and terpinen-4-ol, helping in chemotaxonomy	Lorenzo, Loayza, and Dellacassa (2006)
Cinnamomum osmophloeum Kaneh	Lauraceae	Essential oil profile and enantiomeric purity	Cheng *et al.* (2012)
Solidago canadensis L.	Asteraceae	Enantiomeric distribution of bioactive chiral terpenoids in the essential oil	Chanotiya and Yadav (2008)
Zanthoxylum schinifolium Siebold. & Zucc.	Rutaceae	Enantiomeric separation and composition of α-pinene, nerolidol, β-pinene, linalool, limonene, and citronellal	Seo *et al.* (2009)

the characterization of lavender (*Lavandula angustifolia* P. Miller, Lamiaceae) and citrus [*Citrus limon* (L.) Burman, Rutaceae] oils containing linalyl oxides, linalool, linalyl acetate, borneol, bornyl acetate, α-terpineol, and *cis*- and *trans*-nerolidols, whereas peppermint oil was better analyzed using a 2,3-di-*O*-methyl-6-*O*-*tert*-butyldimethylsilyl-β-CD on PS086 phase, and α- and β-pinenes, limonene, menthone, isomenthone, menthol, isomenthol, pulegone, and methyl acetate were identified (Bicchi *et al*., 1997b). A thorough investigation of the stereochemical correlations of terpenoids was performed by Konig; this study led to the conclusion that when using a heptakis (6-*O*-methyl-2,3-di-*O*-penthyl)-β-CD and octakis (6-*O*-methyl-2,3-di-*O*-penthyl)-γ-CD in polysiloxane, the separation of both enantiomers is common for monoterpenes, less common for sesquiterpenes, and never observed for diterpenes (König, 1998).

Enantioseparation by high resolution capillary gas chromatography (HRC-GC) presents some advantages such as high efficiency, sensitivity and speed of separation, simple detection, unusually high precision and reproducibility, and small amount of sample requirement. Moreover, its main use is related with the characterization of the enantiomeric composition and the determination of the enantiomeric excess and/or ratio of chiral research chemicals, intermediates, metabolites, flavors and fragrances, drugs, pesticides, fungicides, herbicides, pheromones, to name just a few (Dugo *et al*., 1992). However, the sesquiterpenes fractions of many essential oils are complex mixtures with many overlapping peaks in the gas chromatograms. The identification of the enantiomeric composition of sesquiterpene hydrocarbons or enantiomers directly from such a mixture is almost impossible. Therefore, in many cases two- or multidimensional gas chromatography (MDGC) has to be performed with the analysis been more reliable.

4.2.3 Two-Dimensional and Multidimensional Gas Chromatography (2DGC and MDGC)

Generally, in chromatography, separations could be characterized as two- (2D) or multidimensional (MD) when separation of all or some selected groups of the sample's components is repeated in two or more analytical chromatographic columns of different selectivity (Giddings, 1995). Therefore, each dimension of separation is associated with a specific type of stationary phase and with a specific molecular interaction developed between this stationary phase and the analyte solute. In MDGC, key fractions of a sample are selected from the first column and reinjected onto a second one, where, ideally, they should be fully resolved. The instrumentation usually involves the use of a switching valve arrangement and two chromatographic columns. The capillary columns used can be operated in either a single or two distinct GC ovens, with both GC systems commonly equipped with detectors.

In conventional 2DGC using the *heart-cutting* technique, many separations of chiral oil constituents have been performed in the past. For instance, the investigation of the chiral sesquiterpene hydrocarbon germacrene D was found to be the main constituent of the essential oil from the flowering herb of *Solidago canadensis* L. (Asteraceae). The enantioselective investigation of the germacrene D fraction from a GC run using a nonchiral DB-Wax capillary transferred to a 2,6-methyl-3-pentyl-β-cyclodextrin capillary revealing the presence of both enantiomers. This is worth mentioning, because, in most of other germacrene D containing higher plants, nearly, exclusively, the (−)-enantiomer can be found (Kubeczka, 2010).

This previously mentioned 2DGC design, however, has some shortcomings mainly because of the valve used to direct the portion of desired effluent from the first into the second column. The sample comes into contact with the metal surface of the valve body, the pressure drop of both connected columns may be significant, and the use of only one-column oven does not allow a proper adjustment of the temperature for both columns. Hence, to tackle this issue, one of the best approaches has been the development of a two-column oven instrument using a Deans-type pressure balancing interface between the two columns called a "live-T-connection," providing considerable flexibility (Korytár *et al*., 2002). Using this instrument, the enantiomeric composition of several volatile mixtures has been investigated very successfully. As an example, the investigation of the essential oil from *L. angustifolia* P. Miller (Lamiaceae) shall be mentioned, showing the simultaneous stereoanalysis of a mixture of chiral compounds, which can be found in lavender oil, using the column combination Carbowax 20M as precolumn and 2,3-di-*O*-acetyl-6-*O*-*tert*-butyldimethylsilyl-b-cyclodextrin as main column. All the unresolved enantiomeric pairs from the precolumn could be

well separated after transferring them to the chiral main column in a single run. As a result, it was found that most of the characteristic and genuine chiral constituents of lavender oil exhibit a high enantiomeric purity (Kreis and Mosandl, 1992).

Although genius separations can be performed using this so-called *heart-cutting technique*, drawbacks are related to the partial analysis of the sample in the two different dimensions and to the increased analysis time. Introduced in the early 1990s by Liu and Phillips, comprehensive 2DGC (GC × GC) was designed to overcome these limitations by producing a high frequency heart-cutting separation of the entire sample (Liu and Phillips, 1991).

Figure 6 represents a schematic diagram of a GC × GC arrangement in a GC instrument: the instrumental arrangement is simplified with respect to the column coupling, with no valving or switching systems required (Marriott, Shellie, and Cornwell, 2001).

4.2.4 Gas Chromatography–Mass Spectrometry (GC-MS) and Gas Chromatography-Tandem Mass Spectrometry (GC-MS/MS)

The introduction of hyphenation of GC with MS (GC-MS) has been one of the most significant steps towards a reliable characterization of components of volatiles complex mixtures, which could not be completely separated even on high resolution capillary columns. Nowadays, direct coupling of the capillary columns to the ion source of the MS is the most commonly used method. In the GC-MS investigation of sesquiterpenes and monoterpenes, different types of platforms have been incorporated using, for example, quadrupole (Al-Harrasi and Al-Saidi, 2008), ion-trap (IT) (Piskorski *et al.*, 2007), high resolution magnetic sectors (Brown *et al.*, 2006) and time-of-flight (TOF) analyzers (Santos *et al.*, 2006).

In GC-MS, there are two commonly used ionization methods. The most frequent one is electron impact (EI) ionization, using electrons to ionize the compounds during their elution. Initially, the analyte molecule is bombarded with the electron beam to produce a molecular radical ion, which subsequently generates fragment ions. This relatively harsh mode of ionization yields to extended fragmentation and, therefore, allows determination of both relative molecular mass and structural information of the molecule. A very important feature of this ionization mode is its high reproducibility under standard set of conditions, facilitating the generation of spectral libraries, which can be used for the identification of unknown compounds. However, the intense fragmentation is sometimes a limitation in term of sensitivity, which is required for the analysis of very small amount of compounds. The other ionization technique also used in GC-MS analysis and presenting softer characteristics is chemical ionization. In this case, the analyte reacts with a reagent gas introduced in the MS at a low pressure, and as a result, molecular ion adducts are formed with the ionization
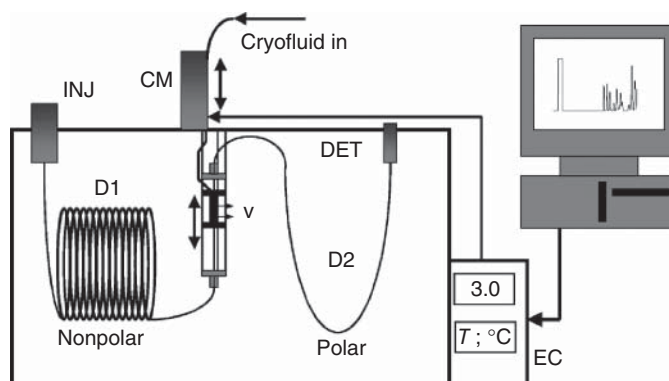
Figure 6 Schematic diagram of the GCXGC arrangement using a cryogenic modulator. *Abbreviations*: CM, cryomodulator; INJ, injector; DET, detector; D1, D2, separation dimensions 1 and 2, respectively; EC, electronic controller with selectable modulation duration (30 = 3.0 s) and cryotrap temperature shown; V, exit vents from the cryotrap. The modulator oscillates as shown by the arrow. Cryofluid CO_2 is provided to the trap via the movement arm. (*Source*: P.J. Marriott *et al.* (2001). Reproduced from Elsevier.)

gas, and less or usually no fragment ions are generated. Various reagent gases (e.g., methane, isobutane, and ammonia) could be used to yield valuable additional information, but ammonia seems to be the optimum gas for sesquiterpenes (Schultze, Lange, and Schmaus, 1992). After a case study of naturally occurring sesquiterpenes lactones in *Achillea* species (Asteraceae), Reznicek *et al*. recommended the combination of CI and EI-MS as a method of choice for the maximum structural information for the identification of sesquiterpene lactones by MS (Reznicek, Kastner, and Glasl, 1997).

The selection of a mass analyzer in GC-MS analysis defines the mass range, mass resolution, sensitivity, spectral collection speed, and cost of the instrument. Nowadays, the majority of GC-MS applications uses one-dimensional capillary GC with quadrupole MS analyzer and EI source. Some reasons of this preference are its ruggedness and reliability in addition to its good library compatibility as quadrupoles are the most commonly used analyzers to generate mass libraries (Mastovska and Lehotay, 2006). Quadrupole GC-MS operates in two scanning modes: full scan and selected ion monitoring (SIM). Specifically regarding the SIM mode, sensitivity is improved for the analysis of target compounds and for discrimination of overlapping GC peaks even though the spectral information is sacrificed. Firstly described in the 1950s, the IT mass spectrometer is a development of the quadrupole family, which operates by traping a range of ion masses detecting and measuring the ions while stored (Wolfgang and Helmut, 1960). This technique that is challenging the linear quadrupole position today is widely used in the research and routine laboratories worldwide. The GC-ITMS offers not only good sensitivity but also the capability to manipulate ions during storage while it is characterized by relatively high mass range and low cost. The very crucial advantage of this analyzer is its capacity of performing tandem MS and MS^n by means of dissociation such as collision-induced dissociation (CID), providing an extra degree of selectivity and an exceptional ability to avoid spectral interferences.

GC tandem MS has been used for the determination of sesquiterpenes in Chinese vetiver [*Chrysopogon zizanioides* (L.) Roberty, Poaceae] oil. A part of this study aimed to confirm the presence of three sesquiterpenes, α-cedrene, acoradiene, and khusimene, which are minor and isomeric compounds. Their EI mass spectra were mixed together in

the GC-MS analysis, and consequently, correct identification by comparison with EI spectra of the standards could not be ensured. This aim was achieved only by GC-EI-MS/MS experiments (Sellier *et al*., 1991).

One of the newer methods proposed to give improved analysis of complex mixtures is time-of-flight mass spectrometry (GC–TOF-MS). Being capable of generating high resolution mass spectra instantaneously, no bias arises from the mismatch between scan rate (duty cycle) and peak abundance changes in the ion source, which may arise with quadrupole mass spectrometers when used for fast GC peaks; so, one should expect uniform mass spectra across the whole peak. For instance, a comprehensive two-dimensional GC coupled to TOF-MS (GC×GC-TOFMS) was used to improve the chemical characterization of ylang-ylang [*Cananga odorata* (Lam.) Hook. f. & Thomson, Annonaceae] essential oil previously reported in 1DGC-MS analysis (Benini *et al.*, 2012) by employing mass selective detector. Based on the enhanced chromatographic separation and the mass spectral deconvolution, 161 individual compounds were identified as characteristic analytes of ylang-ylang flowers. Among these, 75 compounds essentially terpenes, terpenoid esters, and alcohols were reported for the first time in ylang-ylang essential oil during this study (Brokl *et al*., 2013).

4.3 High Performance Liquid Chromatography (HPLC)

HPLC is a widespread and versatile chromatographic technique that has been widely used for the analysis of natural products in complex matrices, such as crude extracts. HPLC coupled with a variety of detectors [ultra violet/visible (UV/vis), diode array detector (DAD), evaporate light scattering detector (ELSD), MS, MS/MS, and nuclear magnetic resonance (NMR)] has been applied for the detection, identification, and quantification of secondary metabolites in plant extracts or other matrices (e.g., biological fluids) and for the fingerprinting and profiling of crude extracts (Table 5) (see **HPLC and Ultra HPLC: Basic Concepts**).

In general, the potentials and the limitations of HPLC analysis are connected to those of detectors used. The major limitation of HPLC as chromatographic technique in sesquiterpenoids field is

Table 5　HPLC analysis of various sesquiterpenoids.

Plant material	Column	Mobile phase	Detection	Purpose	References
Sesquiterpenes					
Celastrus angulatus Maxim. (Celastraceae)	Hypersil Gold C18 (5 μm, 150 × 2.1 mm)	MeCN–H$_2$O (gradient)	DAD, 230 nm ESI(+)-MS/MS (ion trap)	Detection and identification	Wei *et al.* (2012)
Warburgia salutaris (+SLs) (G. Bertol.) Chiov. (Canellaceae)	Luna C18 (3 μm, 150 × 4.60 mm) at 40 °C	MeOH/H$_2$O 95:5 + 0.1% F.H. H$_2$O/MeOH 95:5, + 0.1% F.A. (gradient)	DAD, 230 nm; NMR, 600 MHz; 30 μL flow probe	Detection and isolation	Clarkson *et al.* (2007)
Curcuma rhizomes (Zingiberaceae)	Zorbax ODS (5 μm, 250 × 4.6 mm) at 25 °C	MeCN–H$_2$O (gradient)	DAD, 214/256 nm	Quantification and quality control	Yang, Wang, and Li (2006)
Sesquiterpene phytoalexins *Capsicum annuum* L. and *Nicotiana tabacum* L. (Solanaceae)	Spherisorb CN (5 μm, 100 × 3 mm)	Hexane–isopropanol (97:3)	UV/Vis, 205 nm	Quantification	Moreau, Preisig, and Osman (1992)
Petasites hybridus (L.) G. Gaertn. *et al.* (Asteraceae)	Nucleosil-100 (3 μm, 250 × 4.0 mm) at RT	Hexane–diisopropylether-acetonitrile (gradient)	DAD, 254 nm	Quantification	Debrunner, Neuenschwander, and Bremeisen (1995)
Ylang ylang oil *Solidago altissima* L. (Asteraceae)	Chiralcel OD (5 μm, 250 × 4.60 mm) at 24 °C	Hexane–isopropanol (40:1)	DAD, 255 nm	Detection and isolation	Nishii, Yoshida, and Tanabe (1997)
Daucane esters, *Ferula hermonis* Boiss. (Apiaceae)	Aqua C18 (5 μm, 150 × 4.6mm)	A: H$_2$O; B:MeCN–THF–isopropanol (80/80/20); A:B (48:52)	DAD, 240 nm	Detection and quantification	Abourashed *et al.* (2001)
Sesquiterpene lactones					
Saussurea lappa (Decne.) C. B. Clarke (Asteraceae)	TC-C18 (5 μm, 250 × 4.60 mm) at RT	MeOH–H$_2$O (70:30)	DAD, 225 nm	Quantification	Zhang, Cai, and Liu (2011)
Inula helenium L. (Asteraceae)	Zorbax XDB-C18 (5μm, 250 × 4.6mm) at 40 °C	MeCN–H$_2$O (55:45)	DAD, 210 nm	Quantification	Huo *et al.* (2010)
Saussurea costus (Falc.) Lipsch. (Asteraceae)	SunFire C18 (3.5 μm, 3 × 150 mm)	MeCN–H$_2$O (gradient)	DAD, 220 nm ESI(+)-MS (micro-TOF)	Detection and isolation	Julianti *et al.* (2011)
Achillea species (Asteraceae)	LiChrospher 100 RP 8 (5 μm, 250 × 4.0 mm)	MeCN–H$_2$O (gradient)	DAD, 220/255 nm	Quantification	Glasl *et al.* (1999)
Thapsia garganica L. (Apiaceae)	Luna C18 (3 μm, 150 × 4.60 mm) at 40 °C	MeOH/H$_2$O 95:5 +0.1% F.A. – H$_2$O/MeOH 95:5 +0.1% F.A. (gradient)	DAD, 230 nm; NMR, 500 MHz; CapNMR probe	Detection and isolation	Lambert *et al.* (2006)

Species (Family)	Column	Mobile phase	Detection	Purpose	Reference
Artemisia absinthium L. (Asteraceae)	Zorbax Eclipse XDB-C18	MeCN–H$_2$O 0.085% *o*-phosphoric acid (gradient)	DAD, 205 nm	Quantification	Aberham *et al.* (2010)
Viguiera radula Baker (Asteraceae)	Hypersil ODS (5 μm, 250 × 4.0 mm)	MeOH–H$_2$O (30 : 70/50 : 50)	DAD, 225/265 nm	Detection and isolation	Spring *et al.* (2003)
Lindera aggregata (Sims) Kosterm. (Lauraceae)	UPLC HSS T3 (1.8 μm, 2.1 × 100 mm)	MeCN–H$_2$O+0.1% F.A. (gradient)	UPLC-MS/MS	Quantification	Wu *et al.* (2010)
Eupatorium lindleyanum DC. (Asteraceae)	Zorbax Extend C18 (5 μm, 250 × 4.6 mm) at 30 °C	MeCN–H$_2$O (gradient)	DAD-ESI(+)-MS/MS	Detection and identification	Yang *et al.* (2010)
Vernonia fastigiata Oliv. & Hiern (Asteraceae)	RP-18 (5 μm, 250 × 4.6 mm)	MeCN–H$_2$O (gradient)	LC-MS LC-NMR	Detection and identification	Vogler *et al.* (1998)
Lactuca species (Asteraceae)	RP-18 (5 μm, 250 × 4.6 mm) 35 °C	H$_2$O:+0.1% H$_3$PO$_4$–MeCN/H$_2$O (90 : 10) (gradient)	DAD, 200 nm	Detection	Sessa *et al.* (2000)

the restriction for the analysis of volatile compounds. Therefore, mostly, nonvolatile sesquiterpenoids (esters, acids, sesquiterpene lactones, etc.) could be subjected to HPLC analysis. In particular, the method of choice for the analysis of sesquiterpene lactones as they are nonvolatile and thermolabile compounds is HPLC, mainly on RP. (Cimpan and Gocan, 2002; Merfort, 2002).

4.3.1 High Performance Liquid Chromatography–Ultraviolet/Visible Detector or Diode Array Detector (HPLC-UV/vis or HPLC-DAD)

The hyphenation of HPLC with UV/vis-DAD is broadly applied for natural products analysis. In particular, UV/vis and DAD combining the advantages of simplicity, sensitivity, versatility, and reliability have been widely applied for the detection, identification, and quantification of natural compounds. Especially, DAD detector provides the possibility of online collection of spectroscopic data, UV spectra, generation of UV libraries, and peak purity capabilities exploited for the characterization of the compounds contained in multiple matrices. Hence, hyphenation of HPLC with DAD has been proved a powerful tool for the identification of natural products in complex mixtures. The most important drawback of this technique is the impossibility of analyzing compounds that lack UV chromophores.

Generally, sesquiterpenoids lack UV chromophores, and their detection by UV or DAD detectors is usually carried out using low range wavelengths. In this case, several mobile phase solvents that exhibit high UV cutoffs should be avoided because they might blind the detection of these compounds. For example, using short wavelength at 192 nm, a method was developed and validated for the analysis of underivatized artemisinin and related sesquiterpene lactones in *A. annua* L. (Asteraceae) extracts and artemisinin-based drugs (Ferreira and Gonzalez, 2009). Sesquiterpene phytoalexins such as capsidiol, rishitin, lubimin, and debneyol produced by different species of Solanaceae family have been traditionally analyzed by GC or TLC. However, Moreau *et al*. reported the quatitative analysis of sesquiterpene phytoalexins accumulated by bell pepper (*Capsicum annuum* L.) fruit and tobacco (*Nicotiana tabacum* L.) cell suspensions after elicitation by cellulase using normal phase HPLC and

a cyanopropyl-bonded phase column coupled with UV detection at 205 nm. The UV chromatogram recorded as signal output was essentially a response from the number of carbon–carbon double bonds present in each sesquiterpene compound (Moreau, Preisig, and Osman, 1992).

However, less volatile sesquiterpenoids with strong UV chromophores are more suitable for HPLC-DAD analysis. For instance, a method for the detection of 14 sesquiterpene esters (petasin derivatives) in different parts of *Petasites hybridus* (L.) G.M. et Sch. (Asteraceae) has been developed on a normal phase Nucleosil 100 column using DAD, at 254 nm. Among them, six compounds were quantitatively determined in the different parts of the plant, demonstrating their abundance in all subterranean parts, with the exception of one compound that seems to be located mostly in leaves (Debrunner, Neuenschwander, and Brenneisen, 1995). Similarly, Abourashed *et al*. reported the development of an HPLC-DAD method for the detection and quantification of four daucane sesquiterpene esters of the sesquiterpene alcohol, ferutinol, in plant extracts and commercial samples from various species of the genus *Ferula* (Umbelliferae). The presence of an aromatic ring on the structure of these compounds enabled the UV monitoring at 240 nm (Abourashed *et al*., 2001).

Sesquiterpenes are often optically active; hence, the use of chiral columns in some cases could result in enantioselective separations. For instance, racemic mixture of 1-hydroxy-derivative of germacrene D was analyzed using chiracel OB and OD columns, but no separation was obtained; however, after its transformation to 9-hydroxy-derivative, it was analyzed under the same conditions, and clear enantiomeric separation was achieved through chiracel OD using hexane:isopropanol 40:1 as a mobile phase. The aforementioned method was used for the determination of the optical purity of germacrene D in ylang ylang oil (*Cananga odorata* (Lam.) Hook. f. & Thomson, Annonaceae) and in stems and leaves of *Solidago altissima* L. (Asteraceae) and for the isolation of both enantiomers (Nishii, Yoshida, and Tanabe, 1997).

Furthermore, HPLC is a valuable tool for the analysis of thermolabile sesquiterpenoids, where GC cannot be used. An application for the quantification of sesquiterpenoids, among them two thermolabile compounds, was described by Yang *et al*. during the development of an HPLC-DAD method for the simultaneous determination of sesquiterpenoids in

Curcuma rhizomes (Zingiberaceae) using a Zorbax ODS column at 214 and 256 nm. Although GC-MS methods had been already developed for the quantification of sesquiterpenoids from *Curcuma* species, quantification of the two thermolabile compounds resulted in false results because of their degradation. This validated method was successfully applied to quantify 11 investigated components in 18 samples of three *Curcuma* species (Yang, Wang, and Li, 2006). Similarly, a RP-HPLC-DAD method (RP 8 LiChrospher 100) was developed in order to identify and quantify antiphlogistic sesquiterpenoids such as some thermolabile proazulenes in different *Achillea* species (Asteraceae) (Glasl *et al*., 1999).

Regarding sesquiterpene lactones, despite their physicochemical properties that render these analytes suitable for LC analysis, the presence of chromophores on their structures facilitates significantly their UV or DAD detection. Through the years, many methods have been developed for the detection and quantification of sesquiterpene lactones in complex mixtures with HPLC-UV or DAD principally at short wavelengths using RP chromatography. Spring *et al*. reported the detection of 13 sesquiterpene lactones from glandular trichomes of *Viguiera radula* Baker (Asteraceae) using two isocratic HPLC-DAD methods (Hypersil ODS column) and observing chromatograms recorded at 225 and 265 nm. The extraction of the UV spectra of the observed peaks revealed the presence of sesquiterpene lactones in the glandular trichomes extracts under study (Spring *et al*., 2003).

Another approach for the detection of sesquiterpenoids with weak UV chromophores using HPLC-DAD methodologies includes the pre- or postcolumn chemical derivatization of the target compounds resulting in their transformation into UV-active ones. This approach has been used, for example, for the quantification of artemisinin in different matrices. In this case, a pre- or postcolumn hydrolysis of artemisinin occurs in order to obtain more UV-active compounds, which allows the use of HPLC-UV/vis or DAD apparatus (ElSohly, Croom, and ElSohly, 1987; Qian, Yang, and Ren, 2005). Similarly, derivatization by 9-tryomethylanthracene of sesquiterpene lactones from *Tanacetum parthenium* L. (Asteraceae) containing α-methylenebutyrolactone functions resulted in the transformation of these lactones into compounds with strong chromophores that enabled their

detection in leaf extracts at 369 nm (Dolman *et al*., 1992).

4.3.2 High Performance Liquid Chromatography – Retention Index (RI) and Evaporate Light Scattering Detector (ELSD)

Since the majority of sesquiterpenoids lack chromophores, other detectors compatible with HPLC have been applied for their detection in complex mixtures. The most simple and least expensive of these detectors is the refractive index (RI) detector. Nevertheless, RI presents many drawbacks such as low sensitivity and instability when changes in temperature, pressure, and flow rate occur; hence, its use under gradient elution conditions is not recommended. Some scarce applications of RI detection in sesquiterpenoids field have been reported for the quantification of artemisinin in *A. annua* L. (Asteraceae) extracts using RP chromatographic column and mobile phase consisting of acetonitrile-water in isocratic elution method (60 : 40, v/v). This method has been validated and showed a limit of quantification of 0.1 mg/mL (Lapkin *et al*., 2009). Similarly, the determination of bilobalide together with ginkgolides, the active constituents of *Ginkgo biloba* L. (Ginkgoaceae) leaves and phytopharmaceuticals was performed using RP chromatography and mobile phase consisting of methanol–water (33 : 67, v/v) (Van Beek *et al*., 1991). The aforementioned methods, because of the unstable baseline problems and the low sensitivity, are scarcely used, and methods based on the ELSD detection are preferred.

Recently, ELSD systems have gained popularity for the analysis of compounds possessing poor chromophores and being less volatile than the mobile phase. Detection involves the nebulization of the LC effluent, followed by the evaporation of the more volatile solvents, and then measurement of the light scattering caused by the analytes. The light scattering response is determined by the nature and the size of the analytes in addition to their concentration. ELSD provides not only some advantages toward RI such as compatibility with gradient elution methods but also some disadvantages toward MS detectors in terms of sensitivity (Megoulas and Koupparis, 2005). Because artemisinin constitute a powerful antimalarian agent, many groups have been working on the development of chromatographic methods, such as HPLC-ELSD,

for the quantitative determination of artemisinin and other sesquiterpene derivatives of *Artemisia* species (Liu, Zhou, and Zhao, 2007; Peng, Ferreira, and Wood, 2006). The most sensitive HPLC-ELSD method for the quantitative determination of six artemisinin derivatives was proposed by Avery *et al*. using RP chromatography with isocratic elution method consisting of methanol–water (80 : 20, v/v) at 32 °C, gas pressure adjusted at 2.2 bar and signal gain 12 (Avery, Venkatesh, and Avery, 1999). In general, the target of each analysis defines the suitability of different methods. Referring to a comparative study concerning the development of analytical protocols for the determination of artemisinin and its analogs in different matrices, the HPLC-ELSD method proved to be the most robust for routine quantification of artemisinin in plant extracts and useful for optimization of extraction/purification protocols. Nevertheless, according to this study, the HPLC-UV method (detection at 210 nm) is recommended for the analysis of the purity of bulk artemisinin. In both cases, detection of some impurities in artemisinin mixtures was not possible due to the limited sensitivity of UV and ELSD detectors (Lapkin *et al*., 2009). In fact, it is demonstrated that MS detectors enable the quantitative determination of artemisinin analogs in significantly smaller quantities. For instance, Avula *et al*. report the development of a sensitive LC method using a TOF analyzer MS detector equipped with elecrospray ionization (ESI) source, in positive mode, for the identification and quantification of artemisinin derivatives, in aerial parts of different species of *Artemisia* (Avula *et al*., 2009). Similarly, the sesquiterpene bilobalide together with the diterpene trilactones and ginkgolides were quantitatively determined in *G. biloba* L. (Ginkgoaceae) complex matrices such as extracts and commercial products. Exclusively, RP chromatography was applied using mostly gradient (Camponovo *et al*., 1995; Dubber and Kanfer, 2006; Li and Fitzloff, 2002) and scarcely isocratic (Tang, Wei, and Yin, 2003) elution methods and ELSD as a detector. As an example, Dubber and Kanfer developed an HPLC-ELSD method for the quantitative analysis of the above-mentioned compounds in *G. biloba* solid oral dosage forms. Using a simple gradient method with mobile phase consisting of methanol and water, separation of the selected marker compounds was achieved within 14 min. After the optimization of ELSD conditions, the method was validated and used to assay commercially available *G. biloba* products and proved to be

suitable for routine analysis of such products and quality control purposes (Dubber and Kanfer, 2006).

4.3.3 Liquid Chromatography–Mass Spectrometry (LC-MS)

Different mass analyzers have been used as HPLC detectors for the screening of plant extracts in order to detect and identify sesquiterpenoids and for their quantification in different matrices. The use of LC-MS techniques provides many advantages comparing to the aforementioned LC detectors in terms of sensitivity and selectivity. Furthermore, MS detection offers online important structural information such as molecular weight and diagnostic fragmentation patterns, which are crucial for the structural characterization of different compounds in complex mixtures and dereplication strategies. Dereplication refers to a procedure aiming for the identification of known compounds in complicated mixtures such as plant extracts early in the discovery process based on already known analytical methods and databases. Thus, the coupling of the HPLC with MS instruments and the development of atmospheric pressure ionization (API) sources has led to the introduction of a new powerful approach for the screening and characterization of plant extracts (see **LC and LC-MS: Techniques and Applications**). Regarding small terpenoids, germacrane sesquiterpene lactones were detected and identified in the aerial parts of *Eupatorium lindleyanum* DC. (Asteraceae) using HPLC-DAD-MS techniques; however, the analysis of five reference compounds before the plant extract analysis was necessary in order to compare the spectroscopic data with the reference ones. For the ionization of the germacrane sesquiterpene lactones, an ESI source in positive mode was applied and the identification thereof was mainly based on the detected $[M+Na]^+$ adduct ions. In addition, the study of the reference compounds fragmentation patterns in a micromass Q/TOF mass spectrometer was crucial for the identification of unknown derivatives in the crude extract of *E. lindleyanum* (Yang *et al*., 2010). Similarly, after the extensive study of 29 celangulin sesquiterpenoids and the development of an HPLC-DAD-ESI-MS/MS method for the rapid characterization of celangulins, 36 celangulin derivatives were identified in crude extract of root barks of *Celastrus angulatus* Max.

(Celastraceae). The ESI/MS/MS analysis in positive mode of the celangulins using an IT analyzer enabled the generation of diagnostic fragment ions, which led to the rapid detection of low abundance celangulins in the crude plant extracts (Wei *et al*., 2012).

The LC-MS profiling of plant extracts is a key point for dereplication strategies combined ideally with other hyphenated techniques. In the context of an HPLC-based activity profiling approach for the identification of antitrypanosomal compounds, six sesquiterpene lactones from *Saussurea costus* (Falc.) Lipschitz (Asteraceae), *Laurus nobilis* L. (Myrtaceae), and *Eupatorium cannabinum* L. (Asteraceae) were identified combining the information extracted from their HPLC-DAD and HPLC-ESI(+)-TOF-MS profiles and from the NMR spectra of the targeted compounds. HPLC-MS and HPLC-HRMS profiles of all extracts were performed using ESI ion source, in both positive and negative modes. Interestingly, sesquiterpene lactones from *S. costus* ionized under these conditions were identified based on the $[2M+H]^+$ ions observed in the HRMS spectra, whereas characteristic $[M+Na]^+$ ions corresponding to sesquiterpene lactones were observed in the rest species (Julianti *et al*., 2011). Similarly, information extracted from the HPLC-DAD, LC-ESI-MS, and LC-ESI(+)-Q-TOF profiles of *Ferula communis* L. (Apiaceae) and *Dittrichia viscosa* (L.) W. Greuter (Asteraceae) extracts was important for the determination of the chemical composition of these extracts. In fact, two sesquiterpene lactones (inuviscolide and tomentosin) and three sesquiterpene acids (costic acid, hydroxycostic acid, ilicic acid) were identified from the *D. viscosa* extract, whereas in *F. communis* extracts, three daucane sesquiterpenes (acetoxyferutinin, oxojaeskeanadioyl anisate, and fertidin) were characterized (Mamoci *et al*., 2011).

Since HPLC-MS techniques are characterized by high sensitivity and selectivity, many methods based on this technique have been developed for the quantitative determination of sesquiterpenoids in various matrices. A rapid validated method was proposed by Wu *et al*. in order to evaluate the quality of *Lindera aggregata* (Sims) Kosterm. (Lauraceae) extracts through the simultaneous quantification of five sesquiterpene lactones. A UHPLC-ESI(+)-MS/MS method of 8 min was developed using multiple reaction monitoring (MRM) scan mode in a triple quadruple analyzer and was characterized by good sensitivity. This method was successfully applied for the quantitative determination of *L. aggregata*

samples from different areas of China (Wu *et al*., 2010).

The quantification of artemisinin and its derivatives, as mentioned above, have been studied broadly, using a variety of analytical methods, because of its therapeutic importance (Lapkin *et al*., 2009). However, the quantitation of artemisinin with LC-MS techniques is characterized by higher sensitivity and selectivity (Ivanescu *et al*., 2011; Liu *et al*., 2010; Reale *et al*., 2008; Van Nieuwerburgh *et al*., 2006; Wang *et al*., 2005). Wang *et al*. proposed a validated LC-MS method developed on a quadrupole mass analyzer using ESI ion source in positive mode and SIM scan mode recording the abundance of the $[M-18+H]^+$ ion peak of artemisinin in order to quantify it in different populations of *A. annua* L. (Asteraceae) (Wang *et al*., 2005). A more sensitive LC-ESI(+)-MS/MS method in MRM mode was proposed by Van Nieuwerburgh *et al*. for the quantification of four artemisinin analogs in *A. annua* using a Q-TOF mass analyzer. Detection was achieved by the monitoring of specific MS/MS transitions of these sesquiterpenoids from their precursor $[M+H]^+$ or $[M+NH_4]^+$ ions (Van Nieuwerburgh *et al*., 2006).

The distinction between helenalins and dihydrohelenalin esters, traditionally, is performed by UV detection at 225 nm according to the European Pharmacopoeia. An alternative approach for the differentiation of these compounds in different chemotypes of *Arnica montana* L. (Asteraceae) was proposed by Perry *et al*. A selective LC-ESI(+)-MS/MS method was developed to clearly separate helenalin from dihydrohelenalin esters, including the isomeric ones, taking advantage of the characteristic daughter ions at *m/z* 245 for helenalins and *m/z* 247 for dihydrohelenalins and led to the identification of minor compounds in the crude extracts. Furthermore, the application of this method enabled a more precise quantification (comparing to that of the European Pharmacopoeia method) of the different helenalin and dihydrohelenalin esters in the various chemotypes of *A. montana* (Perry *et al*., 2009).

Most of the sesquiterpenoids are successfully ionized using ESI ion source in positive mode, and mainly, $[M+Na]^+$ adduct ions or $[M+H]^+$ ions are observed. In rare cases, ESI has been reported unable to generate mass spectra of sesquiterpenoids, whereas atmospheric pressure chemical ionization (APCI) was proved to be more efficient for the ionization of these compounds because of their nonpolar nature (Montsko *et al*., 2008; Spring *et al*.,

2001). Moreover, in APCI positive mode, the presence of a volatile acid in the mobile phase seems to result in a higher stability of the protonated ions and an improved signal-to-noise ratio in the MS spectra. For instance, using a RP nonporous column for more rapid separation, the ionization of a santonin standard solution (0.1 µg/mL) was optimized, concluding that concentrations of 1 mM of ammonium acetate in the mobile phase ameliorated significantly the stability and the abundance of the recorded signal (Montsko *et al*., 2008). APCI in negative mode has been also reported for the analysis of sesquiterpene compounds when functional groups present on the molecule enable their ionization in this mode. For example, the quatitative determination of bilobalide and ginkgolides in *G. biloba* extracts and pharmaceutical preparation was performed by negative APCI on a triple quadrupole mass spectrometer using a 14 min method. All the compounds under study were selectively detected by SIM mode of their deprotonated molecules. Thus, this method was proposed for the rapid and effective quantification of the above-mentioned compounds (Jensen *et al*., 2002).

4.3.4 Liquid Chromatography–Nuclear Magnetic Resonance (LC-NMR)

The hyphenation of NMR spectrometers online with HPLC systems represents an additional valuable tool for dereplication processes and rapid structure elucidation of natural products. Providing the advantage of online separation with structural information, avoiding laborious and time-consuming isolation and purification steps, LC-NMR technique facilitates high throughput screening approaches (Wolfender, Ndjoko, and Hostettmann, 2001). Another advantage of LC-NMR includes the lower probability of degradation of sensitive products and introduction of impurities during the classical isolation procedure. For instance, LC-NMR together with other hyphenated techniques may ideally allow the total characterization of plant extracts. However, online LC-NMR measurements suffer from low sensitivity compared to other techniques such as LC-MS, which results in decreased use of the technique. In order to overcome this limitation, three major concepts have been introduced to the development of hyphenated NMR techniques. The first approach was the introduction of new solvent suppression techniques, which indeed

contributed to the suppression of solvent signals with minimal baseline distortions. One important disadvantage of this approach is the possible suppression of signals from the analyte nearby the solvent signal, which would result in loss of structural information. Secondly, the developments in NMR probes design (cryogenic probes, cap-NMR, and microflow NMR) led to a significant increase of the sensitivity. Finally, the use of SPE interface between the LC apparatus and the NMR spectrometer has gained high popularity lately in the natural products field. This methodology enables the targeted trapping of the analyte at the SPE cartridge, the evaporation of the HPLC solvents, following the resolubilization in deuterated solvents and spectra recording (Exarchou *et al*., 2005; Silva Elipe, 2003) (see **On-line and At-line LC-NMR and Related Micro NMR Methods**).

LC-NMR and HPLC-SPE-NMR applications have been reported for the identification of sesquiterpenoids and mainly sesquiterpene lactones in plant extracts (Spring *et al*., 1995, Lambert *et al*., 2006). In most of the cases, multiple injections before NMR analysis are required for the accumulation of larger quantity in the SPE cartridge, which results in a higher sensitivity of the measurements. However, according to Miliauskas *et al*., the use of semipreparative columns is proposed for the analysis of nonpolar and medium polarity compounds, among them sesquiterpenoids providing better results in terms of peak resolution and, therefore, NMR spectra (Miliauskas *et al*., 2006). The active compounds of *Warburgia salutaris* (Bertol. f.) Chiov. (Canellaceae) were identified through HPLC-SPE-NMR analysis of the active fraction of the extract, resulting in the characterization of 11 drimane and coloratane sesquiterpenoids. Separation was carried out using a RP analytical column injecting 1 mg six times successively while trapping of the target peaks was based on the UV chromatogram using HyShere resin GP cartridges. Clarkson *et al*. have also highlighted the importance of the HPLC-SPE-NMR method for sesquiterpene lactones as they are considered relatively unstable substances (Clarkson *et al*., 2007). In this case, during the development of a validated method for the quatitative determination of five sesquiterpene lactones together with two lignans and a flavonoid from *A. absinthium* L. (Asteraceae), the degradation of a sesquiterpene lactone, absinthin, was observed. The three degradation compounds (anabsin, anabsinthin, and the new dimer

30-hydroxyanabsinthin) were identified and characterized by HPLC-MS and HPLC-SPE-NMR and quantified by an established HPLC-DAD method (Aberham *et al.*, 2010).

4.4 Supercritical Fluid Chromatography (SFC)

SFC using a supercritical fluid, most commonly CO_2, as a mobile phase in combination with one or more polar organic solvents constitutes a promising chromatographic technique for the analysis of sesquiterpene derivatives (Merfort, 2002). SFC is suitable for the analysis of thermally labile compounds such as sesquiterpene hydrocarbons because degradation may occur during their analysis by GC operating at high temperatures, whereas HPLC analysis is not preferable for these compounds due to their volatility and nonpolar nature (Morin *et al.*, 1991). For example, SFC analysis of thermosensitive sesquiterpene hydrocarbons and alcohols was performed at 40 °C using bare silica stationary phase. Sesquiterpene hydrocarbons were separated and identified from several essential oils by the developed SFC-Fourier transform infrared (FTIR) method with pure CO_2 as mobile phase (Morin *et al.*, 1991). Sesquiterpene alcohols, among them the geometrical isomers geranyllinalool and nerolidol, were successfully separated in less than 10 min using CO_2 modified with 0.5% of methanol as mobile phase in order to decrease the retention of these compounds on the stationary phase. Identification of studied compounds was performed through UV spectra obtained during the SFC-DAD analysis (Morin, 1994).

Regarding instrumentation, capillary or packed columns have been used as stationary phases in SFC (Taylor, 2009). Analysis of sesquiterpenes with capillary columns is scarcely reported. Köhler *et al.* compared the efficiency of two SFC methods for the determination and quantification of artemisinin and artemisinic acid in *A. annua* L. (Asteraceae) extracts. Both capillary SFC-FID and packed column SFC-ELSD methods were optimized using different stationary phases; the mobile phase in the case of capillary SFC-FID consisted of pure CO_2, while a small percentage of methanol was added to CO_2 when packed column SFC-ELSD method was applied. The only difference observed between these SFC methods was the significantly shorter time analysis of the packed column SFC-ELSD method (Köhler *et al.*, 1997a). In order to determine the differences of four different hop (*Humulus lupulus* L., Cannabaceae) variety essential oils, an open-tubular SFC-FTIR analysis was performed using a SB-Phenyl-5 column and Sc-CO_2. The major constituents (*trans*-caryophyllene, α-humulene, and myrcene) were identified comparing the IR spectra taken as films deposited on AgCl discs and the spectra obtained after chromatographic separation in the flow cell (Auerbach, Kenan, and Davidson, 2000).

Although SFC has been coupled with a variety of detectors, traditionally hyphenated to HPLC or GC (Chester and Pinkston, 2004), the detection of sesquiterpenes using SFC has been mainly carried out by UV or DAD detectors. The introduction of pressure-resistant UV flow cells and a back pressure regulator enabled the coupling of SFC with DAD detectors (Berger and Berger, 2011; Taylor, 2010). For instance, an extensive study for the analysis of different skeleton sesquiterpene lactones (germacranolides, guaianolides, eudesmanolides, seco-eudesmanolides, and helenanolides) from *Carduus benedictus* L. and *Artemisia umbelliformis* Lam. (Asteraceae) was performed by SFC-DAD at 254 nm. In order to optimize the conditions, diverse stationary phases and modifiers were applied, and finally, a polar stationary phase (CN) and CO_2 with methanol and water as mobile phase was found to be the optimum combination for the analysis of these compounds (Bicchi, Balbo, and Rubiolo, 1997a). Under the same conditions, after optimization of fundamental parameters (stationary and mobile phase), the quantification of valerenic acids (sesquiterpene derivatives) together with valepotriates derivatives (diterpenes) in *Valeriana officinalis* L. (Valerianaceae) extracts was obtained by SFC-UV (225 nm), and the results were compared with those obtained by the HPLC-UV method. In terms of efficiency, both the methods were comparable, whereas SFC-UV method was significantly shorter (Bicchi, Binello, and Rubiolo, 2000). Similarly, the determination of parthenolide in feverfew samples after ultrasonication has been performed by SFC (Guo, 2006). Alternative to UV detection, when UV inactive compounds are to be analyzed, ELSD has been applied. Moreover, ELSD is considered more compatible with SFC techniques than with HPLC as a result of the evaporating ease of the mobile phase (Sc-CO_2) (Chester and Pinkston, 2004). For instance, the analysis of artemisinin and its precursor

artemisinic acid have been successively performed on capillary and packed columns using an ELSD detector. The analysis on a aminopropyl silica column was significantly faster than the one on a capillary column (Köhler *et al.*, 1997b). Furthermore, the same group continued with the quatitative determination of these two compounds in SFE extracts of *A. annua* L. (Asteraceae) using the SFC-ELSD method with the aminopropyl silica column (Köhler *et al.*, 1997a).

Finally, the interface of SFC with MS instruments is of great interest; the application of packed columns SFC contributed to overcome the problems of open-tubular SFC-MS couplings. For instance, a SB-Methyl-100 fused-silica capillary column has been applied for the analysis of thyme essential oil and commercial extracts using a new injection technique with a retention gap (Blum, Kubeczka, and Becker, 1997). Specifically, coupling of SFC with APCI source has been gaining popularity during the past years as APCI and SFC are compatible with relatively high flow rates and are better suited techniques for the analysis of low to moderate polarity compounds (Chen, 2009). For instance, the analysis of artemisinin in *A. annua* extracts was carried out by a SFC-APCI-MS method using a packed column polar stationary phase (CN) and CO_2 modified by 10% methanol. The qualitative and quantitative determinations of artemisinin in the extracts were performed using full scan and SIM mode, respectively, in positive ionization mode (Dost and Davidson, 2003).

5 ISOLATION TECHNIQUES

5.1 Classical Techniques

The majority of the reported small terpenoids so far have been isolated by traditional techniques such as column chromatography (CC), medium pressure liquid chromatography (MPLC), HPLC, and preparative thin layer chromatography (pr-TLC). CC, as a routine tool, has been applied for the separation and purification of such compounds using mostly silica gel as a stationary phase and dextran gel (sephadex) (Bhat, Nagasampagi, and Sivakumar, 2005b), whereas the utilization of absorbent resins has been scarcely reported (Lai, Chen, and Tsai, 2005). In addition, the impregnation of silica with silver nitrate has been also used for small terpenoids

separation purposes and found to enhance the resolution of unsaturated terpene derivatives. Williams and Mander in a detailed review presented a number of publications showing that terpenoids can be easily separated using this technology (Williams and Mander, 2001). Because small terpenoids are generally nonpolar compounds, their elution is usually performed using nonpolar to medium polar mobile phase systems with gradually increasing the polarity. Usually, further purification of isolated compounds is achieved by means of pr-TLC. For example, five aromadendrane derivatives were isolated from the CH_2Cl_2 leaves extract of *Xylopia brasiliensis* Sprengel (Annonaceae) after repeated chromatographic steps using silica gel and Sephadex LH-20 (Moreira *et al.*, 2003). Similarly, the isolation of guaianolide, elemanolide, and lindenanolide sesquiterpene lactones from the aerial parts of *Hedyosmum brasiliense* Miq. (Chloranthaceae) was performed using successive silica gel-based column chromatography steps, eluting with mixtures of *n*-hexane, CH_2Cl_2, EtOAc, and acetone (Amoah *et al.*, 2013). In addition, the isolation of 11 sesquiterpene lactones from the methanolic extract of *Lactuca sativa* L var. *anagustata* (Asteraceae) was achieved by repeated silica gel columns and pr-TLC (Han *et al.*, 2010).

An alternative method for the isolation of essential oils constituents or extracts containing small terpenoids includes the prefractionation of the sample by dry column chromatography in order to obtain separately the hydrocarbon compounds from more polar ones such as alcohols, ketones, and other oxygenated compounds (Hostettmann, Marston, and Hostettmann, 1998). In order to isolate separately the components from the hydrocarbon fraction, preparative GC is performed using a nonpolar column, and subsequent further purification is achieved by packed column GC with the most suitable cyclodextrin phase, which is selected after the analysis of the fractions on corresponding capillary columns (König, Bülow, and Saritas, 1999). For instance, prefractionation of the raw essential oil by dry column chromatography with hexane as eluent before preparative GC using packed columns with different modified cyclodextrins led to the isolation of two new sesquiterpene hydrocarbons, petasitene and pethybrene from the rhizomes of *Petasites hybridus* (L.) G.Gaertn., B.Mey. & Scherb. (Asteraceae). Regarding pethybrene, an additional purification step on pr-TLC with $AgNO_3$ precoated plate was required (Saritas, von Reu, and König, 2002). In several cases,

dry column chromatography is performed on deactivated silica gel with water (Hernandez *et al*., 1980; Perales *et al*., 1983).

Moreover, mainly RP, preparative MPLC, and HPLC have been extensively used for the fractionation of essential oils or extracts containing terpene compounds and for the isolation of these compounds. Generally, multiple HPLC separations before GC analysis enable the substantial enrichment of the components in terms of their functional group type and polarity, in addition to the detection of even minor terpenoids that are normally overlapped (Merfort, 2002). For instance, semipreparative-HPLC on octadecyl-bonded silica using gradient elution method consisting of acetonitrile and water was successfully performed in order to separate initially essential oils into fractions of oxygenated monoterpenes, monoterpene hydrocarbons, sesquiterpene hydrocarbons, and oxygenated sesquiterpenoids (Morin *et al*., 1986). Another example concerns the methanolic extract of *Salvia divinorum* Epling & Jativa (Lamiaceae) and involves an initial fractionation step performed by normal phase column chromatography and further fractionation and isolation of fifteen neoclerodane diterpenes, which was carried out by RP-MPLC and HPLC (Shirota, Nagamatsu, and Sekita, 2006).

5.2 Countercurrent Chromatography (CCC)

Discovered by Dr. Yoichiro Ito (Conway, 2011; Ito, 2005a), the CCC (countercurrent chromatography) technique is a liquid–liquid partition, which possesses the uniqueness of using no-solid stationary phase, and separation is achieved between two immiscible liquid phases. Centrifugal or gravitational forces are used to maintain one liquid phase in a coil or train of chambers (stationary), whereas a stream of a second, immiscible phase passes through the system in contact with the stationary liquid phase (Conway, 1991). Therefore, using CCC provides several key advantages. First, the absence of solid adsorbents eliminates the adsorption losses and the formation of artifacts caused by active surfaces. Added to this, CCC techniques rely exclusively on less expensive solvent mixtures comparing to solid packing materials, which, in many cases, are very costly. Furthermore, only small amounts of solvents are required, and large sample loads of a wide range of polarities

can be applied, rendering the process scalable so that sufficient quantities of individual components can be obtained. Finally, in contrast to other chromatographic techniques, a total recovery (up to 100%) of the sample material is guaranteed (Sutherland and Fisher, 2009). Before discussing the application of CCC in small terpenoids and sesquiterpenoids, a short introduction regarding the principle of the technique and some technical aspects are presented.

Nowadays, the development of the modern CCC is split in two main directions: the hydrostatic CCC, often called *high performance centrifugal partition chromatography* (*HPCPC*); and the hydrodynamic CCC, also known as *high speed countercurrent chromatography* (*HSCCC*). The hydrostatic column is composed of partition channels, linked in cascade by interconnecting ducts and in which the two liquid phases are contacted. They are contained in a rotor whose rotation generates a constant gravitational field, *G*, which pushes the denser phase at the bottom of the channel. The hydrostatic CCC columns possess a very good retention of the liquid stationary phase in the channels, but only the connecting ducts contain the mobile phase. Because there is only one phase, the chromatographic mobile–stationary phase exchanges are not possible: the connecting ducts can be considered as dead volumes (Berthod, 2007; Ito, 2005b). Hydrodynamic CCC columns, exclusively developed by Ito, are composed of open tubes coiled on one or several spools, which are enclosed in a rotor (Ito, 2005a, b). In the very successful type-*J* synchronous planetary motion multilayer coil separation column, a planetary gear mounted on the column holder axis is engaged to a stationary sun gear rigidly fixed to the centrifuge framework producing the planetary motion. Consequently, the column holder rotates around its own axis and, at the meantime, revolves around the centrifuge axis at the same angular velocity (synchronous) in the same direction. Two major functions for performing CCC separation are provided by this particular type of planetary motion of the column: (i) a rotary seal-free elution system so that the mobile phase is continuously eluted through the rotating separation column and (ii) the production of a unique hydrodynamic motion of two solvent phases within the rotating multilayer coiled column mainly owing to the Archimedean screw effect. This force tends to push the lighter liquid phase toward one end of the coil called *the head* and the other *the tail*. The hydrodynamic equilibrium is created only if the mobile phase enters

through the right side of the coil. If the mobile phase is the lower denser liquid, it should enter in the coil through the head side, against the Archimedean force, in the head-to-tail direction. If the mobile phase is the upper lighter liquid, it should enter the CCC column through its tail in the tail-to-head direction. Otherwise, the CCC column will not retain any liquid stationary phase, thus precluding separation (Berthod, 2007; Ito, 2005b).

Generally, it is not possible to recommend one type of CCC columns or another. It depends on the separation or purification demands. If the liquid system is difficult to retain (e.g., aqueous two-phase liquid systems), a hydrostatic CCC column would be recommended. If high efficiency is needed, a hydrodynamic CCC column may be preferred as it provides more theoretical plates.

The two-phase solvent system selection for the separation of target compounds from a complex mixture is the most critical step in CCC, where searching for a suitable two-phase solvent system may be estimated as 90% of the entire isolation work (Ito, 2005a). The CCC user is able to choose solvents from an enormous number of possible combinations, and without consulting the previous literature involving the separation of similar compounds, this task can be very time consuming. A suitable two-phase solvent system has to fulfill the following conditions: (i) the analytes should be soluble and stable in the system; (ii) the solvent system should form two phases with nearly equal volume ratios to avoid excessive waste of solvent; (iii) the solvent system should provide a suitable partition coefficient to the analytes; and (iv) the solvent system should yield satisfactory retention of the stationary phase in the column. The most common procedure for the selection of the optimum solvent system is the test of already described systems in the literature (Foucault and Chevolot, 1998; Marston and Hostettmann, 2006; Yoon, Chin, and Kim, 2010). It is strongly advised to face the challenge of inventing new solvent system conditions only when none of the ones in the specialist literature is successful.

Very critical for a suitable solvent system selection is the partition coefficient (K), which is defined as the ratio of solute distributed between the mutually equilibrated two solvent phases. It is usually expressed as the concentration of solute in the stationary phase divided by that of the mobile phase. The suitable K values for CCC are $0.5 \leq K \leq 1.0$. In CCC, either the upper or the lower phase can be chosen as the stationary phase (depending on head-to-tail vs tail-to-head

elution as discussed earlier). Before deciding which phase is to be used as the stationary phase, the user, may temporarily express the partition coefficient as $K_{U/L} = C_U/C_L$, where C_U is the solute concentration in the upper phase and C_L that of the lower phase. If, for example, the $K_{U/L} = 2$, the lower phase should be used as the stationary phase, which gives $K = 0.5$.

In 1991, Oka et al. introduced two series of two-phase solvent systems that facilitate the systematic search for the solvent systems suitable for CCC. The n-hexane–ethyl acetate–n-butanol–methanol–water systems provided a broad range of hydrophobicity, whereas the chloroform–methanol–water systems were extremely useful for separations of various natural products with moderate hydrophobicity. In the case of the hexane-containing system, n-butanol was added only if the sample was more hydrophilic. In total, Oka et al. described 27 two-phase solvent systems, giving a broad polarity range. Most of the solvent systems provided close to $1:1$ volume ratios of the two phases and reasonable settling times. The exploration of a suitable solvent system may be, therefore, initiated with n-hexane–ethyl acetate–methanol–water $(1:1:1:1)$ or chloroform–methanol–water $(10:3:7)$. If the K values of the compounds of interest are too high, the search should be directed towards the more hydrophobic solvent systems in the hexane series, and if the K values are too low, the search should be directed towards the more hydrophilic solvent systems until the proper K values are obtained (Oka, Oka, and Ito, 1991). However, if K is 1, the analyte will elute at the retention volume equal to the column capacity regardless of the retention volume of the stationary phase. At $K < 1$ or $K > 1$, the analyte would elute before or after the elution of one column capacity volume, respectively. The retention volume (V_R) of the analyte may be predicted more accurately using the elution volume of the solvent front (V_{SF}) and the total column capacity (V_C) using the following equation: $V_R = V_{SF} + K(V_C - V_{SF})$.

Another model, similar to the approach of Oka et al., has been adopted by Margraff, starting with a quaternary mixture of n-heptane–ethyl acetate–methanol–water $(1:1:1:1)$ and ending with the binary ethyl acetate–water $(1:1)$ for polar compounds or n-heptane–methanol $(1:1)$ for less polar samples. Among these lines, Conway has developed specific diagrams, namely ternary diagrams, which constitute efficient means to search

for CCC solvent systems. Ternary diagrams suitable for CCC applications have been collected together (Foucault, 1994). These give the exact composition of both the mobile and the stationary phases and allow the desired quantity of each phase to be prepared independently. The solubility curve separates the miscible region from the two-phase region.

A variety of examples are available in the literature regarding the isolation and purification of sesquiterpenes and other small terpenoids using CCC (Table 6). For instance, rupestonic acid, which is the "marker compound" for the chemical evaluation and quality control of *Artemisia rupestris* L. (Asteraceae) and its products, has been isolated successfully applying a two-step HSCCC method (Ma *et al.*, 2005). However, this method had some limitations such as the noncalculation of the partition coefficient (*K*), leading to the isolation of rupestonic acid in two steps and time and solvent wasting. Furthermore, the acetic acid, which was used in the stationary phase, led to several complications such as formation of salts or degradation of the constituents. Trying to tackle this problem, Yang *et al*. carefully selected a suitable acid-free two-phase solvent system composed of *n*-hexane–ethyl acetate–methanol–water (3 : 5 : 3 : 5), and its application together with a better management of the flow rate not only led to the increase of the amount of the crude sample loaded, the peak resolution, and the yield but also avoided multistep separation, which leads to low recovery. Meanwhile, the potential degradation of the compounds was avoided since acid was not used (Yang *et al*., 2009).

The use of this technology is intensively growing particularly in China and contributing to the modernization of the Chinese herbal medicine (Sutherland and Fisher, 2009). Few examples can be cited: HSCCC was applied to isolate and purify bioactive compounds from the essential oil of the rhizomes of *Curcuma wenyujin* Y. H. Chen & C. Ling (Zingiberaceae), a traditional Chinese medicinal plant whose rhizomes and tubers specified as two herbal medicines in China Pharmacopoeia, *Rhizoma curcumae* and *Radix curcumae* (Yan *et al*., 2005). Two sesquiterpenes germacrone and curdione were successfully isolated and purified using two-phase solvent system composed of petroleum ether–ethanol–diethyl ether–water (5 : 4 : 0.5 : 1) in tail-to-head elution mode; 62 mg of germacrone and 93 mg of curdione were obtained from 658 mg of the essential oil each at over 95% purity.

HSCCC was also successfully applied for the separation and purification of sesquiterpene lactone compounds from the light petroleum extract of the Chinese medicinal herb *Radix linderae* [*Lindera aggregata* (Sims) Kosterm., Lauraceae] for the first time. Linderalactone and lindenenol, which are central components of this plant, were obtained in a one-step separation (Sun, Sun, and Liu, 2006). Another example refers at the application of a preparative two-dimensional countercurrent chromatography (2D-CCC) system for simultaneous separation and purification of oridonin and ponicidin from the crude extract of one of the most widely used medicinal plants in China, *Rabdosia rubescens* (Hemsl.) Hara (Lamiaceae), was presented by Lu *et al*. (2007). This system was based on the use of an analytical HSCCC instrument in the first dimension (1st-D) and a preparative upright countercurrent chromatography (UCCC) column in the second dimension (2nd-D). The interface was an in-house column-switching system with a sample loop and a 2nd-D column having a 10-fold capacity comparing to the 1st-D one. In this way, almost the whole interesting region from the first dimension was transferred online to the second dimension for further separation. The use of a pair of two-phase solvent systems composed of *n*-hexane–ethyl acetate–methanol–water (1 : 5 : 1 : 5 and 3 : 5 : 3 : 5) in the two dimensions permitted the simultaneous separation of oridonin and ponicidin with 95% purity, satisfactory resolution, and peak capacity.

5.3 Supercritical Fluid Chromatography (SFC)

As mentioned earlier, SFC is a relatively new technique offering multiple advantages to the analysis of complex mixtures; however, its application is wider covering not only analytical but also preparative needs. The specific advantages of SFC results in a growing interest and could be considered as an emergent technique, nowadays (Taylor, 2009). Preparative separation with SFC provides many advantages comparing to LC or GC because of the properties of Sc-CO_2, which is used as a mobile phase (Gerd, 2005). In SFC, the lower operating temperatures, the higher diffusivities of the solutes, and the lower viscosity of the eluent are observed. In addition, SFC provides advantages compared to LC such as: fast

Table 6 Suggested CCC solvent systems for the isolation of sesquiterpenoids and diterpenoids.

Categories	Isolated compounds	Source/Family (Part/Extract)	Solvent systems (volume ratio)	Conditions	References
Sesquiterpenes and sesquiterpenes lactones	Germacrone and curdione	*Curcuma wenyujin* Y. H. Chen & C. Ling/Zingiberaceae (**rhizome/essential oil, SD**)	Petroleum ether–ethanol–diethyl ether–water (5:4:0.5:1)	Ascending mode, forced elution of the lower phase	Yan *et al.* (2005)
	Rupestonic acid	*Artemisia rupestris* L./Asteraceae (**roots/dichloromethane**)	*n*-Hexane–ethyl acetate–methanol–water (3:5:3:5)	Descending mode, normal elution	Yang *et al.* (2009)
	Nootkatone	*Alpinia oxyphylla* Miquel/Zingiberaceae (**fruits/essential oil, SDE**)	*n*-Hexane–MeOH–water (5:4:1)	Descending mode, normal elution	Xie *et al.* (2009)
	β-Caryophyllene	*Vetix negundo* L. var. *heterophylla* (Franch.) Rehd/Verbenaceae (**leaf/essential oil, SD**)	*n*-Hexane–dichloromethane–acetonitrile (10:3:7)	Descending mode, normal elution	Xie *et al.* (2008)
	Costunolide and dehydrocostus lactone	*Aucklandia lappa* Decne/Asteraceae (**roots/light petroleum**)	Light petroleum–methanol–water (5:6.5:3.5)	Ascending mode, normal elution	Li, Sun, and Liu (2005)
	Atractylon and atractylenolide III	*Atractylodes macrocephala* Koidz./Asteraceae (**roots/ethyl acetate**)	Light petroleum–ethyl acetate–ethanol–water (4:1:4:1)	First descending, dual mode	Zhao and He (2006)
	Lactucopicrin	*Cichorium glandulosum* Boiss. et Huet./Asteraceae (**roots/ethyl acetate**)	*n*-Hexane–ethyl acetate–methanol–water (1.5:5:2.75:5)	Descending mode, normal elution, first step	Wu *et al.* (2007)
	11β,13-dihydrolactucin and lactucin		Ethyl acetate–methanol–water (20:1:20)	Descending mode, normal elution, second step	
	Linderalactone and lindenenol	*Lindera strychnifolia* (Sieb. et Zucc) F. Villar/Lauraceae (**roots/light petroleum**)	Light petroleum–ethyl acetate–methanol–water (5:5:6:4)	Descending mode, normal elution	Sun, Sun, and Liu (2006)
	Eremophilane-type sesquiterpenes	*Ligularia atroviolacea* (Franchet) Handel-Mazzetti/Asteraceae (**roots/ethanol**)	*n*-Hexane–ethyl acetate–ethanol–water (4:1:4:1)	Descending mode, normal elution	Yun Shi *et al.* (2008)
Diterpenes	Oridonin and ponicidin	*Rabdosia rubescens* (Hemsl.) Hara/Lamiaceae (**ethanol extract**)	*n*-Hexane–ethyl acetate–methanol–water (1:5:1:5) and (3:5:3:5)	2D-CCC, descending mode, and elution extrusion	Lu *et al.* (2007)

16-O-methylcafestol, 16-O-methylkahweol	*Coffea canephora* Pierre ex Froehner var. *robusta*/Rubiaceae (**green and roasted beans/MTBE, Soxhlet**)	Hexane–ethyl acetate–ethanol–water (5:2:5:2)	Descending mode, normal elution	Scharnhop and Winterhalter (2009)
Kahweol, dehydrokahweol, cafestol, and dehydrocafestol	*Coffea arabica* L./Rubiaceae (**green** and **roasted beans/MTBE, Soxhlet**)	Hexane–ethyl acetate–ethanol–water (5:2:5:2)	Descending mode, normal elution	Scharnhop and Winterhalter (2009)
Kaurenoic acid and polyalthic acid	*Copaifera glycycarpa* Ducke/Leguminoseae (**oleoresin**)	n-Hexane–acetonitrile–ethyl acetate (1:1:0:4)	Ascending mode, normal elution	De Souza *et al.* (2010)
ent-Labdane diterpenes (andrographolide and neoandrographolide)	*Andrographis paniculata* (Burm.F.) Wall. ex Nees/Acanthaceae (**extract of leaves**)	n-Hexane–ethyl acetate–methanol–water (1:4:2.5:2.5)	Descending mode, normal elution	Du, Jerz, and Winterhalter (2003)
Diterpenoid alkaloids	*Aconitum coreanum* (H. Léveillé) Rapaics/Ranunculaceae (**ethanol extract**)	Ethyl acetate–n-butanol–water (7:2:2:7)	Descending mode, normal elution	Tang *et al.* (2008)
Triptonide-like diterpenoids	*Tripterygium wilfordii* Hook.f/Celastraceae (**ethanol extract**)	n-Hexane–ethyl acetate–methanol–water (3:2:3:2)	Descending mode, normal elution	Peng *et al.* (2008)
Tanshinones	*Salvia miltiorrhiza* Bung/Lamiaceae (**ethyl acetate extract, reflux**)	Light petroleum–ethyl acetate–methanol–water (6:4:6.5:3.5)	Descending mode, normal elution	Sun *et al.* (2011)
Labdane diterpenes	*Juniperus communis* L./Cupressaceae (**berries/methanol**)	Chloroform:methanol: isopropanol:water (5:6:1:4)	First descending, dual mode	Martin *et al.* (2006)

analysis time, easy recovery of products by simple decompression, low consumption of organic solvents, and wider range of applicability (Li and Hsieh, 2008).

In 1990, Yamauchi and Saito fractionated cold-pressed lemon peel oil [*Citrus limon* (L.) Burm.f., Rutaceae] by SFC system using a silica column. A compound-based fractionation was performed, resulting in three different fractions containing the hydrocarbon compounds, the esters, and the alcohols and the aldehydes, respectively. The fractionation seems to be the result of modification of the mobile phase, which, at the beginning, consisted of pure CO_2 operating under two different pressure conditions and then a small percentage of ethanol modifier 0.05% was added (Yamauchi and Saito, 1990). Moreover, the isolation of the major constituent, the sesquiterpene (+)-davanone, has been performed from natural davana oil (*Artemisia pallens* Wall. ex DC., Asteraceae) by carbon dioxide-based semipreparative packed column. The purified davanone exhibited high optical rotation that proves the purity of this compound. The semipreparative separation was developed on a semipreparative 2-ethylpyridine column, whereas detection was obtained using a UV detector at 205 and 214 nm. The separation of the davanone was performed under 120 atm at 45 °C using a gradient elution program with methanol as modifier (Coleman *et al.*, 2007). Finally, the successful fractionation of rosemary (*R. officinalis* L., Lamiaceae) supercritical extract was performed by preparative SFC using a LC-Diol packed column (25 cm × 10 mm I.D., 5 µm) at 80 °C column temperature, 130 bar pressure, and 10% ethanol as modifier of the mobile phase (CO_2). The two different fractions obtained, one rich in carnosic acid (a diterpene compound) and the other rich in oxygenated terpenes such as camphor, borneol, and verbenone, were found to have enhanced antioxidant and antimicrobial activity, respectively (Ramírez *et al.*, 2006). Although SFC provides many advantages and facilitates drastically the isolation procedure of hydrophilic compounds (Bamba, 2008), there are limited data published in natural products area and more specifically related to sesquiterpenoids.

Overall, the isolation of natural products involves a number of fractionation and purification steps. Because a large number of small terpenoids are known to be labile, the possibility of decomposition or rearrangement of these sensitive compounds by elevated temperature, irradiation, or acidic conditions has to be considered during the different steps of isolation procedure.

6 STRUCTURE ELUCIDATION

Terpenoids, biosynthetically synthesized from several isoprene units usually modified by the addition of functional groups and ring formation, constitute a structurally diversified class of chemical compounds. Their structure determination is usually carried out through the combination of spectroscopic methods (UV, IR, CD, NMR, and MS). In addition, the presence of several chiral carbons on terpenoids structures necessitates the use of x-rays for the description of their absolute configuration. Hereafter, a brief discussion will be made regarding the identification of sesquiterpenes and small terpenoids using MS and NMR techniques as they are the most commonly used. Because identification procedure is highly structure dependent and there are many chemical classes of small terpenoids and sesquiterpenoids, only some general information will be given.

MS is a valuable technique not only for the analysis and characterization of the organic molecules but also for the structural characterization thereof as it provides many structural data. Specifically, low molecular weight terpenoids (monoterpenoids, sesquiterpenoids, and diterpenoids), compounds usually smaller than 500 Da, have been studied by MS either directly or after GC and/or LC separation. Normally, GC-MS apparatus are equipped with mass spectrometers using hard ionization techniques, whereas LC-MS instruments use soft ionization techniques.

Spectroscopic data of terpenoids are usually extracted from the GC-MS analysis of these compounds equipped with EI or CI sources. EI mainly and in some cases CI sources provide useful structural information depending on the derived fragmentation pattern. For instance, the study of sesquiterpene lactones with various structural types from *Achillea* spp. (Asteraceae) using both EI and CI techniques led to the characterization of structural features of these compounds. Particularly, EI-MS provided useful structural information for different types of sesquiterpene lactones (matricin derivatives, matricarin derivatives, artabsin derivatives, 3-oxa-artabsin derivatives 7,8-guaianolide derivatives, guaianolide-endo-peroxides, and eudesmanolides). Characteristic fragments were observed

using EI in positive mode corresponding to the loss of acyl groups (acetyl, tiglyl, or angelyl) or the elimination of water in the case of hydroxyl groups presence on the structure. In addition, in the case of guaianolide-endo-peroxides, a typical ion at m/z 111 was observed, along with a loss of molecular oxygen. Concerning molecular ion intensities, in some cases, M^+ ions are obtain (matricin, matricarin, and 3-oxaartabsin derivatives), whereas in other cases (artabsin derivatives, 7,8-guaianolides, guaianolide-endo-peroxides, and eudesmanolides), molecular ions are hardly detectable. On the other hand, CI-MS with ammonia as reactant gas yields quasimolecular adduct ions $[M + NH_4]^+$ with high intensity for each type of sesquiterpene lactone, which could be used for the estimation of their molecular weight (Reznicek *et al.*, 1997). A variety of monoterpenes containing mono-, bi-, and tricyclic derivatives have been studied by GC-EI-MS. Diagnostic peaks with 100% relative abundance at m/z 93 in most of the cases and at m/z 121 and 81 in some other cases, corresponding to specific fragments, have been observed and could be used for the detection of these compounds (Yermakov *et al.*, 2010).

CI-MS with methane or isobutane as reagent gases seems to fail to afford molecular weight information for a large number of mono- and sesquiterpene alcohols and their esters, whereas the use of ammonia as a reagent gas leads to the formation of abundant $[M+18]^+$ quasimolecular adduct ions helpful for their identification. Alcohols and esters could be distinguished by mean of CI spectra according to their characteristic fragmentation patterns. Alcohols show a mass difference of 18 mass units, and esters display the difference of the respective acid in the patterns of the CI spectra. Specifically, oxygenated terpenoids have been also studied using CI with OH^- as reagent ion. With some exceptions, the $[M–H]^-$ ion is sufficiently stable and observed with a high relative abundance. Nucleophilic displacement by the OH^- ion produces carboxylate ions from terpene esters, giving information about the acid part of the ester (Lange and Schultze, 1987).

On the other hand, applications of soft ionization techniques such as API [ESI, APCI, and atmospheric pressure photoionization (APPI)] in terpenoids do not provide extensive fragmentation and are, therefore, from a spectroscopic point of view, not very informative. Fragmentation of molecular or pseudomolecular ions usually observed with soft ionization techniques can be obtained by in-source fragmentation or by tandem mass spectrometry using IT or triple quadrupole analyzers. In the latter case, fragmentation arises through collision with noble gases and techniques such as CID.

Small terpenoids have been studied using both ESI and APCI methods, and different results have been excluded depending on the structures of interest. Taxoids are a class of diterpenoids with significant biological importance; hence, their structural features and fragmentation behavior have been studied using both ESI and APCI in positive mode. Characteristic fragments were observed depending on the substitution of each molecule, whereas the 5/7/6-type taxoids underwent characteristic losses of 58 or 118 from ions produced by both APCI and ESI sources. A number of fragmentation rules have been developed, which could facilitate the rapid screening and structural characterization of taxoids in plant extracts by HPLC/MS (Ye and Guo, 2005). Another very important class of compounds, ginkgolides together with bilobalide, have been investigated by MS using APCI (Jensen *et al.*, 2002; Mauri *et al.*, 2003) and ESI (Sun *et al.*, 2005) in either negative or positive mode. Mostly, pseudomolecular ions such as $[M+H]^+$, $[M+NH_4]^+$, $[M+Na]^+$, and $[M-H]^-$ were observed. In addition, on-source fragmentation and MS/MS spectra have contributed to the structural elucidation of these compounds through their fragmentation patterns and for the analysis of these compounds using LC-MS interfaces (Lang *et al.*, 2004; Lu *et al.*, 2002).

Toward the same direction, aiming to the identification of small terpenoids and sesquiterpenoids, NMR spectroscopy gained a universal applicability as it plays a crucial role for the structural characterization by offering a greater variety of accessible experimental parameters. The structural elucidation is a demanding task due to the large variety of sesquiterpene substitution pattern and their configuration isomerism. A distinctive advantage of the NMR technique is that the measurements can be made on the native molecule without any introduction of foreign isotopes or reporter groups that might alter the structure, thus giving rise to artifacts (see **NMR of Small Molecules**). Specifically, for sesquiterpenoids and other small molecular weight-related compounds, the structure elucidation using NMR is rather complicated, especially due to their particular structural features allowing extended

degree of isomerization, which increase drastically the possible candidate structures.

The nonpolar nature of sesquiterpenoid derivatives imposes the utilization of nonpolar deuterated solvents such as dichloromethane-d_2, chloroform-d_3, toluene-d_8, benzene-d_6, in addition to relatively polar such as acetone-d_6, methanol-d_4, and dimethylsulfoxide-d_6, when oxygenated compounds are analyzed. However, rearrangements of sesquiterpenes may occur when chloroform is used as a solvent for NMR investigations (König, Bülow, and Saritas, 1999).

Generally, the complexity of ^{1}H NMR spectra of sesquiterpenes depends strongly on the numbers and the character of substituents. The relative intensity of signals, their multiplicity, and chemical shift values provide initial structure information. Because, usually, aromatic features are not incorporated on their basic structure, the proton signals of sesquiterpenes are generally shielded (0–3 ppm), affording crowded ^{1}H NMR spectra in this region. However, some peaks may be also detected relatively up-field when oxygenated or unsaturated features such as the *exo*-methylenes and oxygen-containing species are present. Even if ^{1}H NMR spectra offers informative data, which may even orient the further experimentation, ^{13}C NMR remains the method of choice for the determination of the terpenoids structure in general and for the sesquiterpenes, in particular. Similarly to ^{1}H NMR spectra, ^{13}C NMR spectra and decoupled ^{13}C NMR DEPT (Distortionless Enhancement by Polarization Transfer) spectra follow the same motif. Most of the carbons are resonated in high fields (10–50 ppm), whereas, only in oxygenated or unsaturated derivatives, characteristic carbons are resonated

in the low field area. However, the number of carbons and the observed chemical shifts are very critical factors for determining terpenoid structures. 2D NMR experiments, and especially correlation spectroscopy (COSY), nuclear Overhauser effect correlation spectroscopy (NOESY), and total correlation spectroscopy (TOCSY), are commonly used, not only establishing the vicinity of protons and carbons but also contributing to the determination of stereochemistry and configuration aspects. Hereafter, a characteristic example is given revealing the challenging nature of the identification procedure and the complicated structural features of small terpenoids as well.

This example concerns the structure determination of sesquiterpenoids isolated from *Senecio argunensis* Turcz. (Asteraceae) (Xie *et al.*, 2010). Specifically, two new isodaucane sesquiterpenoids have been isolated and characterized together with the known one artabotrol **2** (Fleischer, Waigh, and Waterman, 1997). ^{1}H and ^{13}C NMR and DEPT, ^{1}H–^{1}H COSY, HMBC (Heteronuclear multiple-bond correlation spectroscopy), and NOESY NMR experiments were used for the study. According to both 1D and 2D NMR spectra, the significant differences observed between the chemical shifts of isodauc-7(14)-en-6,10-diol **1** and **2** were due to the different stereochemistry. The NOESY cross peaks observed between Me-15 and H-6 strongly suggest that Me-15 and H-6 were in β orientation, whereas cross peaks between H-10 and H-5 and between H-5 and H-4 indicated the α orientation of H-4, H-5, and H-10 (Figure 7). The large coupling constants for H-10 (dd, J = 11.5, 5.0 Hz) also suggested the β orientation of the hydroxy group at C-10. Thus, the structure of **1** was assigned as isodauc-7(14)-en-6α,10β-diol.
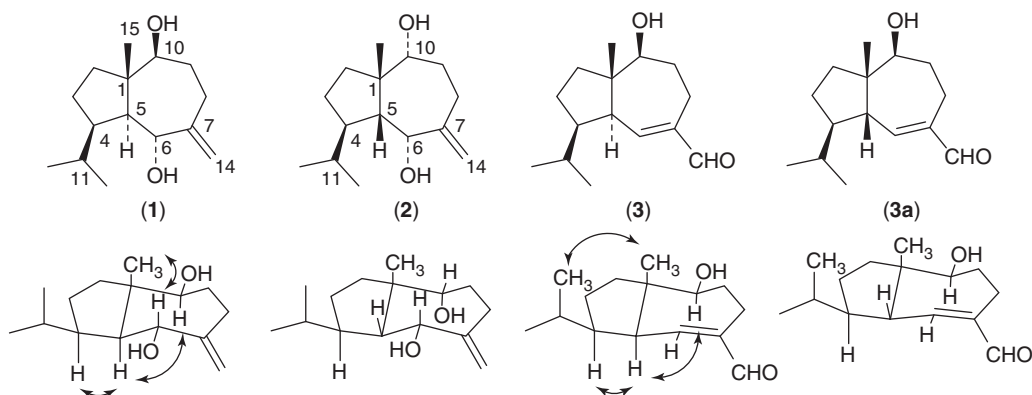
Figure 7   Structures of isodaucane sesquiterpenoids (arrows indicate NOE correlations).

From the same plant material, using the HSQC (Heteronuclear Single Quantum Correlation), ^{1}H–^{1}H COSY, and HMBC spectra, the plane structure of compound **3** was assigned as 10-hydroxyisodauc-6-en-14-al, which is identical with the plane structure of aphanamol II **3a** (Nishizawa *et al*., 1984). The relative stereochemistry of **2** was deduced by NOESY spectrum (Figure 7). In this spectrum, Me-15 and H-5 did not show any NOESY correlation, indicating a *trans*-configuration of Me-15 and H-5, whereas the correlations between Me-15 and Me-12, H-5 and H-10, and H-5 and H-4 were observed. These NOESY correlations suggested that Me-15, isopropyl group, and the hydroxy group at C-10 were all at β orientation and H-5 and H-10 were at α orientation. The large coupling constants for H-10 (dd, $J = 11.5, 4.0$ Hz) also supported the β orientation of the hydroxy group at C-10. From these data, the structure of **3** was established as 10β-hydroxyisodauc-6-en-14-al, a 5-epimer of **3a**.

From this specific example and the information given previously, it becomes obvious that the structure elucidation of small terpenoids and sesquiterpenoids is a demanding, however, appealing task. Furthermore, combination of spectroscopic techniques are required and in deep investigation of the unambiguous structure elucidation thereof.

7   CONCLUSION

Sesquiterpenoids and other small terpenoids encompass one of the most variable and important subgroup of terpenoids essential for every plant organism. Apart from their physiological role, they display significant physicochemical and biological properties, which enable their exploitation as medicinal, cosmetic, or agrochemical agents. Due to their high potential, historically early enough have been placed under scientific investigation, which resulted in the discovery of numerous sesquiterpenoids and small terpenoids characterized by a wide array of biological and pharmacological properties. Toward this effort, phytochemical analysis plays a central role in modern biodiscovery process.

Technological advances and upgraded instrumentation, in addition to new methodological concepts and approaches recently introduced in natural products analysis and chemistry, have been used in the area of small terpenoids and sesquiterpenoids. This technological evolution facilitates drastically their study in all successive steps from extraction to structure elucidation. Modern extraction techniques such as PLE, SFE, SWE, or MAE are proven to be more efficient, fast, less costly overall, and reproducible compared to traditional ones while offering many alternatives that increase indirectly the selectivity and credibility. Similarly, separation procedures have been optimized and accelerated especially by the introduction of technologically advanced CCC apparatus and preparative SFC instruments, enabling the more straightforward and efficient purification specifically in the area of terpenoids. Moreover, technical advances, methodological optimizations, and increased rate of instrumental hyphenation are currently enabling the more complete, accurate, reproducible and effective analysis, and characterization of complicated mixtures containing terpenoids. Several 1D and 2D, GC and LC systems hyphenated to multiple detectors, analyzers, and platforms are, nowadays, incorporated, which opens new directions in dereplication and in the discovery of novel terpenoid leads.

REFERENCES

d'Acampora Zellner, B., Dugo, P., Dugo, G., *et al*. (2010) Analysis of essential oils, in *Handbook of Essential Oils: Science, Technology, and Applications*, eds. K. H. C. Baser and G. Buchbauer, CRC Press, Boca Raton, pp. 165–183.

Aberham, A., Cicek, S. S., Schneider, P., *et al*. (2010) *J. Agric. Food Chem.*, **58**(20), 10817–10823.

Abourashed, E. A., Galal, A. M., El-Feraly, F. S., *et al*. (2001) *Planta Med.*, **67**(7), 681–682.

Adlard, E. (1987) *Chromatographia*, **23**(9), 691.

Al-Harrasi, A. and Al-Saidi, S. (2008) *Molecules*, **13**(9), 2181–2189.

Ali, M. S., Ibrahim, S. A., Ahmed, S., *et al*. (2007) *Chem. Biodiv.*, **4**(1), 98–104.

Ali, N. A. A., Marongiu, B., Piras, A., *et al*. (2011) *Nat. Prod. Res.*, **25**(14), 1366–1369.

Almeida, P., Mezzomo, N. and Ferreira, S. S. (2012) *Food Bioprocess. Tech.*, **5**(2), 548–559.

Amoah, S. K. S., de Oliveira, F. L., da Cruz, A. C. H., *et al*. (2013) *Phytochemistry*, **87**, 126–132.

Attaway, J. A., Barabas, L. J. and Wolford, R. W. (1965) *Anal. Chem.*, **37**(10), 1289–1290.

Auerbach, R. H., Kenan, D. and Davidson, G. (2000) *J. AOAC Int.*, **83**(3), 621–626.

Avery, B. A., Venkatesh, K. K. and Avery, M. A. (1999) *J. Chromatogr. B Biomed. Sci. Appl.*, **730**(1), 71–80.

Avula, B., Wang, Y.-H., Smillie, T., *et al*. (2009) *Chromatographia*, **70**(5), 797–803.

Bakkali, F., Averbeck, S., Averbeck, D., *et al*. (2008) *Food Chem. Toxicol.*, **46**(2), 446–475.

Bamba, T. (2008) *J. Sep. Sci.*, **31**(8), 1274–1278.

Barboni, T., Venturini, N., Paolini, J., *et al*. (2010) *Food Chem.*, **122**(4), 1304–1312.

Barnhart, R. K. (1995) , in *The Barnhart Concise Dictionary of Etymology*, ed. R. K. Barnhart, HarperCollins Publishers, New York.

Bedoya, L. M., Abad, M. J. and Bermejo, P. (2008) *Curr. Signal Transd. T*, **3**(2), 82–87.

Benini, C., Ringuet, M., Wathelet, J.-P., *et al*. (2012) *Flavour Frag. J.*, **27**, 356–366.

Benthin, B., Danz, H. and Hamburger, M. (1999) *J. Chromatogr. A*, **837**(1–2), 211–219.

Berger, T. A. and Berger, B. K. (2011) *J. Chromatogr. A*, **1218**(16), 2320–2326.

Berthod, A. (2007) *J. Liq. Chromatogr. Relat. Technol.*, **30**, 1447–1463.

Bhandari, P., Gupta, A. P., Singh, B., *et al*. (2005) *J. Sep. Sci.*, **28**(17), 2288–2292.

Bhat, S. V., Nagasampagi, B. A. and Sivakumar, M. (2005a) Terpenoids, in *Chemistry of Natural Products*, Springer, Berlin, pp. 115–116.

Bhat, S. V., Nagasampagi, B. A. and Sivakumar, M. (2005b) *Chemistry of Natural Products*, Springer, Berlin, pp. 127–130.

Bicchi, C., Balbo, C. and Rubiolo, P. (1997a) *J. Chromatogr. A*, **779**(1–2), 315–320.

Bicchi, C., D'Amato, A., Manzin, V., *et al*. (1997b) *Flavour Fragr. J.*, **12**(2), 55–61.

Bicchi, C., Binello, A. and Rubiolo, P. (2000) *Phytochem. Anal.*, **11**(3), 179–183.

Bicchi, C., Cagliero, C. and Rubiolo, P. (2011) *Flavour Fragr. J.*, **26**(5), 321–325.

Blum, C., Kubeczka, K.-H. and Becker, K. (1997) *J. Chromatogr. A*, **773**(1–2), 377–380.

Bohlmann, J., Meyer-Gauen, G. and Croteau, R. (1998) *Proc. Natl. Acad. Sci.*, **95**(8), 4126–4133.

Bortolomeazzi, R., Berno, P., Pizzale, L., *et al*. (2001) *J. Agric. Food Chem.*, **49**(7), 3278–3283.

Bousbia, N., Abert Vian, M., Ferhat, M. A., *et al*. (2009a) *Food Chem.*, **114**(1), 355–362.

Bousbia, N., Vian, M. A., Ferhat, M. A., *et al*. (2009b) *J. Food Eng.*, **90**(3), 409–413.

Boutekedjiret, C., Bentahar, F., Belabbes, R., *et al*. (2003) *Flavour Fragr. J.*, **18**(6), 481–484.

Bouvier, F., Rahier, A. and Camara, B. (2005) *Prog. Lipid Res.*, **44**(6), 357–429.

Brokl, M., Fauconnier, M.-L., Benini, C., *et al*. (2013) *Molecules*, **18**, 1783–1797.

Brown, A. E., Riddick, E. W., Aldrich, J. R., *et al*. (2006) *J. Chem. Ecol.*, **32**(11), 2489–2499.

Bruno, S. (1961) *Farmaco*, **16**, 481–486.

Camel, V. (2001) *Analyst*, **126**(7), 1182–1193.

Camponovo, F. F., Wolfender, J.-L., Maillard, M. P., *et al*. (1995) *Phytochem. Anal.*, **6**(3), 141–148.

Cane, D. E. (1990) *Chem. Rev.*, **90**(7), 1089–1103.

Carlson, L. H. C., Machado, R. A. F., Spricigo, C. B., *et al*. (2001) *J. Supercrit. Fluids*, **21**(1), 33–39.

Casas, L., Mantell, C., Rodríguez, M., *et al*. (2007) *J. Supercrit. Fluids*, **41**(1), 43–49.

Castello, G., Moretti, P. and Vezzani, S. (2009) *J. Chromatogr. A*, **1216**(10), 1607–1623.

Chan, C.-H., Yusoff, R., Ngoh, G.-C., *et al*. (2011) *J. Chromatogr. A*, **1218**(37), 6213–6225.

Chanotiya, C. S. and Yadav, A. (2008) *Nat. Prod. Commun.*, **3**(2), 263–266.

Chaturvedi, D. (2011) Sesquiterpene lactones: structural diversity and their biological activities, in *Opportunity, Challenge and Scope of Natural Products in Medicinal Chemistry*, ed. V. K. Tiwari, Research Signpost, Kerala (India), pp. 313–334.

Chemat, S., Lagha, A., AitAmar, H., *et al*. (2004) *Flavour Fragr. J.*, **19**(3), 188–195.

Chen, R. (2009) *Chromatogr. Today*, **2**(1), 11–13.

Cheng, B. H., Lin, C. Y., Yeh, T. F., *et al*. (2012) *J. Agric. Food Chem.*, **60**(31), 7623–7628.

Chester, T. L. and Pinkston, J. D. (2004) *Anal. Chem.*, **76**(16), 4606–4613.

Chizzola, R. (2010) *Nat. Prod. Commun.*, **5**(9), 1477–1492.

Cimpan, G. and Gocan, S. (2002) *J. Liq. Chromatogr. Relat. Technol.*, **25**, 2225–2292.

Claeson, P., Zygmunt, P. and Högestätt, E. D. (1991) *Pharmacol. Toxicol.*, **69**, 173–177.

Clarkson, C., Madikane, E. V., Hansen, S. H., *et al*. (2007) *Planta Med.*, **73**(06), 578–584.

Coleman, W. M., Dube, M. F., Ashraf-Khorassani, M., *et al*. (2007) *J. Agric. Food Chem.*, **55**(8), 3037–3043.

Conway, W. D. (1991) *J. Chromatogr. A*, **538**(1), 27–35.

Conway, W. D. (2011) *J. Chromatogr. A*, **1218**(36), 6015–6023.

Cordell, G. A. (1976) *Chem. Rev.*, **76**(4), 425–460.

Cornwell, C. P., Reddy, N., Leach, D. N., *et al*. (2000) *Flavour Fragr. J.*, **15**(5), 352–361.

Cramers, C. A., Janssen, H.-G., van Deursen, M. M., *et al*. (1999) *J. Chromatogr. A*, **856**(1–2), 315–329.

Croteau, R., Kutchan, T. M. and Lewis, N. G. (2000) Natural products (secondary metabolites), in *Biochemistry & Molecular Biology of Plants*, eds. B. Buchanan, W. Gruissem and R. Jones, American Society of Plant Physiologists, Rockville, MD, pp. 1250–1318.

Da Porto, C. and Decorti, D. (2009) *Ultrason. Sonochem.*, **16**(6), 795–799.

Dai, D. N., Hoi, T. M., Thang, T. D., *et al*. (2012) *Nat. Prod. Commun.*, **7**(2), 231–234.

Darjazi, B. B. (2011) *J. Med. Plants Res.*, **5**(13), 2839–2847.

Dawidowicz, A. L., Rado, E., Wianowska, D., *et al*. (2008) *Talanta*, **76**(4), 878–884.

De Souza, P. A., Rangel, L. P., Oigman, S. S., *et al*. (2010) *Phytochem. Anal.*, **21**(6), 539–543.

De, L. S. G., Neto, J. M. M., Cito, A. M. G. L., *et al*. (2009) *J. Chil. Chem. Soc.*, **54**(1), 55–57.

Debrunner, B., Neuenschwander, M. and Brenneisen, R. (1995) *Pharm. Acta Helv.*, **70**(2), 167–173.

Demarne, F. E., Viljoen, A. M. and Van, d W J J A. (1993) *J. Essent. Oil Res.*, **5**(5), 493–499.

van Den Dool, H. and Kratz, P. D. (1963) *J. Chromatogr. A*, **11**, 463–471.

Deng, C., Wang, A., Shen, S., *et al*. (2005) *J. Pharm. Biomed. Anal.*, **38**(2), 326–331.

Deng, C., Liu, N., Gao, M., *et al*. (2007) *J. Chromatogr. A*, **1153**(1–2), 90–96.

Diaz-Maroto, M. C., Diaz-Maroto Hidalgo, I. J., Sanchez-Palomo, E., *et al.* (2005) *J. Agric. Food Chem.*, **53**(13), 5385–5389.

Doimo, L., Fletcher, R. J. and D'Arcy, B. R. (1999) *J. Essent. Oil Res.*, **11**, 291–299.

Dolara, P., Corte, B., Ghelardini, C., *et al.* (2000) *Planta Med.*, **66**(04), 356–358.

Dolman, D. M., Knight, D. W., Salan, U., *et al.* (1992) *Phytochem. Anal.*, **3**(1), 26–31.

Donelian, A., Carlson, L. H. C., Lopes, T. J., *et al.* (2009) *J. Supercrit. Fluids*, **48**(1), 15–20.

Dong, J.-Y., Ma, X.-Y., Cai, X.-Q., *et al.* (2013) *Phytochemistry*, **85**, 122–128.

Dost, K. and Davidson, G. (2003) *Analyst*, **128**(8), 1037–1042.

Drew, D., Krichau, N., Reichwald, K., *et al.* (2009) *Phytochem. Rev.*, **8**(3), 581–599.

Du, Q., Jerz, G. and Winterhalter, P. (2003) *J. Chromatogr. A*, **984**(1), 147–151.

Dubber, M. J. and Kanfer, I. (2006) *J. Pharm. Biomed. Anal.*, **41**(1), 135–140.

Dugo, G., Lamonica, G., Cotroneo, A., *et al.* (1992) *Perf. Flav.*, **17**, 57–74.

Elakovich Stella, D. (1988) Terpenoids as models for new agrochemicals, in *Biologically Active Natural Products*, American Chemical Society, Washington, DC, pp. 250–261.

Eller, F. J. and Taylor, S. L. (2004) *J. Agric. Food Chem.*, **52**(8), 2335–2338.

ElSohly, H. N., Croom, E. M. and ElSohly, M. A. (1987) *Pharm. Res.*, **4**(3), 258–260.

Emura, M., Nohara, I., Toyoda, T., *et al.* (1997) *Flavour Fragr. J.*, **12**(1), 9–13.

Esclapez, M. D., García-Pérez, J. V., Mulet, A., *et al.* (2011) *Food Eng. Rev.*, **3**(2), 108–120.

Exarchou, V., Krucker, M., van Beek, T. A., *et al.* (2005) *Magn. Reson. Chem.*, **43**(9), 681–687.

Facundo, V. A., Pinto, A. C. and Rezende, C. M. (2005) *Flavour Frag. J.*, **20**(1), 93–95.

Ferreira, J. F. S. and Gonzalez, J. M. (2009) *Phytochem. Anal.*, **20**(2), 91–97.

Fleischer, T. C., Waigh, R. D. and Waterman, P. G. (1997) *J. Nat. Prod.*, **60**(10), 1054–1056.

Fojtová, J., Lojková, L. and Kubáň, V. (2008) *J. Sep. Sci.*, **31**(1), 162–168.

Foucault, A. P. (1994) Solvent system in centrifugal partition chromatography, in *Centrifugal Partition Chromatography*, ed. A. P. Foucault, Marcel Dekker, New York, pp. 71–97.

Foucault, A. P. and Chevolot, L. (1998) *J. Chromatogr. A*, **808**(1–2), 3–22.

Frank, H., Nicholson, G. J. and Bayer, E. (1977) *J. Chromatogr. Sci.*, **15**(5), 174–176.

Gabriels, M. and Plaizier-Vercammen, J. A. (2003) *J. Chromatogr. Sci.*, **41**(7), 359–366.

Gabriels, M. and Plaizier-Vercammen, J. (2004) *J. Chromatogr. Sci.*, **42**(7), 341–347.

Gamiz-Gracia, L. and Luque de Castro, M. D. (2000) *Talanta*, **51**(6), 1179–1185.

Gerd, B. (2005) *J. Food Eng.*, **67**(1–2), 21–33.

Gershenzon, J. and Dudareva, N. (2007) *Nat. Chem. Biol.*, **3**(7), 408–414.

Giddings, J. C. (1995) *J. Chromatogr. A*, **703**(1–2), 3–15.

Gil-Av, E., Feibush, B. and Charles-Sigler, R. (1966) *Tetrahedron Lett.*, **7**(10), 1009–1015.

Glasl, S., Kastner, U., Jurenitsch, J., *et al.* (1999) *J. Chromatogr. B Biomed. Sci. Appl.*, **729**(1–2), 361–368.

Glasl, S., Gunbilig, D., Narantuya, S., *et al.* (2001) *J. Chromatogr. A*, **936**(1–2), 193–200.

Gocan, S. (2009) Terpenoids: TLC analysis, in *Encyclopedia of Chromatography*, 3rd edn, ed. J. Cazes, CRC Press, Boca Raton, pp. 2299–2303.

Gocan, S. and Cimpan, G. (2005) *J. Liq. Chromatogr. Relat. Technol.*, **27**, 1377–1411.

Golmakani, M.-T. and Rezaei, K. (2008) *Eur. J. Lipid Sci. Technol.*, **110**(5), 448–454.

Gonzalez-Coloma, A., Martín, L., Mainar, A. M., *et al.* (2012) *Phytochem. Rev.*, **11**(4), 433–446.

Goodman, J. and Walsh, V. (2001) *The Story of Taxol: Nature and Politics in the Pursuit of an Anti-Cancer Drug*, Cambridge University Press, Cambridge, p. 115.

Grantham, P. J. and Douglas, A. G. (1980) *Geochim. Cosmochim. Acta*, **44**(11), 1801–1810.

Guo, Y. (2006) *Anal. Lett.*, **39**(9), 2055–2059.

Guy, I., Charles, B., Guinaudeau, H., *et al.* (2001) *J. Essent. Oil Res.*, **13**(3), 200–201.

Han, Y.-F., Cao, G.-X., Gao, X.-J., *et al.* (2010) *Food Chem.*, **120**(4), 1083–1088.

Handa, S. S. (2008) An overview of extraction techniques for medicinal and aromatic plants, in *Extraction Technologies for Medicinal and Aromatic Plants*, eds. S. S. Handa, S. P. S. Khanuja, G. Longo and D. D. Rakesh, Internation centre for science and high technology ICS-UNIDO, Trieste, pp. 21–54.

Haring, H. G., Rijkens, F., Boelens, H., *et al.* (1972) *J. Agric. Food Chem.*, **20**(5), 1018–1021.

Hemalatha, S., Mandal, V. and Mohan, Y. (2007) *Pharmacogn. Rev.*, **1**(1), 7–18.

Hernandez, A., Pascual, C., Sanz, J., *et al.* (1980) *Phytochemistry*, **21**(12), 2909–2911.

Herrero, M., Mendiola, J. A., Cifuentes, A., *et al.* (2010) *J. Chromatogr. A*, **1217**(16), 2495–2511.

Heuskin, S., Godin, B., Leroy, P., *et al.* (2009) *J. Chromatogr. A*, **1216**(14), 2768–2775.

Hiltunen, R. and Laakso, I. (1995) *Flavour Fragr. J.*, **10**(3), 203–210.

Hoigné, J., Windmer, H. and Gäumann, T. (1963) *J. Chromatogr. A*, **11**, 459–462.

Holm, Y. and Hiltunen, R. (1997) *Flavour Fragr. J.*, **12**(5), 335–339.

Holm, Y., Laakso, I. and Hiltunen, R. (1994) *Flavour Fragr. J.*, **9**(5), 223–227.

Holm, Y., Laakso, I., Hiltunen, R., *et al.* (1997a) *Flavour Fragr. J.*, **12**(4), 241–246.

Holm, Y., Vuorela, P. and Hiltunen, R. (1997b) *Flavour Fragr. J.*, **12**(6), 397–400.

Hostettmann, K., Marston, A. and Hostettmann, M. (1998) Special column chromatography, in *Preparative Chromatography Techniques: Applications in Natural Product Isolation*, ed. C. Messerschmidt, Springer-Verlag, Berlin 2nd Completely Revised and Enlarged Edition, pp. 33–39.

Huang, L., Zhong, T., Chen, T., *et al.* (2007) *Rapid Commun. Mass Spectrom.*, **21**(18), 3024–3032.

Hudaib, M., Gotti, R., Pomponio, R., *et al.* (2003) *J. Sep. Sci.*, **26**(1–2), 97–104.

Huo, Y., Shi, H., Li, W., *et al.* (2010) *J. Pharm. Biomed. Anal.*, **51**(4), 942–946.

Ito, Y. (2005a) *J. Chromatogr. A*, **1065**(2), 145–168.

Ito, Y. (2005b) *Sep. Purif. Rev.*, **34**, 131–154.

Itokawa, H., Masuyama, K., Morita, H., *et al.* (1993) *Chem. Pharm. Bull.*, **41**(6), 1183–1184.

Ivanescu, B., Vlase, L., Corciova, A., *et al.* (2011) *Nat. Prod. Res.*, **25**(7), 716–722.

Ivanovic, J., Ristic, M. and Skala, D. (2011) *J. Supercrit. Fluids*, **57**(2), 129–136.

Jensen, A. G., Ndjoko, K., Wolfender, J.-L., *et al.* (2002) *Phytochem. Anal.*, **13**(1), 31–38.

Juchelka, D., Steil, A., Witt, K., *et al.* (1996) *J. Essent. Oil Res.*, **8**, 487–497.

Juergens, A., Feldhaar, H., Feldmeyer, B., *et al.* (2006) *Biochem. Syst. Ecol.*, **34**(2), 97–113.

Julianti, T., Hata, Y., Zimmermann, S., *et al.* (2011) *Fitoterapia*, **82**(7), 955–959.

Kaškonienė, V., Kaškonas, P., Jalinskaitė, M., *et al.* (2011) *Chromatographia*, **73**(1), 163–169.

Kaufmann, B. and Christen, P. (2002) *Phytochem. Anal.*, **13**(2), 105–113.

Khan, V. A., Dubovenko, Z. V. and Pentegova, V. A. (1983) *Khim. Drev.*, **4**, 34–42.

Kimbaris, A. C., Siatis, N. G., Daferera, D. J., *et al.* (2006) *Ultrason. Sonochem.*, **13**(1), 54–60.

Kjeldsen, F., Christensen, L. P. and Edelenbos, M. (2003) *J. Agric. Food Chem.*, **51**(18), 5400–5407.

Köhler, M., Haerdi, W., Christen, P., *et al.* (1997a) *J. High Resolut. Chromatogr.*, **20**(2), 62–66.

Köhler, M., Haerdi, W., Christen, P., *et al.* (1997b) *Phytochem. Anal.*, **8**(5), 223–227.

König, W. A. (1998) *Chirality*, **10**(5), 499–504.

König, W. A., Lutz, S., Mischnick-Lübbecke, P., *et al.* (1988a) *J. Chromatogr. A*, **447**, 193–197.

König, W. A., Lutz, S. and Wenz, G. (1988b) *Angew. Chem. Int. Ed. Engl.*, **27**(7), 979–980.

König, W. A., Rieck, A., Hardt, I., *et al.* (1994) *J. High Resolut. Chromatogr.*, **17**(5), 315–320.

König, W. A., Bülow, N. and Saritas, Y. (1999) *Flavour Frag. J.*, **14**(6), 367–378.

Korytár, P., Janssen, H.-G., Matisová, E., *et al.* (2002) *Trends Anal. Chem.*, **21**(9–10), 558–572.

Kováts, E. (1958) *Helv. Chim. Acta*, **41**(7), 1915–1932.

Kowalska, T., Sherma, J. and Waksmundzka-Hajnos, M. (2008) Overview of the field of TLC in phytochemistry and the structure of the book, in *Thin Layer Chromatography in Phytochemistry*, eds. T. Kowalska, J. Sherma and M. Waksmundzka-Hajnos, CRC Press, Boca Raton, pp. 5–9.

Kpoviessi, D. S. S., Gbaguidi, F. A., Kossouoh, C., *et al.* (2011) *J. Med. Plant Res.*, **5**(18), 4640–4646.

Kreis, P. and Mosandl, A. (1992) *Flavour Fragr. J.*, **7**(4), 187–193.

Kubeczka, K.-H. (1981) Standardization and analysis of essential oils, in *A Perspective of the Perfumes and Flavours Industry in India: Proceedings of the Fifth Seminar of the Perfumes and Flavours Association of India*, ed. S. Jain, Perfumes Flavours Association of India, New Delhi.

Kubeczka, K.-H. (2010) History and sources of essential oil research, in *Handbook of Essential Oils: Science, Technology, and Applications*, eds. K. H. C. Baser and G. Buchbauer, CRC Press, Boca Raton, pp. 3–38.

Kula, J., Masarweh, A. and Gora, J. (1996) *J. Essent. Oil Res.*, **8**(4), 453–454.

Kuzuyama, T. and Seto, H. (2003) *Nat. Prod. Rep.*, **20**(2), 171–183.

Lachenmeier, D. W. (2007) *Food Res. Int.*, **40**(1), 167–175.

Lai, S.-M., Chen, I. W. and Tsai, M.-J. (2005) *J. Chromatogr. A*, **1092**(1), 125–134.

Lalli, J. Y. Y., Viljoen, A. M., Baser, K. H. C., *et al.* (2006) *J. Essent. Oil Res.*, **18** Special Edition, 89–105.

Lambert, M., Wolfender, J.-L., Staerk, D., *et al.* (2006) *Anal. Chem.*, **79**(2), 727–735.

Lang, Q., Wai, C.M., Ang, C.Y., *et al.* (2004) *J. AOAC Int.*, **87**, 815–826.

Lange, G. and Schultze, W. (1987) *Flavour Frag. J.*, **2**, 63–73.

Lapkin, A. A., Walker, A., Sullivan, N., *et al.* (2009) *J. Pharm. Biomed. Anal.*, **49**(4), 908–915.

Lawrence, M. B. (1968) *J. Chromatogr. A*, **38**, 535–537.

Li, W. and Fitzloff, J. F. (2002) *J. Pharm. Biomed. Anal.*, **30**(1), 67–75.

Li, F. and Hsieh, Y. (2008) *J. Sep. Sci.*, **31**(8), 1231–1237.

Li, A., Sun, A. and Liu, R. (2005) *J. Chromatogr. A*, **1076**(1–2), 193–197.

Liu, Z. and Phillips, J. B. (1991) *J. Chromatogr. Sci.*, **29**(6), 227–231.

Liu, C.-Z., Zhou, H.-Y. and Zhao, Y. (2007) *Anal. Chim. Acta*, **581**(2), 298–302.

Liu, N. Q., Choi, Y. H., Verpoorte, R., *et al.* (2010) *Phytochem. Anal.*, **21**(5), 451–456.

Lo Presti, M., Ragusa, S., Trozzi, A., *et al.* (2005) *J. Sep. Sci.*, **28**(3), 273–280.

Lorenzo, D., Dellacassa, E., Atti-Serafini, L., *et al.* (2000) *Flavour Fragr. J.*, **15**(3), 177–181.

Lorenzo, D., Loayza, I. and Dellacassa, E. (2006) *Flavour Fragr. J.*, **21**(1), 95–98.

Lu, D., Wei, P., Ouyang, P., and Chen, J. (2002) *J. Chin. Pharm. Sci.*, **11**, 26–30.

Lu, Y., Sun, C., Liu, R., *et al.* (2007) *J. Chromatogr. A*, **1146**(1), 125–130.

Lu, D. Y., Cao, Y., Li, L., *et al.* (2011) *J. Pharm. Anal.*, **1**(3), 203–207.

Lucchesi, M. E., Chemat, F. and Smadja, J. (2004a) *Flavour Fragr. J.*, **19**(2), 134–138.

Lucchesi, M. E., Chemat, F. and Smadja, J. (2004b) *J. Chromatogr. A*, **1043**(2), 323–327.

Luo, X., Jiang, Y., Fronczek, F. R., *et al.* (2011) *Pharm. Biol.*, **49**(5), 464–470.

Luque de, C. and Capote, F. P. (2007) Applications of ultrasound-based detection techniques, in *Techniques and Instrumentation in Analytical Chemistry*, eds. C. Luque de and F. P. Capote, Elsevier, Amsterdam, pp. 351–388.

Luque de Castro, M. D. and Garca-Ayuso, L. E. (1998) *Anal. Chim. Acta*, **369**(1–2), 1–10.

Luque de Castro, M. D., Jiménez-Carmona, M. M. and Fernández-Pérez, V. (1999) *TrAC Trends Anal. Chem.*, **18**(11), 708–716.

Ma, Y., Aisha, H. A., Liao, L., *et al.* (2005) *J. Chromatogr. A*, **1076**(1–2), 198–201.

Mahesh, A., Jeyachandran, R., Cindrella, L., *et al.* (2010) *Acta Biol. Hung.*, **61**(2), 175–190.

Mahmoud, S. S. and Croteau, R. B. (2002) *Trends Plant Sci.*, **7**(8), 366–373.

Mamoci, E., Cavoski, I., Simeone, V., *et al.* (2011) *Molecules*, **16**(3), 2609–2625.

Marriott, P. J., Shellie, R. and Cornwell, C. (2001) *J. Chromatogr. A*, **936**(1–2), 1–22.

Marston, A. and Hostettmann, K. (2006) *J. Chromatogr. A*, **1112**(1–2), 181–194.

Martin, V. J. J., Pitera, D. J., Withers, S. T., *et al.* (2003) *Nat. Biotech.*, **21**(7), 796–802.

Martin, A. M., Ferreira Queiroz, E., Marston, A., *et al.* (2006) *Phytochem. Anal.*, **17**(1), 32–35.

Mastovska, K. and Lehotay, S. J. (2006) *J. Agric. Food Chem.*, **54**(19), 7001–7008.

Mauri, P., Minoggio, M., Iemoli, L., *et al.* (2003) *J. Pharm. Biomed. Anal.*, **32**, 633–639.

Megoulas, N. C. and Koupparis, M. A. (2005) *Crit. Rev. Anal. Chem.*, **35**(4), 301–316.

Merfort, I. (2002) *J. Chromatogr. A*, **967**(1), 115–130.

Miliauskas, G., van Beek, T. A., de Waard, P., *et al.* (2006) *J. Chromatogr. A*, **1112**(1–2), 276–284.

Mishra, D., Bisht, G., Mazumdar, P. M., *et al.* (2010) *Pharm. Biol.*, **48**(11), 1297–1301.

Mondello, L., Casilli, A., Tranchida, P. Q., *et al.* (2003) *J. Agric. Food Chem.*, **51**(19), 5602–5606.

Montsko, G., Boros, B., Takatsy, A., *et al.* (2008) *Chromatographia*, **67**(5), 467–470.

Moon, H. I. and Zee, O. (2011) *Immunopharmacol. Immunotoxicol.*, **33**(2), 338–341.

Moreau, R. A., Preisig, C. L. and Osman, S. F. (1992) *Phytochem. Anal.*, **3**(3), 125–128.

Moreira, I. C., Lago, J. H. G., Young, M. C. M., *et al.* (2003) *J. Braz. Chem. Soc.*, **14**, 828–831.

Morin, P. (1994) *Fresenius J. Anal. Chem.*, **348**(4), 327–328.

Morin, P., Caude, M., Richard, H., *et al.* (1986) *J. Chromatogr. A*, **363**(1), 37–56.

Morin, P., Pichard, H., Caude, M., *et al.* (1991) *J. Chromatogr. A*, **464**, 125–137.

Moronkola, D. O., Atewolara-Odule, O. C. and Olubomehin, O. O. (2009) *Afr. J. Pharm. Pharmacol.*, **3**(9), 458–462.

Mosandl, A. and Juchelka, D. (1997) *J. Essent. Oil Res.*, **9**, 5–12.

Mottaleb, M. A. and Sarker, S. D. (2012) Accelerated solvent extraction for natural products isolation, in *Natural Products Isolation*, eds. S. D. Sarker and L. Nahar, Humana Press, New Jersey, pp. 75–87.

Mustafa, A. and Turner, C. (2011) *Anal. Chim. Acta*, **703**(1), 8–18.

Nakamura, M. J., Monteiro, S. S., Bizarri, C. H. B., *et al.* (2010) *Biochem. Syst. Ecols.*, **38**(6), 1170–1175.

Nishii, Y., Yoshida, T. and Tanabe, Y. (1997) *Biosci. Biotechnol. Biochem.*, **61**(3), 547–548.

Nishizawa, M., Inoue, A., Hayashi, Y., *et al.* (1984) *J. Org. Chem.*, **49**(19), 3660–3662.

Nowak, G. (1993) *Chromatographia*, **35**(5), 325–328.

Nowak, G., Dawid-Pac, R., Urbanska, M., *et al.* (2011) *Acta Soc. Bot. Pol.*, **80**(3), 193–196.

Oka, F., Oka, H. and Ito, Y. (1991) *J. Chromatogr. A*, **538**(1), 99–108.

Ong, E. S., Cheong, J. S. and Goh, D. (2006) *J. Chromatogr. A*, **1112**(1–2), 92–102.

Ono, M., Yamashita, M., Mori, K., *et al.* (2008) *Food Sci. Technol. Res.*, **14**(5), 499–508.

Ormeno, E., Goldstein, A. and Niinemets, U. (2011) *Trends Anal. Chem.*, **30**(7), 978–989.

Ottavioli, J., Bighelli, A., Casanova, J., *et al.* (2009) *Spectrosc. Lett.*, **42**(8), 506–512.

Özek, T. and Demirci, F. (2012) Isolation of natural products by preparative gas chromatography, in *Natural Products Isolation*, eds. S. D. Sarker and L. Nahar, Humana Press, New Jersey, pp. 275–300.

Ozel, M. Z. and Kaymaz, H. (2004) *Anal. Bioanal. Chem.*, **379**(7), 1127–1133.

Ozel, M. Z., Gogus, F. and Lewis, A. C. (2003) *Food Chem.*, **82**(3), 381–386.

Palá-Paúl, J., Pérez-Alonso, M. J., Velasco-Negueruela, A., *et al.* (2002) *J. Chromatogr. A*, **947**(2), 327–331.

Pala-Paul, J., Copeland, L. M., Brophy, J. J., *et al.* (2009) *Nat. Prod. Commun.*, **4**(7), 983–986.

Paraschos, S., Magiatis, P., Gousia, P., *et al.* (2011) *Food Chem.*, **129**(3), 907–911.

Peng, C. A., Ferreira, J. F. S. and Wood, A. J. (2006) *J. Chromatogr. A*, **1133**(1–2), 254–258.

Peng, A., Li, R., Hu, J., *et al.* (2008) *J. Chromatogr. A*, **1200**(2), 129–135.

Perales, A., Martinez-Ripoll, M., Fayos, J., *et al.* (1983) *J. Org. Chem.*, **48**(26), 5318–5321.

Pereira, C. and Meireles, M. A. (2010) *Food Bioprocess. Tech.*, **3**(3), 340–372.

Pereira, R. A., MdGB, Z. and MdNdC, B. (2010) *J. Essent. Oil Bear. Plants*, **13**(4), 440–450.

Perry, N. B., Burgess, E. J., Rodriguez Guitian, M. A., *et al.* (2009) *Planta Med.*, **75**(6), 660–666.

Persson, M., Sjodin, K., Borg-Karlson, A.-K., *et al.* (1996) *Phytochemistry*, **42**(5), 1289–1297.

Picman, A. K. (1986) *Biochem. Syst. Ecol.*, **14**(3), 255–281.

Picman, K. A., Ranieri, R. L., Towers, G. H. N., *et al.* (1980) *J. Chromatogr. A*, **189**(2), 187–198.

Pietsch, M. and König, W. A. (1997) *J. High Resolut. Chromatogr.*, **20**(5), 257–260.

Pietsch, M. and König, W. A. (2000) *Phytochem. Anal.*, **11**(2), 99–105.

Piskorski, R., Hanus, R., Vašíčková, S., *et al.* (2007) *J. Chem. Ecol.*, **33**(9), 1787–1794.

Pripdeevech, P., Wongpornchai, S. and Marriott, P. J. (2010) *Phytochem. Anal.*, **21**(2), 163–173.

Qian, G.-P., Yang, Y.-W. and Ren, Q.-L. (2005) *J. Liq. Chromatogr. Relat. Technol.*, **28**(5), 705–712.

Ramírez, P., García-Risco, M. R., Santoyo, S., *et al.* (2006) *J. Pharm. Biomed. Anal.*, **41**(5), 1606–1613.

Ravid, U., Putievsky, E., Katzir, I., *et al.* (1992) *Flavour Fragr. J.*, **7**(4), 235–238.

Ravid, U., Putievsky, E. and Katzir, I. (1994) *Flavour Fragr. J.*, **9**(2), 85–87.

Ravid, U., Putievsky, E., Katzir, I., *et al.* (1997a) *Flavour Fragr. J.*, **12**(4), 293–296.

Ravid, U., Putievsky, E., Katzir, I., *et al.* (1997b) *Flavour Fragr. J.*, **12**(2), 109–112.

Reale, S., Pace, L., Monti, P., *et al.* (2008) *Nat. Prod. Res.*, **22**(4), 360–364.

Reich, E. and Widmer, V. (2009) *Planta Med.*, **75**(7), 711–718.

Reverchon, E. and De Marco, I. (2006) *J. Supercrit. Fluids*, **38**(2), 146–166.

Reznicek, G., Kastner, U. and Glasl, S. (1997) *Phytochem. Anal.*, **8**(1), 9–13.

Roberts, S. C. (2007) *Nat. Chem. Biol.*, **3**(7), 387–395.

Rodrigues, M. R. A., Caramão, E. B., dos Santos, J. G., *et al.* (2002) *J. Agric. Food Chem.*, **51**(2), 453–456.

Rodriguez, E., Towers, G. H. N. and Mitchell, J. C. (1976) *Phytochemistry*, **15**(11), 1573–1580.

Rogerio, A. P., Andrade, E. L., Leite, D. F. P., *et al.* (2009) *Br. J. Pharmacol.*, **158**(4), 1074–1087.

Rout, K., Mishra, S. and Sherma, J. (2009) *Acta Chromatogr.*, **21**(3), 443–452.

Ruzicka, L. (1953) *Cell. Mol. Life Sci.*, **9**(10), 357–367.

Safaralie, A., Fatemi, S. and Sefidkon, F. (2008) *J. Chromatogr. A*, **1180**(1–2), 159–164.

Saini, R., Jaitak, V., Guleria, S., *et al.* (2012) *J. Essent. Oil. Res.*, **24**(3), 315–320.

Santos, A. M., Vasconcelos, T., Mateus, E., *et al.* (2006) *J. Chromatogr. A*, **1105**(1–2), 191–198.

Saritas, Y., von Reu, S. H. and König, W. A. (2002) *Phytochemistry*, **59**(8), 795–803.

Saroglou, V., Dorizas, N., Kypriotakis, Z., *et al.* (2006) *J. Chromatogr. A*, **1104**(1–2), 313–322.

Sayadi, L., Missaoui, I., Jamoussi, B., *et al.* (2010) *Open Chem. Biomed. Meth. J.*, **3**, 18–24.

Scharnhop, H. and Winterhalter, P. (2009) *J. Food Compos. Anal.*, **22**(3), 233–237.

Schultze, W., Lange, G. and Schmaus, G. (1992) *Flavour Fragr. J.*, **7**(2), 55–64.

Schurig, V. (1977) *Angew Chem.*, **89**(2), 113–114.

Seaman, F. (1982) *Bot. Rev.*, **48**(2), 121–594.

Sellier, N., Cazaussus, A., Budzinski, H., *et al.* (1991) *J. Chromatogr. A*, **557**, 451–458.

Senseman, S. A. and Ketchersid, M. L. (2000) *Arch. Environ. Contam. Toxicol.*, **38**(3), 263–267.

Seo, H. Y., Shim, S. L., Ryu, K. Y., *et al.* (2009) *Food Sci. Biotechnol.*, **18**(1), 18–24.

Sessa, R. A., Bennett, M. H., Lewis, M. J., *et al.* (2000) *J. Biol. Chem.*, **275**(35), 26877–26884.

Shinde, D. B., Chavan, M. J. and Wakte, P. S. (2011) HPTLC in herbal drug quantification, in *High-Performance Thin-Layer Chromatography (HPTLC)*, ed. M. Srivastava, Springer, Berlin Heidelberg, pp. 117–139.

Shirota, O., Nagamatsu, K. and Sekita, S. (2006) *J. Nat. Prod.*, **69**(12), 1782–1786.

Shirsath, S. R., Sonawane, S. H. and Gogate, P. R. (2012) *Chem. Eng. Process.*, **53**, 10–23.

Shotipruk, A., Kaufman, P. B. and Wang, H. Y. (2001) *Biotechnol. Prog.*, **17**(5), 924–928.

Silva Elipe, M. V. (2003) *Anal. Chim. Acta*, **497**(1–2), 1–25.

Silva-Brandao, K. L., Solferini, V. N. and Trigo, J. R. (2006) *Biochem. Syst. Ecol.*, **34**(4), 291–302.

Singh, J. (2008) Maceration, percolation and infusion techniques for the extraction of medicinal and aromatic plants, in *Extraction Technologies for Medicinal and Aromatic Plants*, eds. S. S. Handa, S. P. S. Khanuja, G. Longo and D. D. Rakesh, Internation centre for science and high technology ICS-UNIDO, Trieste, pp. 67–82.

Sonsuzer, S., Sahin, S. and Yilmaz, L. (2004) *J. Supercrit. Fluids*, **30**(2), 189–199.

Sonwa, M. M. and König, W. A. (2001) *Phytochemistry*, **58**(5), 799–810.

Sparr Eskilsson, C. and Bjorklund, E. (2000) *J. Chromatogr. A*, **902**(1), 227–250.

Spring, O., Buschmann, H., Vogler, B., *et al.* (1995), *Phytochemistry*, **39**, 609–612.

Spring, O., Zipper, R., Reeb, S., *et al.* (2001) *Phytochemistry*, **57**(2), 267–272.

Spring, O., Zipper, R., Conrad, J., *et al.* (2003) *Phytochemistry*, **62**(8), 1185–1189.

Stashenko, E. E., Jaramillo, B. E. and Martinez, J. R. (2004a) *J. Chromatogr. A*, **1025**(1), 93–103.

Stashenko, E. E., Jaramillo, B. E. and Martinez, J. R. (2004b) *J. Chromatogr. A*, **1025**(1), 105–113.

Sun, Y., Li, W., Fitzloff, J.F., and van Breemen, R.B. (2005) *J. Mass. Spectrom.*, **40**, 373–379.

Sun, Q., Sun, A. and Liu, R. (2006) *J. Liq. Chromatogr. Relat. Technol.*, **29**(1), 113–121.

Sun, A., Zhang, Y., Li, A., *et al.* (2011) *J. Chromatogr. B*, **879**(21), 1899–1904.

Sutherland, I. A. and Fisher, D. (2009) *J. Chromatogr. A*, **1216**(4), 740–753.

Tabanca, N., Demirci, B., Kirimer, N., *et al.* (2005) *J. Chromatogr. A*, **1097**(1–2), 192–198.

Tam, C. U., Yang, F. Q., Zhang, Q. W., *et al.* (2007) *J. Pharm. Biomed. Anal.*, **44**(2), 444–449.

Tang, C., Wei, X. and Yin, C. (2003) *J. Pharm. Biomed. Anal.*, **33**(4), 811–817.

Tang, Q., Yang, C., Ye, W., *et al.* (2008) *Phytochem. Anal.*, **19**(2), 155–159.

Taylor, L. T. (2009) *J. Supercrit. Fluids*, **47**(3), 566–573.

Taylor, L. T. (2010) *Anal. Chem.*, **82**(12), 4925–4935.

Upton, R. T. (2010) *J. AOAC Int.*, **93**(5), 1349–1354.

Van Beek, T. A., Scheeren, H. A., Rantio, T., *et al.* (1991) *J. Chromatogr. A*, **543**, 375–387.

Van Nieuwerburgh, F. C. W., Vande Casteele, S. R. F., Maes, L., *et al.* (2006) *J. Chromatogr. A*, **1118**(2), 180–187.

Verma, R. K., Singh, A. K., Srivastava, P., *et al.* (2009) *J. Liq. Chromatogr. Relat. Technol.*, **32**, 2437–2450.

Vian, M. A., Fernandez, X., Visinoni, F., *et al.* (2008) *J. Chromatogr. A*, **1190**(1–2), 14–17.

Villar, A., Rios, J. L., Simeon, S., *et al.* (1984) *J. Chromatogr. A*, **303**, 306–308.

Vogler, B., Klaiber, I., Roos, G., *et al.* (1998) *J. Nat. Prod.*, **61**(2), 175–178.

Wabo, H. K., Tane, P. and Connolly, J. D. (2006) *Biochem. Syst. Ecol.*, **34**(7), 603–605.

Wagschal, K. C., Pyun, H. J., Coates, R. M., *et al.* (1994) *Arch. Biochem. Biophys.*, **308**(2), 477–487.

Wang, L. and Weller, C. L. (2006) *Trends Food Sci. Technol.*, **17**(6), 300–312.

Wang, X., Liu, Y.-S., Nair, U. B., *et al*. (1997) *Tetrahedron: Asymmetry*, **8**(23), 3977–3984.

Wang, M., Park, C., Wu, Q., *et al*. (2005) *J. Agric. Food Chem.*, **53**(18), 7010–7013.

Wang, Z., Ding, L., Li, T., *et al*. (2006) *J. Chromatogr. A*, **1102**(1–2), 11–17.

Wang, G., Tian, L., Aziz, N., *et al*. (2008) *Plant Physiol.*, **148**(3), 1254–1266.

Wang, S., Jin, D.-Q., Xie, C, *et al*. (2013) *Food Chem.*, **141**, 2075–2082.

Wei, S., Gao, J., Wu, W., *et al*. (2012) *Phytochem. Anal.*, **23**(1), 23–33.

Widmer, V., Handloser, D. and Reich, E. (2007) *J. Liq. Chromatogr. Relat. Technol.*, **30**, 2209–2219.

Widmer, V., Reich, E. and Ankli, A. (2008) Medicines and dietary supplements produced from plants, in *Thin Layer Chromatography in Phytochemistry*, eds. T. Kowalska, J. Sherma and M. Waksmundzka-Hajnos, CRC Press, Boca Raton, pp. 40–42.

Williams, C. M. and Mander, L. N. (2001) *Tetrahedron*, **57**(3), 425–447.

Withers, S. and Keasling, J. (2007) *Appl. Microbiol. Biotechnol.*, **73**(5), 980–990.

Wolfender, J.-L., Ndjoko, K. and Hostettmann, K. (2001) *Phytochem. Anal.*, **12**(1), 2–22.

Wolfgang, P. and Helmut, S. Paul, assignee (1960). Apparatus for separating charged particles of different specific charges. United States.

Wu, H., Su, Z., Yang, Y., *et al*. (2007) *J. Chromatogr. A*, **1176**(1–2), 217–222.

Wu, Y., Zheng, Y., Liu, X., *et al*. (2010) *J. Sep. Sci.*, **33**(8), 1072–1078.

Wu, P., Su, M.-X., Wang, Y., *et al*. (2012) *Phytochemistry*, **76**, 133–140.

Wüst, M., Rexroth, A., Beck, T., *et al*. (1997) *Flavour Fragr. J.*, **12**(6), 381–386.

Xie, J., Wang, S., Sun, B., *et al*. (2008) *J. Liq. Chromatogr. Relat. Technol.*, **31**(17), 2621–2631.

Xie, J., Sun, B., Wang, S., *et al*. (2009) *Food Chem.*, **117**(2), 375–380.

Xie, W.-D., Niu, Y.-F., Lai, P.-X., *et al*. (2010) *Chem. Pharm. Bull.*, **58**(7), 991–994.

Xu, J., Guo, Y., Zhao, P., *et al*. (2012). *Fitoterapia*, **83**, 801–805.

Yamauchi, Y. and Saito, M. (1990) *J. Chromatogr. A*, **505**(1), 237–246.

Yan, J., Chen, G., Tong, S., *et al*. (2005) *J. Chromatogr. A*, **1070**(1–2), 207–210.

Yang, F. Q., Li, S. P., Chen, Y., *et al*. (2005) *J. Pharm. Biomed. Anal.*, **39**(3–4), 552–558.

Yang, F. Q., Wang, Y. T. and Li, S. P. (2006) *J. Chromatogr. A*, **1134**(1–2), 226–231.

Yang, Y., Kayan, B., Bozer, N., *et al*. (2007) *J. Chromatogr. A*, **1152**(1–2), 262–267.

Yang, Y., Gu, D., Yili, A., *et al*. (2009) *Phytochem. Anal.*, **21**(2), 205–209.

Yang, N. Y., Duan, J. A., Shang, E. X., *et al*. (2010) *Phytochem. Anal.*, **21**(2), 144–149.

Yatagai, M., Sato, T. and Takahashi, T. (1985) *Biochem. Syst. Ecol.*, **13**(4), 377–385.

Ye, M. and Guo, D-a. (2005) *Rapid Commun. Mass Spectrom.*, **19**, 818–824.

Yermakov, A., Khlaifat, A., Qutob, H., *et al*. (2010) *Chem. Sci. J.*, **1**, 1–10, doi 10.4172/2150-3494.1000005.

Yoon, K. D., Chin, Y.-W. and Kim, J. (2010) *J. Liq. Chromatogr. Relat. Technol.*, **33**(9–12), 1208–1254.

Yun Shi, S., Ping Zhang, Y., Long Huang, K., *et al*. (2008) *J. Liq. Chromatogr. Relat. Technol.*, **31**(6), 828–837.

Zdero, C. and Bohlmann, F. (1990) *Plant Syst. Evol.*, **171**(1), 1–14.

Zellner, B. A., Bicchi, C., Dugo, P., *et al*. (2008) *Flavour Fragr. J.*, **23**(5), 297–314.

Zhang, Q., Cai, D. and Liu, J. (2011) *J. Chromatogr. B*, **879**(26), 2809–2814.

Zhannan, Y., Shiqiong, L., Quancai, P., *et al*. (2009) *Chromatographia*, **69**(7–8), 785–790.

Zhao, C. and He, C. (2006) *J. Sep. Sci.*, **29**(11), 1630–1636.

Zhong, R., Zhang, Z., Xiao, Z., *et al*. (2006) *Acta Bot. Yunnanica*, **28**(2), 208–214.

Zhou, J. Z., Kou, X. and Stevenson, D. (1999) *J. Agric. Food Chem.*, **47**(3), 1018–1022.

Zhou, X., Li, Z., Liang, G., *et al*. (2007) *J. Pharm. Biomed. Anal.*, **43**(2), 440–444.

Zhou, L., Xu, M., Yang, C.-R., *et al*. (2011) *Helv. Chim. Acta*, **94**(2), 218–223.

Ziémons, E., Goffin, E., Lejeune, R., *et al*. (2005) *J. Supercrit. Fluids*, **33**(1), 53–59.

Zúñiga, B., Guevara-Fefer, P., Herrera, J., *et al*. (2005) *Planta Med.*, **71**(EFirst), 825–828.

Zwenger, S. and Basu, C. (2008) *Biotechnol. Mol. Biol. Rev.*, **3**(1), 1–7.

Analysis of Plant Saponins

Justyna Krzyzanowska, Mariusz Kowalczyk and Wieslaw Oleszek

Institute of Soil Science and Plant Cultivation, State Research Institute, Pulawy, Poland

1   INTRODUCTION

Saponins are naturally occurring compounds, which can be found in different tissues of a large number of plant species belonging to nearly 100 families. They occur predominantly in angiosperms but may also be found in some lower plants, such us ferns and algae as well as in certain marine invertebrates. They can be found in roots, shoots, flowers, and seeds of plants (Hostettmann and Marston, 1995; Oleszek and Marston, 2000).

The formation of soapy lather when shaken in water solution has been the feature of saponins. This property has for a long time been used in identification of saponin-containing plant species as well as for their quantification. The height of the froth, when shaken in a glass tube, and the time of its disappearance was a semiquantitative test. A number of species were misclassified this way as saponin-containing plants because of the fact that some other plant components may also form froth when shaken in water solution. Tests that are more accurate were needed to prove this fact.

In the Orient, plants rich in saponins were used as soap substitute in folk medicine. For this reason, some common names of saponin-rich plant species were derived from this feature, for example, soaproot (*Chlorogalum pomeridianum* Kunth), soapbark (*Quillaja saponaria* Molina), soapwort (*Saponaria officinalis* L.), soapberry (*Sapindus saponaria* L.), soapnut (*Sapindus mukorossi* L.), and soapjacob (*Glinus lotoides* L.). For this reason, saponins found some industrial interest mainly as surface-active or foaming agents (Oleszek and Hamed, 2010). In food industry, saponins were for a long time considered antinutritional compounds because of their throat irritating properties and bitterness. Hence, the food processing was targeted to removal of these undesirable chemicals from the diet. They were also being removed from cultivated species by breeding, for example, low saponin alfalfa varieties. Nowadays, saponins are considered in some cases as health beneficial food components because of their cholesterol-lowering and anticancer properties, for example, soybean, garlic, and onion. Some find commercial application as adjuvants, taste modifiers, emulsifiers, sweeteners, or precursors for hormone synthesis.

Chemically, saponins form a very diverse group of compounds with two basic types of nonsugar aglycone skeletons (sapogenins) – steroidal or triterpene (Figure 1). To the group of saponins, the steroidal glycoalkaloids of solanidans and spirosolan class are also being classified. The aglycones can be substituted with different functional groups (–H, –OH, –COOH, –CH$_3$), which makes great diversity of possible structures present in plants. This diversity can be further multiplied by the number of sugar chains, which can be attached as one chain (monodesmosides), two chains (bidesmosides), or even three chains (tridesmosides) (Greek *desmos* = chain). Also, the composition of sugar chain (pentoses,
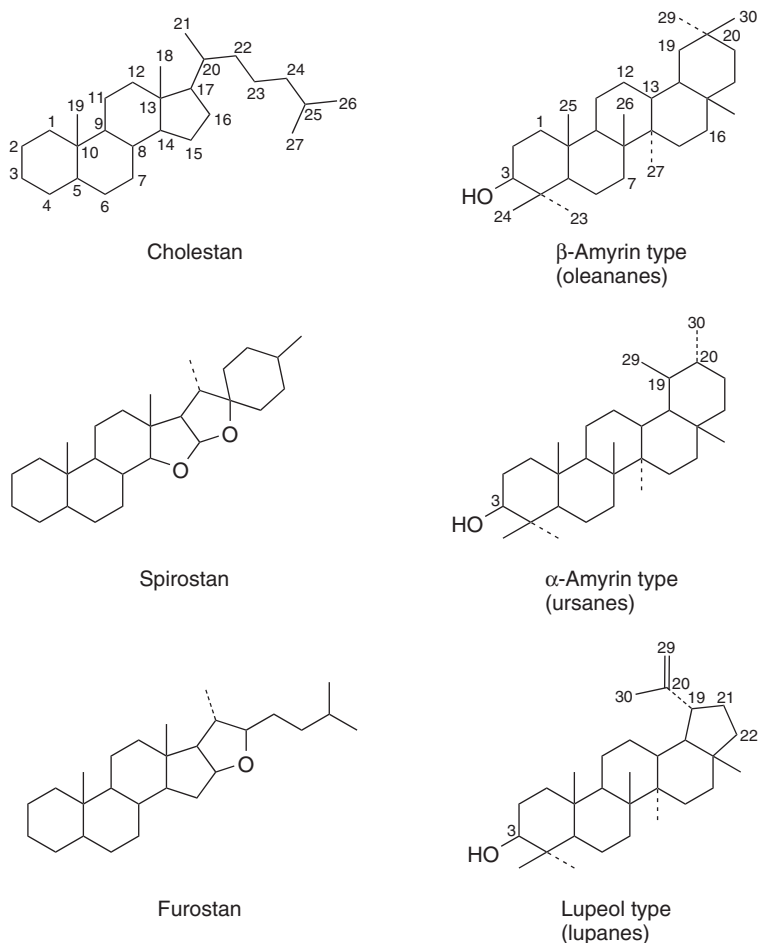
Figure 1 Most common aglycone structures of triterpene and steroidal saponins.

hexoses, deoxyhexoses, and uronic acids) and its shape (linear, branched) can further increase the structural diversity. Some saponin structures can be acylated or acetylated with different chemical groups. Thus, the number of possible structures of these compounds and their diversity in plants is huge. They can rarely be found in plant material as a single component, but rather they occur in multicomponent mixtures, differing even between plant organs (roots, aerial parts, and flowers). In oat plants, for example, both steroidal and triperpene structures can be found depending on the plant part. Thus, the isolation of saponins from such a mixture into single compounds for structure elucidation and biological activity measurements is still a challenge.

2 EXTRACTION AND PURIFICATION

Extraction has been one of the crucial points of saponin analysis. As some saponins can be quite sensitive to numerous extraction parameters (heat, pH, and methanol presence), selection of appropriate conditions is essential for getting unchanged picture of saponin profile. The heat may degrade very sensitive acyl substitutions, whereas pH may influence both acyl and sugar chains. The use of methanol as solvent quite often leads to the formation of $-OCH_3$ derivatives and lactones, which are being identified in plants but are in fact artifacts of the extraction process. Thus, for extraction protocol, cold extraction supported by sonication, with EtOH-water should be strongly recommended.

Early work on saponins included hot extraction of defatted plant samples with aqueous alcohol. This was followed by evaporation of solvent and extraction of saponins from water solution by liquid–liquid extraction with butanol (Wall *et al*., 1952). This protocol has a number of disadvantages. First of all, evaporation of aqueous solution containing saponins has been a very difficult process because of froth formation. Second, liquid–liquid extraction with butanol allows some purification of saponins from highly polar compounds (predominantly polysaccharides), but the resulting extract when condensed produces a dark-brown solid-rich product with many interfering compounds. To further purify this extract from small water-soluble molecules, dialysis can be successfully applied (Massiot *et al*., 1988). Liquid–liquid extraction with BuOH may also not be complete. Some highly polar saponins (bidesmosides and tridesmosides), containing 7–8 sugar components, hardly form BuOH phase and may be completely lost or extraction may not be quantitative. This has been observed in zhanic acid tridesmoside present in alfalfa aerial parts (Oleszek *et al*., 1992). In the Waal's procedure, tridesmoside, in spite of its high concentration in some cultivars, remained in water and could not be identified.

The recommended alternative to liquid–liquid extraction is solid-phase extraction (SPE) using different adsorbents (C18, C8, XAD-resins, etc.) (Oleszek, 1988). In this method, condensed extracts (no more that 20% alcohol content) can be loaded onto water-preconditioned solid phase, which can then be washed with solvents containing increasing concentrations of alcohol in water. Careful selection of the ratio of alcohol : water allows very good purification of saponin fractions. The extract loaded onto C18 support was washed first with water, which removed all polysaccharides. In the next step, the column was washed with 40% MeOH, giving phenolic (phenolic acids, flavonoids, and some isoflavones) fractions and finally, saponins can be washed out with MeOH. It should be emphasized that for saponins of different plant origin, MeOH : water ratio should be carefully established. This can be done by washing the column with gradually increasing the concentration of MeOH in water (5% increase), followed by the analysis of the fraction with thin-layer chromatography (TLC). In fact, this protocol should be preliminarily run for each batch of solid phase and for different extracts,

as some parameters of the extract, for example, pH, may influence SPE purification.

3 THIN-LAYER CHROMATOGRAPHY

Owing to the fact that most of the saponins (excluding avenacins from oats and glycyrrhizin from licorice) do not possess chromophores and cannot be detected with ultraviolet (UV) detectors, TLC has been a very useful technique in saponin analysis (Oleszek, Kapusta, and Stochmal, 2008). This can be successfully used for both quantification and as a supportive technique in the saponin separation with column chromatography (CC). The number of sugars in saponin molecule determines their polarity, which can range widely. Thus, the analysis of these compounds is still a challenge and the selection of TLC adsorbents and solvents is crucial. There is a rather limited choice of ready-to-use TLC adsorbents for saponin analysis. The TLC plates covered with different forms of silica gel are most often used. In some cases, reversed phases C18 and C8 are also of use and are complementary to silica gel plates. The most frequently used solvents for TLC of saponins depend on the adsorbent (Table 1). For silica gel, different proportions of chloroform–methanol–water are most frequently used. Other alcohols, for example, ethanol, isopropanol, and their mixtures, can substantially improve separation. Saponins possessing higher polarity are better separated using acidified solvents, and those having acidic aglycone or uronic acid in the molecule can be successfully separated using ethyl acetate–acetic acid–water as a solvent system.

Best separation of glycoalkaloids can be obtained in ammonified (NH_4OH) solvent system. In some cases, improvement of separation can be obtained by application of ethyl acetate–pyridine–water or diethylamine.

For the separation of aglycones obtained from acid hydrolysis of saponins, most frequently, diisopropyl ether–acetone (75 : 30), benzene–methanol (9 : 1 → 98 : 2), petroleum ether–chloroform–acetic acid (7 : 2 : 1), and hexane–acetone (1 : 4) are the solvents of choice.

3.1 Detection

Most of the saponins do not possess chromophores, allowing their visualization under UV light. Thus, for

Table 1 Adsorbents and solvent systems used for TLC separation of saponins.

Sample	Column	Solvent
Albizia saponins	Silica gel 60 F	$CHCl_3$-MeOH-AcOH-H_2O (15:8:3:2)
		$CHCl_3$-MeOH-H_2O (65:40:1)
Albiziasaponins A–E	Silica gel 60 F	$CHCl_3$-MeOH-H_2O (15:3:1)
	RP-18 WF	MeOH-H_2O (7:3)
Allium saponins	Silica gel	*n*-BuOH-AcOH-H_2O (60:15:25)
Argania saponins	Silica gel	n-BuOH-AcOH-H_2O (12:3:5)
Arnica saponins	Silica gel 60	CH_2Cl_2-EtOAc (9:1)
	RP-18	MeOH
Astragalus saponins	RP-18	MeOH-H_2O (8.5:1:5)
Azadirachta saponins	DC-cards	Petrol-EtOAc (7:3)
	Silica gel GF60	Petroleum ether-EtOAc (6.5:3.5)
Capsicoside E–G	Silica gel	*n*-BuOH-AcOH-H_2O (12:3:5)
Carpolobia saponins	Silica gel 60 F	$CHCl_3$-MeOH-AcOH-H_2O (15:8:3:2)
		CH_2Cl_2-MeOH (19:1)
Chelioclinium saponins	Silica gel	Hexane-EtOAc (4:1) with 0.05% AcOH
Chenopodium saponins	Silica gel	*n*-BuOH-HOAc-H_2O (60:15:25)
Cussonia saponins	Silica gel 60 F	MeOH-CH_2Cl_2-AcOH-H_2O (40:55:3:2)
Cussosaponins A–E	Silica gel	CH_2Cl_2-MeOH-H_2O (17:6:1)
Dammarane saponins	Silica gel 60 F	$CHCl_3$-MeOH-H_2O (65:35:6)
Dioscorea saponins	Silica gel	$CHCl_3$-MeOH-H_2O (13:7:1)
		$CHCl_3$-MeOH (9:1)
Draconins A–C	Silica gel 60 F	*n*-hexane-AcOEt (20:1, 7:3)
		Toluene-PrOH (20:1)
		$CHCl_3$-MeOH (10:1)
Eranthisaponins A and B	Silica gel 60 F	$CHCl_3$-MeOH-H_2O (14:8:1)
	RP-18	MeCN-H_2O (2:5, 4:1, 1:3)
Eucalyptus saponins	Silica gel 60 F	*n*-Hexane-CH_2Cl_2-EtOAc (16:16:1)
Fomefficinic acid A–E	Silica gel	$CHCl_3$-MeOH (95:5)
Gambeya saponins	Silica gel	$CHCl_3$-MeOH-H_2O (8:5:1)
Glinus saponins	Silica gel 60 F	MeOH-H_2O (4:1), $CHCl_3$-MeOH-H_2O (70:50:4)
	RP-18	MeOH-H_2O (4:1)
	Silica gel Sil G-100	$CHCl_3$-MeOH-H_2O (70:50:4)
Harpulia saponins	Silica gel 60 F	$CHCl_3$-MeOH-H_2O (12:8:1)
		$CHCl_3$-MeOH-HCOOH (65:35:1)
Ilex saponins	Silica gel	*n*-BuOH-AcOH-H_2O (65:15:25)
		$CHCl_3$-MeOH-H_2O (70:30:3)
		$CHCl_3$-MeOH-*n*-PrOH-H_2O (5:6:5:1:4) upper phase
Lotoidesides A–F	Silica gel 60 F	$CHCl_3$-MeOH-H_2O (13:7:1, 28:12:1)
Lupinus saponins	Silica gel	Hexane-Me_2CO (75:25, 60:40, 40:60, 10:90)
Lyciantosides A–C	Silica gel	*n*-BuOH-AcOH-H_2O (60:15:25)
		$CHCl_3$-MeOH-H_2O (7:3:0.3)
Maytenus saponins	Silica gel	100% CH_3CN
Medicago saponins	Silica gel	EtOAc-AcOH-H_2O (7:2:2)
Meryta saponins	Silica gel 60 F	$CHCl_3$-MeOH-H_2O-EtOAc (28:35:5:32)
Soyasaponin	Silica gel	$CHCl_3$-MeOH-H_2O (65:35:10) lower phase
Panax saponins	Silica gel 60	$CHCl_3$-MeOH-H_2O (13:7:1)
Panax saponins	Silica gel 60 F	$CHCl_3$-MeOH-H_2O (10:3:1)
	RP-18	MeOH-H_2O (7:3)
Pastuchosides A–E	Silica gel	$CHCl_3$-MeOH-H_2O (26:14:3)
		n-BuOH-AcOH-H_2O (4:1:5)
		$CHCl_3$-MeOH (20:1)
		CH_2Cl_2-MeOH-H_2O (50:25:5)
Pittoviridoside	Silica gel	*n*-BuOH-EtOH-H_2O (5:1:4)
Pittoviridoside	C-18	70% MeOH
	Silica gel	EtOAc-EtOH-H_2O (7:2:1, 7:10:1)

Table 1 (*Continued*)

Sample	Column	Solvent
Silene saponins	Silica gel	$CHCl_3$-MeOH-AcOH-H_2O (15:8:3:2)
Solanum saponins	C-18	MeCN-H_2O (7:3)
Symplocos saponins	Silica gel 60 F	$CHCl_3$-MeOH (9:1)
Solanum spp.	Silica gel F	MeOH or EtOH-$CHCl_3$ (2:1)
		$CHCl_3$-MeOH-2%NH_4OH (70:30:5)
		Pyridine-EtOAc-H_2O (1:3:3) (upper)
		n-BuOH-Me_2CO-H_2O (4:5:1)
		Me_2CO-MeOH (3:5)
		EtOH-HOAc-H_2O (19:1:1)
Tribulus saponins	Silica gel 60 F	n-BuOH-AcOH-H_2O (60:15:25)
		$CHCl_3$-MeOH-H_2O (40:9:1)
Tripterygium saponins	Silica gel 60 F	Hexane-EtOH (2:1)
		$CHCl_3$-hexane (8:2)
Tuberosides N–U	Silica gel	n-BuOH-AcOH-H_2O (4:5:1)
		$CHCl_3$-MeOH-H_2O (7:3:0.5)

Source: W. Oleszek, *et al.* (2006). Reproduced from Elsevier.

their visualization, spray reagents are necessary. Till date, over 50 reagents have been proposed, of which most often used are the following sprays:

Anisaldehyde-Sulfuric Acid. Mix 0.5 mL of anisaldehyde with 10 mL of glacial AcOH and add 85 mL of MeOH and (carefully with stirring) 5 mL of H_2SO_4.

Antimony(III)chloride. Prepare 20% solution of $SbCl_3$ in $CHCl_3$.

Chloramine-Trichloroacetic Acid. Mix 10 mL of freshly prepared 3% aqueous chloramine T solution (synonyms sodium sulfamid chloride or sodium tosylchloramid) with 40 mL of 25% ethanolic trichloroacetic acid.

Ehrlich Reagent. Mix 1 g of *p*-dimethylaminobenzaldehyde with 50 mL of 36% HCl and 50 mL of EtOH.

Kedde Reagent. Mix 5 mL of the freshly prepared 3% ethanolic 3,5-dinitrobenzoic acid and 5 mL of 2 N NaOH.

Komarovsky Reagent. Mix 1 mL of 50% ethanolic H_2SO_4 with 10 mL of 2% methanolic 4-hydroxybenzaldehyde (reagent should be prepared shortly before use).

Liebermann–Burchard Reagent. Carefully add 5 mL of acetic anhydride and then 5 mL of concentrated H_2SO_4 to 50 mL of absolute EtOH, while cooling in ice.

Vaniline-Phosphoric Acid Reagent. One gram of vanillin dissolved in 100 mL of 50% phosphoric acid.

Vaniline-Sulfuric Acid Reagent. Solution I: 5% ethanolic H_2SO_4; solution II: 1% ethanolic vanillin. Spray the plate with solution I and then immediately with solution II.

Most of above-listed reagents (excluding Kedde reagent) should be sprayed on hot plates after being carefully dried from the solvents, heated in the oven at 105–110°C. After spraying, the plates need heating again for 5–10 min for color development (Table 2). Plates should not be overheated or kept too long in the oven after spraying as they may produce dark spots instead of colors. Most of the spray reagents produce color spots, which are not very stable, which might be important if documentation of separation is being made. Most stable are the vanillin-based sprayers.

The Ehrlich reagent has been very useful for distinction between furostanol and spirostanol steroidal saponins; furostanols produce pink-red spots, whereas spirostanols are not visualized. The reagents containing aromatic aldehydes, for example, anisaldehyde and vanillin, give strong color products with aglycones and that is why they can be used both for TLC and for colorimetric procedures.

The saponins separated on TLC plates can also be visualized with blood reagent. Some saponins show hemolytic activity as they form complexes with the walls of the red blood cell, causing their disruption. This effect appears both in the saponin solution and on TLC plates. Thus, for preparing blood reagents, 10 mL of 3.6% sodium citrate is mixed with 90 mL of the fresh bovine blood. Two milliliters of this

Table 2 Color reaction of saponins with spray reagents.

Reagent	Color of spots
Triterpenes and steroids	
Liebermann–Burchard	Blue, green, pink, brown, yellowish visible light also under UV
Vanillin-sulfuric acid	Blue, blue-violet, yellowish
Vanilin-phosphoric acid	Red-violet in visible light and reddish or blue in UV 365
Anisaldehyde-sulfuric acid	Blue red-violet in visible light and reddish or blue in UV 365
Komarowsky reagent	Blue, yellow, red
Antimony (III) chloride (Carr–Price reagent)	Pink, purple in visible light, red-violet, green, blue in UV 365
Ehrlich reagent	Red coloration of furostanol derivatives, spirostanols do not react
Water	Sterols give white spots
Blood reagent	White zones on the reddish background
Cucurbitacines	
Liebermann–Burchard	Blue, green, pink, brown, yellowish visible light also under UV
Vanillin-sulfuric acid	Blue, blue-violet, yellowish
Cardiac glycosides	
Kedde reagent	Blue to red-violet in visible light
Chloramine-trichloroacetic acid	Blue, blue-green, yellow-green in UV 365
Vanillin-sulfuric acid	Blue, blue-violet, yellowish

solution is mixed with 30 mL of gelatin dissolved in phosphate buffer of pH 7.4 (4.5 g per 100 mL). Developed TLC plates need to be carefully dried from the solvent residue and then covered with a layer of gelatin/blood solution. After a few hours, white spots can be seen on the plates, indicating the presence of saponins. Care has to be taken as some hydrophobic compounds prevent plates from wetting properly, and in fact, these areas can be mistakenly identified as hemolytic spots.

3.2 TLC – Densitometry

TLC in one-dimensional (1D) and two-dimensional (2D) modes is a powerful technique, which has been used successfully in the separation and determination of a large number of saponins in plant extracts. There are two major problems occurring with the application of these techniques. First, it is necessary to run in parallel the appropriate standards, which minimize the variation between different plates and color reactions with spraying reagents. Second, the spot detection must be done by means of sophisticated instrumentation for data acquisition and handling to scan the whole plate surface at high speed. The online coupling of a computer with a dual-wavelength flying-spot scanner and 2D analytical software is the method of choice. The developing systems and the spray reagent used in this protocol are the same as listed earlier. Linear relationships

between the peak area and the amount of standard saponins can be found in the range 1–5 µg per spot with recovery being at the level of 98% and standard deviation of around 3–5%. The method is sufficiently accurate for quality control and is particularly suitable for assays in series. This method enables determination of a large number of samples, most often does not require any tedious cleanup steps before analysis, and is highly recommended in pharmaceutical quality control practice.

3.3 TLC – Colorimetry

The quantitative analysis of saponins can be performed with a combination of TLC and colorimetry. TLC in this mode can play a dual role. It can be used as a means of confirmation of saponin presence in the extract when colorimetry is being used for saponin quantitation in a crude plant extract. Alternatively, TLC can be used for separation of saponins, followed by scraping TLC bands, their extraction with alcohol followed by colorimetry by treatment with specific reagents. Most frequently used colorants include Ehrlich or vanillin reagents and measurements are made at $\lambda = 515$–560 nm. The major disadvantage of this procedure is the fact that some other components of the extract matrix, such as sterols and bile acids with hydroxyl group at C3, may give a color reaction with the reagent, providing misleading information. An anisaldehyde–sulfuric acid–ethyl

acetate reagent gives a color reaction with steroidal sapogenins, which is in general free of any influence from interfering compounds. To partially avoid these problems, preliminary purification of saponins from other components of extract matrix is highly recommended.

The second major disadvantage of using silica gel TLC before colorimetric determination is irreversible absorption of a considerable portion of the saponins in the stationary phase and the final quantities present in the plant may be underestimated. To use this procedure for routine analysis, it first has to be calibrated against a more sophisticated technique, for example, high performance liquid chromatography (HPLC).

3.4 TLC – Biological Activity Measurement

Owing to the tedious isolation and identification procedures used for separation of individual saponins, there is a need to introduce fast screening preliminary tests (Marston, 2011). Most of these tests involve TLC bioautography, based on localization of antibacterial or antifungal activity on chromatograms. Early methods of bioautography were based on the so-called agar-diffusion techniques in which active compounds were transferred from TLC layer to an inoculated agar plate through a diffusion process during the contact of these two layers. In this way, antibacterial, antifungal, antiprotozoal, antiphage, phage-inducing, antiviral, and cytotoxic activities of different saponins could be preliminarily measured, as indicators for further activity-guided isolation. The major disadvantage of the agar-diffusion technique was differential diffusion of compounds from the chromatogram to the agar plate because of their structure-dependent water solubility.

This technique can be further improved by the direct bioautographic detection on the TLC plate. In this test, the TLC plates (20×20 cm) are being developed in appropriate solvent system, dried, and sprayed (18 mL per plate) with cultures of bacteria or fungi suspended in TSB (trypticase-soy-broth). After spraying, the plates are incubated in a humid environment at 25°C overnight, dried, and sprayed again with the following solutions: (i) aesculin spray (aesculin 0.2% w/v; ammonium ferric citrate 0.1% w/v; yeast extract 0.5% w/v in distilled water) or (ii) tetrazolium salt (2,3,5-triphenyltetrazolium chloride 20 mg mL^{-1}). On the plates sprayed with the bacteria and subsequently with an aesculin spray, the hydrolysis of aesculin resulted in the development of a brown color, while the zones of inhibition of the bacteria remained colorless. On the plates sprayed with tetrazolium salt, the metabolically active bacteria convert the tetrazolium salt into an intensively colored formazan and the antibacterial compounds appear as clear spots against a colored background.

The method can be further simplified for testing antibacterial activity against, for example, *Bacillus subtilis* and *Escherichia coli*. In these methods, the TLC plates were developed in solvent systems, thoroughly dried, and a suspension of microorganisms in a suitable broth was applied to the TLC surface. Incubation in a humid environment permits growth of the bacteria. The activity of the compound is determined by measuring the zones of inhibition visualized by appropriate reagents. In case of *B. subtilis* and *E. coli*, visualization can be obtained by spraying with tetrazolium salt.

Modification of biological activity measurement and semiquantitation of saponins can be done by hemolytic micromethod. The cow's blood is stabilized with sodium citrate (3.65% w/v) and mixed with gelatin solution. For this, gelatin (4.5 g) is dissolved in 100 mL of isotonic buffered solution and 75 mL of this is mixed with 20 mL of blood. Gelatin/blood mixture is spread on a glass plate (10×20 cm) to a thickness of 0.5 mm, and after coagulation, the plates are used for tests. Saponin samples (10 μL) or mashed plant material is placed in localized areas on the gelatin/blood-covered plates, and after 20 h, the widths of the resulting hemolytic rings are measured. A ring of standard saponin is measured in parallel on each plate (Figure 2).
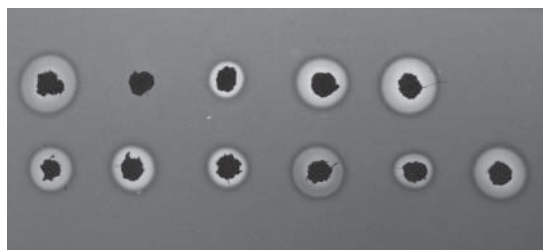
Figure 2 Hemolytic micromethod for saponin occurrence/concentration in plant materials.

3.5 TLC as Supporting Technique in Saponin Isolation

The separation of any crude mixture containing a number of compounds of similar polarity is still quite a challenge. In the case of saponins, this is even more difficult, as these compounds do not possess chromophores and the separation cannot be monitored with the UV detector. Some other kind of detectors, for example, refractive index (RI) or light scattering detector (LSD) can be effectively used. But in many cases, they do not perform satisfactorily. The TLC in many cases can be the detection method of choice. The separation with low pressure CC can be successfully monitored with TLC.

3.6 TLC in Sugar and Aglycone Component Analysis

Some preliminary information on saponin structure can be obtained by classical hydrolysis of saponins followed by TLC analysis of hydrolysis products. Different aqueous solutions of sulfuric or hydrochloric acids are routinely used for hydrolysis. The hydrolysis is performed at high temperature (boiling under reflux or heating in closed ampoules) for different periods of time. Some saponins hydrolyze readily after 10–20 min, whereas others need prolonged heating to complete the process. Prolonged heating has some disadvantages, as some saponins when hydrolyzed produce no single aglycone but a mixture of aglycones, most of which are artifacts. This is the case when soybean saponins are hydrolyzed and soyasapogenols B, C, D, E, and F are being formed.

The hydrolysis mixture can be analyzed with TLC for aglycone structure and sugar composition. For that, sugars need to be separated from aglycones by liquid/liquid extraction with appropriate solvents. Usually, chloroform or ethyl acetate removes aglycones, whereas sugars stay in water solution. The sugars and aglycones can also be easily separated by SPE on C18 cartridges. Sugars pass readily through water-preconditioned cartridge, whereas aglycones bind C18 and can be removed with MeOH.

When standards are available, the 1D TLC performed in several solvent systems and different sprayers may provide the preliminary information on the chemical nature of compounds present in the extract. The aglycones and sugars can be compared

(a)   (b)
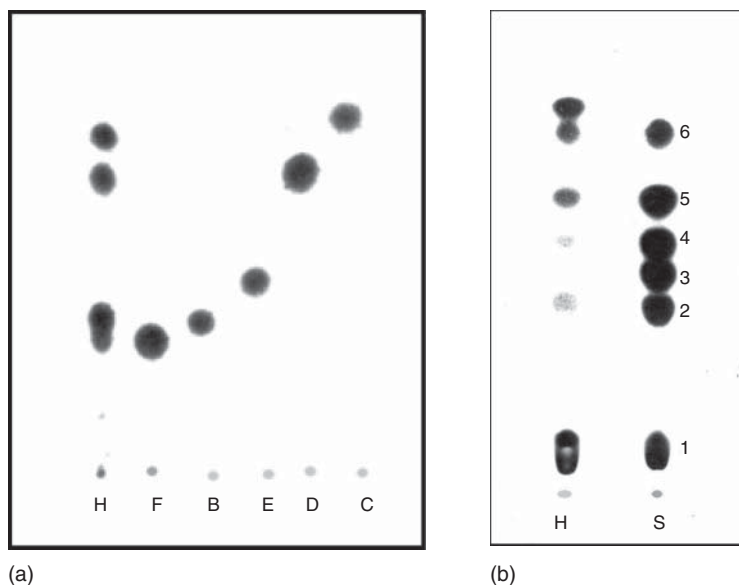
Figure 3   TLC of the hydrolysis products of soybean saponins. (a) Soyasapogenol B, C, D, E, and F standards, and H saponin hydrolysate; solvent system: petroleum ether–chloroform–acetic acid (7 : 2 : 1), sprayer Liebermann–Burchard. (b) Sugar standards S (1, glucuronic acid; 2, galactose; 3, glucose; 4, arabinose; 5, xylose; and 6, rhamnose); solvent system benzene–n-butanol–pyridine–water (1 : 5 : 3 : 3), sprayer: solution I – 20 mL acetone + 0.1 mL of saturated silver nitrate, after drying solution II – 0.5 N NaOH in EtOH and then solution III – 10% aqueous $Na_2S_2O_3$.

by TLC with reference standards. Figure 3 shows the TLC analysis of the aglycones and sugars of saponins present in *Trifolium* spp. The color reaction with different sprayers may also provide additional information on aglycone structure.

4 LOW PRESSURE COLUMN CHROMATOGRAPHY FOR SAPONIN ISOLATION

4.1 Open Column Chromatography

Saponins occur in plant materials most often not as a single compound but rather as a mixture of a number of glycosides differing in polarities. They may possess from one to nine sugar components, which can be additionally acylated with different chemical groups. Thus, the separation of a single compound from such a diverse mixture for structure elucidation or biological activity studies is still a challenge. This can be performed by the combination of low pressure chromatography on different stationary phases. For this purpose, proper careful selection of stationary and mobile phases has been essential for successful work. There is still limited number of possibilities regarding stationary phases. Most frequently, the first step of crude extract separation employed Sephadex LH-20 molecular filtration, which allowed preliminary separation of a complex matrix of the extract into saponins and accompanying impurities. This preliminary separation may result in partition of saponin mixtures into subfractions containing few compounds. Water solution of saponins can also be passed through a porous polymer gel column Diaion HP-20. Washing this column with water removes some impurities and adsorbed saponins, which are being eluted with MeOH. Plant extract could also be separated to different classes of phytochemicals by selective SPE/fractionation on RP-18 and RP-8 short column.

Purified saponins can be further separated using CC on silica gel with mobile phase composed of different combinations of $CHCl_3$–MeOH–H_2O and EtOAc–MeOH–H_2O or on reversed-phase silica gel RP-18, with mobile phase composed of MeOH–H_2O or MeCN–H_2O and their acidified (AcOH, TFA) modifications. These separations rarely result in a single compound. Most often a combination of normal followed by reversed-phase separation is necessary. Final purification can also be accomplished

using HPLC systems with normal or reversed-phase columns, but this procedure is rather tedious and does not allow obtaining bigger amounts of pure saponins.

It is rather much more difficult to separate steroidal saponins. For this separation, a combination of CC with solid phases of different polarities should be combined. On C18 solid phase, quite often, a number of different polarity saponins can be run together. Thus, separation using more polar stationary phases is required in this case (Oleszek *et al.*, 2001).

Higher quantities of separated compounds could be obtained from high speed counter-current chromatography, the system working without any solid support, with separation based on fast partitioning effects of the analytes between two immiscible liquid phases. With this technique, irreversible absorption effects and formation of artifacts were minimal (Marston and Hostettmann, 2006).

4.2 Preparative Low Pressure Chromatography

Separation of saponins of similar molecular weights and polarity cannot be performed on open columns where solid-phase grain ranges between 25 and 65 μm to allow unforced solvent flow. For the separation of such saponin mixtures, an analytical grade solid phase (5–10 μm) has to be applied. Such a small grain requires forced solvent flow, which can be obtained using preparative chromatographs. As presented in Figure 4, good separation of five saponins could be obtained on the semipreparative C18 column (250 × 10 mm, 5 μm) using acetonitrile in 0.1% formic acid. The shape of the solvent gradient must be carefully selected for a particular type of saponins. In such a system, injection volume of 0.2 mL may contain up to 15 mg of saponins. The detection can be performed with LSD detector to get a satisfactory response factor. The width of the saponin peak as well as complete separation allows precise setting of the collection windows and the process can run automatically. After several runs, a few micrograms of each saponin could be obtained. Their purity was high enough to run mass and nuclear magnetic resonance (NMR) spectra for structure elucidation.

However, it is obvious that such a separation should be performed after careful cleanup of saponin fraction. As the longevity of the semipreparative column strongly depends on the purity of loaded matrix, before loading to the column, the saponin fraction
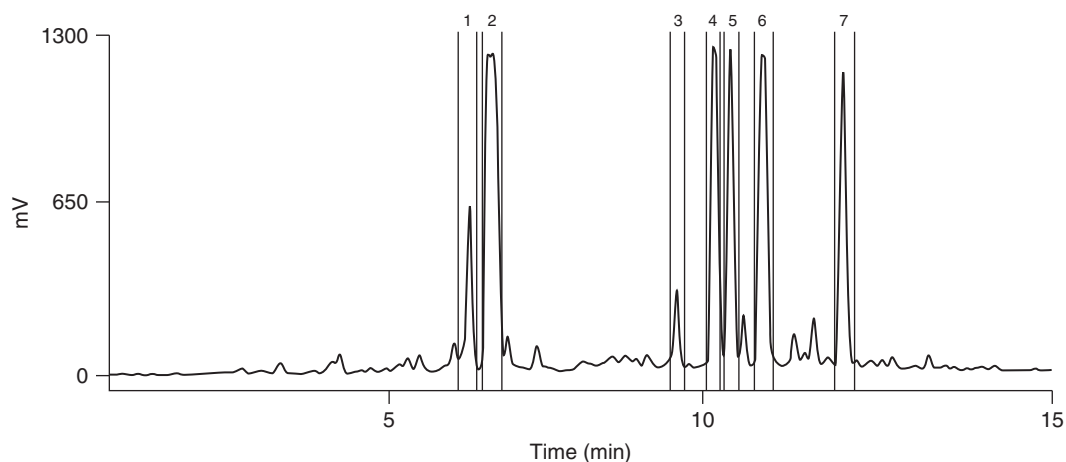
Figure 4   Preparative separation of seven saponins from *Saponaria officinalis* L. Column: Kromasil 100-5-C18, 250 mm × 10 mm, 5 μm; detector prepELS II GILSON; flow rate 5 mL min^{-1}; injection volume – 0.2 mL with a sample concentration of 60 mg mL^{-1}.

should be preliminary purified with different kinds of SPE. It is recommended that first purification is performed on the same solid phase as appears on the semipreparative column. Thus, fractionation on cheaper C18 flesh chromatography is a method of choice. Most of the components of the saponin matrix that bind to C18 will remain on the first sorbent (see the SPE method mentioned earlier). Saponin fraction from the SPE purification can be further cleaned on Sephadex LH-20 before the final semipreparative separation. The LH20 column should be stabilized with aqueous MeOH (e.g., 40% MeOH depending on saponin type) and saponin fraction is loaded onto the column. The column is washed with, for example, 40% MeOH to obtain several saponin subfractions. These subfractions are usually highly purified and do not interfere with solid phase of semipreparative column giving satisfactory separations.

5   HIGH PERFORMANCE LIQUID CHROMATOGRAPHY (HPLC)

HPLC is an analytical method commonly used for the identification and determination of the saponins. Many protocols and detection techniques are elaborated. This technique is rapid, selective, and highly sensitive. Although HPLC may be considered to be a mature technology, steady progress improves the separation capability and productivity of the method and sample analysis. Moreover, by combining it with other techniques such as mass spectrometry (MS) or NMR spectroscopy, the saponin structure can be fully elucidated.

Normal- and reverse-phase HPLC columns have been successfully used for the separation of saponins. Silica gel columns were applied for the separation of soyasaponins (Ireland and Dziedzic, 1986), soybean sapogenins (Ireland and Dziedzic, 1985), and ginsenosides (Besso *et al*., 1979; Sticher and Soldati, 1979). Both C8 and C18 reversed-phase columns were employed; however, C18 has been commonly used. The C18 was applied for numerous analysis (Table 3). However, the separation of sapogenins from Agave (Higgins, 1976) and avenacosides from *Avena sativa* L. (Kesselmeier and Strack, 1981) were performed on C8 columns. The separation of saponins can also be achieved using other stationary phases. Selective carbohydrate, borate anion-exchange, and hydroxyapatite [Ca(PO)(OH)] supports have been used. Carbohydrate (Bushway and Storch, 1982) and NH-modified (Saito *et al*., 1990) columns have been shown to be very effective in the separation of glycoalkaloids. The NH-modified column was also used for the separation of steroid saponins from *Dioscorea* (Xu and Lin, 1985). Octadecylsilylated silica (ODS) column were applied for the separation of major saponin components from *Dizygotheca kerchoveana* Hort & Veitch (Melek *et al*., 2004). Moreover, oleane saponins, from *Sanicula elata* Ham. var *chinensis* Makino (Matsushita *et al*., 2004a) and similarly those from *Hydrocotyle sibthorpioides* Lam. (Matsushita

Table 3 The high performance liquid chromatographic separation of saponins.

Saponin	Column	Solvent system
Acutangulosides A–F	RP-18, phenyl	MeOH-H$_2$O
Aesculiside A	ODS YGW C18	MeCN-H$_2$O-HOAc
Albiziasaponins A–E	YMC-Pack ODS-A	MeCN-H$_2$O-HOAc
Asterosaponins	RP-18	MeOH-H$_2$O
Avicins D and G	Intersil RP-18	MeOH-H$_2$O
	Flurosep-RP-phenyl	MeCN-H$_2$O
Barringtogenol C	C-18 201SP510	MeCN-H$_2$O-TFA
Barringtonia saponins	YMC-Pack ODS-AQ	MeOH-H$_2$O
Boswellic acid	RP-18	MeCN
Capsicosides E–G	μ-Bondapak C-18	MeOH-H$_2$O
Celtis triterpenes	YMC J'sphere ODS-H80	MeCN-H$_2$O
Certonardosides A–J	YMC-Pack ODS	MeOH-H$_2$O
	YMC-Pak ODS	MeOH-H$_2$O
	C18-5E Shodex	MeOH-H$_2$O
	YMS-Pak C-8	MeOH-H$_2$O
Conyzasaponins I–Q	Pegasil ODS-II	MeOH-H$_2$O
		MeCN-H$_2$O
Chenopodium saponins	μ-Bondapak C-18	MeOH-H$_2$O
Cucumariosides	Silasorb C-18	CH$_3$COCH$_3$-H$_2$O
Dioscorea saponins	Cosmogel C-18	MeOH-H$_2$O
Echinocystic acid glc.	ODS	MeCN-H$_2$O-TFA
Elburzensosides	μ-Bondapak C-18	MeOH-H$_2$O
Ficus triterpenes	Silica gel	EtOAc-hexane
Flos saponins	Zorbax SB-C18	MeCN-H$_2$O-OHAc
Ginsenosides	YMC-Pack ODS-AQ	MeCN-H$_2$O
	YMC-Pack ODS-AQ303	MeCN-10 mM K-phosphate buffer
	Synergi Hydro-RP	MeCN-H$_2$O
Glinus saponins	RP-18	MeCN-H$_2$O
Hederagenin saponins	Sigel	CHCl$_3$-MeOH
	YMC R&D ODS	MeOH-H$_2$O-TFA
Helleborus saponins	μ-Bondapak C-18	MeOH-H$_2$O
Hydrocotylosides I–VII	ODS	MeCN-H$_2$O
	Capcell Pak Ph	MeCN-H$_2$O-TFA
Hydroxyimberic acid saponins	PrepLC	MeCN-H$_2$O-TFA
Ilex saponins	μ-Bondapak C-18	MeOH-H$_2$O
Jenisseensosides A–C	Lichrospher RP-18	MeCN-H$_2$O-TFA
Jujubogenins	μ-Bondapak RP-18	MeOH-H$_2$O
Justicosides A–D	YMC ODS	MeCN-H$_2$O
Lotoidesides A–F	XTerra C-18	MeOH-H$_2$O-TFA
Lycianosides A–C	μ-Bondapak C-18	MeOH-H$_2$O
Lupanes	Kromasil Sil	Cyclohexane-EtOAc
Lupinus saponins	Reliasil C-18	MeCN-H$_2$O-TFA
Maesa saponins	Hypersil BDS C18	NH4OAc-MeOH-MeCN-H$_2$O
Medicagenic acid saponins	XTerra RP-18	MeCN-MeOH-H$_2$O
Medicago saponins	RP-18	MeCN-H$_2$O-HOAc
	RP-18	MeCN-H$_2$O-HOAc
Mimengosides C–G	Pegasil OGS-II	MeOH-H$_2$O
Morolic acid	RP-18 201SP	MeCN-H$_2$O-TFA
Nepheliosides I–VI	ODS-AQ	MeOH-H$_2$O
Notoginsenosides	YMC-Pak ODS-A	MeCN-H$_2$O-HOAc
Oenothera triterpenoids	Lichrosorb Diol	*n*-Hexane-EtOAc
Oleanolic acid saponins	Develosil PhA	MeCN-H$_2$O-TFA
	RP-18 201SP	MeCN-H$_2$O-TFA
	YMC R&D ODS	MeOH-H$_2$O-TFA
Pachyelasides A–D	Asahipack GS-320	MeOH

Table 3   (*Continued*)

Saponin	Column	Solvent system
Phytolaccagenic acid glc.	STR Prep-ODS 20	MeOH-H$_2$O
Pittoviridoside	YMC ODS-A	MeOH-tetrahydrofuran-H$_2$O-HOAc
Protoaescigenin saponins	C-18 201SP510	MeCN-H$_2$O-TFA
Protobassic acid saponins	ODS Zorbax	MeCN-H$_2$O
Pulsatilla saponins	Spherisorb ODS 2	MeCN-H$_2$O
Quinonemethide triterpenoids	RP-18	MeOH-H$_2$O-H$_3$PO$_4$
Saniculasaponins I–XI	ODS	MeCN-H$_2$O
Scabiosaponins A–K	Pegasil ODS	MeOH-H$_2$O
Serianic acid saponins	STR Prep-ODS 20	MeOH-H$_2$O
Soyasapogenols A and B	ODS RP-18	MeCN-PrOH-H$_2$O-HOAc
Soyasaponins	Lichroprep RP-18	MeCN-H$_2$O-TFA
	μ-Bondapak RP-18	MeOH-isoPrOH-H$_2$O-HCOOH
	Zorbax eclipse XDB-C18	MeCN-H$_2$O
	Aquasil RP-18	MeCN-H$_2$O-HOAc
Steroidal saponins	Lichrospher C-18	MeCN-H$_2$O
	Lichrospher C-18	MeCN-H$_2$O
Symplocososides	YMC-pack ODS-A	MeOH-H$_2$O
		MeCN-H$_2$O
Synallactosides	Silasorb C-18	EtOH-H$_2$O
Ternstroemiasides A–F	YMC ODS-H80	MeOH-H$_2$O
Tripterygium saponins	YMC-park SIL-06	Hexane-EtOAc
Triquetrosides	RP-18	MeOH-H$_2$O
Tropeosides A and B	RP-18	MeOH-H$_2$O
Ursolic acid	Nova-Pak RP-18	MeOH-H$_2$O-HOAc
Zafaral	RP-18	MeCN-H$_2$O

et al., 2004b) were also successfully separated on the ODS column.

The HPLC separation and detection of saponins could be affected using a variety of mobile phases and additives. The mixtures of water and methanol or water and acetonitrile are the most common in case of reverse-phase saponin chromatography. However, the commonly used methanol hinders the UV detection of saponins because of the overlap in their signals. This follows from the fact that saponins exhibit absorption only at low UV wavelengths, close to the border of UV wavelengths absorbed by methanol. The mobile phase consisting of methanol and water was used for the determination of various saponins. Equally often, linear gradient of acetonitrile and water has been employed. While in case of normal-phase gradient of hexane and isopropanol and hexane and ethyl acetate, solvent mixtures consisting of *n*-heptane : butanol : acetonitrile : water and chloromethane : methanol : water were applied for the separation of various saponins (Table 3).

Moreover, small amounts of trifluoroacetic acid, acetic acid, or formic acid are used as a pH modifier. Trifluoroacetic acid may be a bad choice in the case of HPLC-MS analysis.

The most common flow rate for the HPLC analysis of saponins is $1\,\mathrm{mL\,min^{-1}}$. Typically, time of 30–70 min gives satisfactory separation of these compounds.

5.1   UV Detection

The lack of a chromophore in the molecules of most saponins hampers their detection in UV light. Nonspecific detection at 200–210 nm is possible. But at this wavelength other components of the sample may overlap with the saponins, making their determination difficult. HPLC-UV method results in both a low sensitivity of detection and limiting the choice of solvents and the use of mobile phase modifiers required for improved separation. Despite this, numerous published data are based on recording HPLC profiles at 200–210 nm by applying mobile phase consisting of aqueous methanol. However, the mobile phase consisting of acetonitrile and water is a better choice, because acetonitrile has much lower absorption at lower wavelengths when compared with methanol.

The UV detection can be successfully applied only for a few saponins, which have absorption maxima in the UV range. Glycyrrhizin, glycyrrhetinic acid, and sweet saponin present in *Glycyrrhiza glabra* L. can be detected at 254 nm. For 2,3-dihydro-2,5-dihydroxy-6-methyl-4-pyrone (DDMP)-conjugated soyasaponins, which have an UV absorption maximum at 295 nm, UV detection can also be applied. Similar absorption has been shown by avenacins occurring in *A. sativa* roots. The sensitivity of this detection mode has been satisfactory and, depending on the nature of the saponin, ranges from 50 ng for avenacoside B to 300 ng for ginseng saponins (Oleszek, 2002; Oleszek and Stochmal, 2010).

5.2 Derivatization

The alternative to nonspecific UV detection of the majority of saponins could be precolumn derivatization that introduces a chromophore to their molecule and facilitates UV detection at higher wavelengths. Several different attempts have been made to introduce such a chromophore based on the substitution of different functional groups present in the saponin molecules.

Derivatization with bromophenacyl bromide in the presence of crown ether has been the most applied, however, only to the triterpenes that possess a carboxylic functional group in the molecule. This group can be situated either on the aglycone or on a sugar substituent (uronic acids). The absorption maxima of derivatized compounds range between 257 and 261 nm, depending on the solvent. Such a derivatization method has been successfully used for the determination of oleane saponins in *Phytolacca dodecandra* L'Her. or medicagenic and zhanic acid as well as soyasapogenol B glycosides from *Medicago sativa* L. Other derivatization protocols were listed in detail in the earlier scientific literature reviews (Oleszek, 2002; Oleszek and Stochmal, 2010).

5.3 Evaporative Light Scattering Detection and Charged Aerosol Detection

The methods applying evaporative light scattering detector (ELSD) and charged aerosol detector (CAD) are independent of the presence of chromophore groups, which enables the analysis of weakly or non-UV active and nonvolatile compounds. Therefore, both are good alternatives to the UV detection of saponins and were successfully applied for their analysis. The principle of ELSD was described in the late 1970s, whereas CAD is a more recent developed. Both detectors are mass dependent and share similar principles: use of a pneumatic nebulizer and a heated tube where the solvent evaporates in a detection chamber. However, instead of measuring light scattering as in ELSD, CAD uses an electrometer to measure the electrical charge of the charged particles using a secondary stream of nitrogen passing through a corona discharge needle. Generally, CAD is more sensitive when compared with ELSD system; however, this difference in sensitivity is strongly influenced by the HPLC mobile-phase composition (Vehovec and Obreza, 2010). The sensitivity of ELSD and CAD detectors is determined based on the signal-to-noise ratio. Several factors such as the flow rate of the nebulizer gas, flow rate of the mobile phase, and various concentrations of its additives could affect their sensitivity. The evaporating temperature is also an important adjustable parameter affecting the signal response in case of ELSD detector, similarly to the CAD range.

A HPLC method with ELSD detection was used for the determination of ginsenosides (Rg_1, Re, Rb_1, Rc, Rb_2, and Rd) in *Panax notoginseng* (Burkill) F.H. Chen. The analysis was performed on an ODS RP-18 column utilizing a gradient elution profile consisting of acetonitrile and water in the 60-min program. The ELSD used was set at an evaporating temperature of 35°C and air pressure of 3.4 bar. The detection limit of the saponins was 50 ng (Li and Fitzloff, 2001). Such detector was also applied for separation of platycosides in Platycodi Radix, the root of *Platycodon grandiflorus* (Jacq.) A.DC. used as a food and traditional oriental medicine (Ha *et al.*, 2006; Yoo *et al.*, 2011). In the first case, separation of saponins was achieved on a C18 column. The mobile phase consisted of a gradient of aqueous acetonitrile. The optimum nebulizer gas nitrogen pressure and evaporating temperature in these studies were determined to be 2.5 bar and 70°C, respectively. The second case described the simultaneous determination of 18 platycosides. The optimum conditions for their separation were achieved on an ODS column utilizing gradient elution of eluent A consisting of 30 mM ammonium acetate buffer, pH 4.81 : methanol : acetonitrile, 75 : 5 : 20 (v/v/v),

and solvent B 69 : 5 : 26 (v/v/v). Among the 18 platycosides, platycoside E showed the highest concentration ($2.00\,\mathrm{mg\,g^{-1}}$), followed by polygalacin D2 ($1.77\,\mathrm{mg\,g^{-1}}$) and 3″-O-acetylplatyconic acid A ($1.35\,\mathrm{mg\,g^{-1}}$). The sum of these three compounds is recommended for the quality control of Platycodi Radix for medicinal purposes (Yoo et al., 2011). Saponins profiles of different parts of P. notoginseng were determined using the HPLC-ELSD system. Eight major saponins, notoginsenoside R_1 and ginsenosides Rg_1, Re, Rb_1, Rc, Rb_2, Rb_3, and Rd, were successfully separated on a C18 column using an acetonitrile and water gradient. The drift tube temperature was set at 60°C and the flow rate of the nebulizer nitrogen gas was $1.4\,\mathrm{mL\,min^{-1}}$. They were also quantitatively compared among the different parts of the plant, for example, root, fiber root, rhizome, stem, leaf, flower, and seed. The chromatograms showed that there was significant difference between underground (root, fiber root, and rhizome) and aerial (leaf and flower) parts from P. notoginseng (Wan et al., 2006). Moreover, the validated HPLC-ELSD method was successfully used for the analysis of lupane- and oleane-type saponins (pulsatilloside E, anemoside B4, cussosaponin C, pulsatilla saponin H, and hederacolchiside E) in Pulsatilla koreana Nakai (Lee et al., 2010), seven saponins (macranthoidins A and B, macranthosides A and B, dipsacoside B, and two hederagenin glycosides) in Flos Lonicerae, the dried buds of several species of the genus Lonicera, a commonly used traditional Chinese medicine (Chai, Li, and Li, 2005), steroidal saponins of Asparagus racemosus Willd. (Patricia et al., 20082008), and soya saponins (Lin and Wang, 2004).

The HPLC-CAD system was used for the determination of P. koreana triterpenoid saponins. Analytes were separated on a C18 column with gradient elution of methanol and water, at a flow rate of $0.8\,\mathrm{mL\,min^{-1}}$, the temperature was set at 30°C. The method provided excellent resolution and sensitivity for the four triterpenoid saponins within 26 min of analysis. Linear calibration curves were obtained within the concentration range $2–200\,\mathrm{\mu g\,mL^{-1}}$ for pulsatilloside E, anemoside B4, and cussosaponin C and $5–500\,\mathrm{\mu g\,mL^{-1}}$ for pulsatilla saponin H. The limit of detection (LOD) and limit of quantification (LOQ) were 0.04–0.2 and $2–5\,\mathrm{\mu g\,mL^{-1}}$, respectively. The validity of the developed HPLC-CAD method was confirmed by satisfactory values of linearity, intra- and interday accuracy, and precision

(Yeom et al., 2010). In another study, HPLC-CAD was applied for the quantitative analysis of notoginsenoside R_1 and ginsenosides Rg_1, Re, Rb_1, Rg_2, Rh_1, and Rd in 30 batches of P. notoginseng samples. Moreover, different methods of their detection (UV, ELSD, and CAD) were compared. Comparison of the peak area of seven compounds of the same concentration showed that the response of CAD is apparently higher than the other two. Mean values of peak area were calculated as 7.66 times of CAD to ELSD while 16.76 times of CAD to UV. The sensitivity, which is the main performance index of a detector, was investigated; LODs and LOQs of the three detectors were examined. Results clearly showed that sensitivity of CAD was higher than that of ELSD or UV. Furthermore, the CAD exhibited a steadier baseline in gradient elution compared with UV at 203 nm (Bai et al., 2009). Liquid chromatography with CAD, ELSD, and UV detection was also applied for the determination of ginsenosides Rg_1, Re, Rb_1, Rc, Rb_2, Rb_3, and Rd in Panax ginseng C.A. Mey. Similarly, it was shown that CAD had higher sensitivity than UV and ELSD. Furthermore, it also had better linearity and reproducibility when compared with ELSD system (Wang et al., 2009). Both CAD and ELSD detectors were also applied for the detection of saikosaponins in Bupleuri Radix, the dried root of Bupleurum falcatum L. The best sensitivity for CAD was achieved with 0.1 mM ammonium acetate at pH 4.0 in the mobile phase with a flow rate of $1.0\,\mathrm{mL\,min^{-1}}$ and the CAD range at 100 pA, whereas that for ELSD was achieved with 0.01% acetic acid in the mobile phase with a flow rate at $0.8\,\mathrm{mL\,min^{-1}}$. The sensitivity for CAD was two to six times better than that of ELSD (Eom et al., 2010).

6   ULTRA-PERFORMANCE LIQUID CHROMATOGRAPHY (UPLC)

The arguments mentioned above confirm the usefulness and validity of using HPLC for the analysis of saponins in various biological samples. However, in recent years, ultra-performance liquid chromatography (UPLC) has been increasingly used for the determination of such compounds in various plant extracts. The biggest advantage of the UPLC method is a significant reduction in the time of analysis. It should be noted that the speed of the chromatographic separation is usually 5–10 times faster and without reducing resolution when compared with HPLC.

Moreover, the amount of mobile phase used for the analysis is also greatly reduced. The quantity of the sample or extract needed for a typical UPLC separation is much smaller than that in the case of conventional HPLC.

The UPLC analysis uses shorter columns (usually 50–100 mm vs 150 × 250 mm used in HPLC scale), with a smaller inner diameter (2.1 vs 4.6 mm) and with less graining in the column bed (1.7 vs 5 µm, respectively). The use of very fine size of separation material significantly increases the efficiency of the column. Through all these factors, high resolution and faster analyses can be provided. Furthermore, the time of analysis is significantly shortened, from about 30–80 min in case of conventional HPLC to as much as 5–20 min in the case of UPLC. The total flow of the mobile phase is also reduced due to the much smaller diameter and shorter columns. The separations usually have been performed at a flow rate of 0.250–0.350 versus 1–1.5 mL min^{-1} used in a HPLC system. The separation of the plant extract is carried out at an increased linear flow of the mobile phase. Chromatography columns have been packed by smaller beads with smaller interstitial spaces, which increase the resistance to solvent flow. Thus, higher solvent pressures are required to use. Typically, pressures of up to 400 bar are required for HPLC with 5-µm column beads, whereas up to 1000 bar is required for UPLC with 1.7-µm beads.

The separation of saponins using UPLC was developed for the *P. notoginseng* extract. This analysis allowed the simultaneous determination of 11 saponins: notoginsenoside R_1 and ginsenosides Rg_1, Re, Rf, Rb_1, Rg_2, Rc, Rb_2, Rb_3, Rd, and Rg_3. Moreover, a comparison was made between the UPLC and the HPLC analyses, both performed on C18 columns. The time of analysis was 12 versus 60 min, the column temperature was maintained at 45 versus 40°C, the flow rate was set at 0.35 versus 1.5 mL min^{-1}, and the injection volume was 1 versus 10 µL for the UPLC and HPLC analyses, respectively. In both the cases, the standards and samples were separated using a gradient mobile phase consisting of water and acetonitrile and the detection wavelength was set at 203 nm. UPLC analysis needed only 12 min for good separation of 11 saponins in *P. notoginseng*, which is only one-fifth of the time required for the HPLC analysis. Furthermore, UPLC analysis was characterized by better resolutions of ginsenoside Rg_1 and Re as well as ginsenoside Rb_2 and Rb_3 (Guan, Lai, and Li, 2007). Eight triterpene saponins (cauloside H, leonticin D, cauloside G, cauloside D, cauloside B, cauloside C, cauloside A, and saponin PE) from roots of *Caulophyllum thalictroides* (L.) Michx were separated and determined within 8.0 min using the UPLC method performed on an Acquity UPLC™ BEH Shield RP18 column (50 mm × 2.1 mm, 1.7 µm) and within 35 min using the HPLC method on a Synergi MAX RP column (150 mm × 4.6 mm; 4 µm). The ELSD detection was used in both cases. The mobile phase consisted of ammonium acetate (50 mM) and acetonitrile at a flow rate of 0.25 or 1.0 mL min^{-1}, respectively. The limits of detection for the saponins were between 7 and 10 µg mL^{-1} using UPLC-ELSD as compared to between 10 and 35 µg mL^{-1} using HPLC-ELSD system (Avula *et al.*, 2011). Ruscogenin content in various *Ruscus* spp. was determined by UPLC with UV detection at a wavelength of 200 nm. Analysis were performed on C18 column utilizing a mobile phase consisting of acetonitrile and water (90 : 10 v/v), at a constant flow rate of 0.3 mL min^{-1} (Güvenç, şatir, and Coşkun, 2007).

Plant extracts are a complex mixture of saponins, which differ in the type of aglycone and sugar chains. Moreover, sugar rings can be substituted in various places of aglycone; they may contain various sugars and their sequence in the chain can also be variable. For these reasons, the number of common combinations is enormous. Even in one plant, it is possible to find a mixture of several glycosides, whose structures differ according to the tested organ: seeds, roots, aerial parts, or flowers. Therefore, the UPLC system directly coupled to MS has been in routine use in several laboratories because it provides much more information. Enhanced selectivity, sensitivity, and resolution and faster analyses have made these techniques the predominate technology for both quantitative and qualitative analyses. It is a very useful tool for the examination of a large number of plant samples, which is necessary in case of metabolome studies and when comparing numerous species or mutants, or different parts of the plants, where the differences in both chemical composition and content of these compounds are suspected and quality control is required. Moreover, it also enables the initial determination of the structures of new compounds synthesized under *in vitro* conditions and absent in the native plants. The full structure, that is, a combination and sequence of individual sugar molecules in the glycosidic chain attached to the aglycone, can be determined based on the data obtained by the NMR analysis.

7 MASS SPECTROMETRY OF SAPONINS

7.1 Gas–Liquid Chromatography–Mass Spectrometry

Because of their low volatility and high molecular mass, saponins are generally not suited for GC/MS (gas chromatography–mass spectrometry) analyses. Although derivatization can be applied to increase volatility, because of the attachment of groups from a derivatizing agent, this frequently involves significant increase in the molecular mass of an analyte. The highest detectable mass-to-charge ratio (*m/z*) of a typical gas chromatograph–single quadruple mass spectrometer system very rarely exceeds 1000 mu, which further limits the application of this technique for the analysis of saponins. These limitations can, however, be circumvented by an indirect analysis method, in which the analyte is first hydrolyzed and then the aglycone and monosaccharides constituting the glycone part of a saponin are analyzed separately. The latter study is often employed during the characterization of novel saponins. Absolute configurations of monosaccharides can be determined using appropriate derivatization or a chiral GC column. A good example of one such procedure is given by Luo, Chen, and Kong (2011), where, after the hydrolysis of each saponin, derivatization method of Hara, Okabe, and Mihashi (1987) is applied to convert aldoses into trimethylsilyl esters of methyl 2-(polyhydroxyalkyl)-thiazolidine-4(*R*)-carboxylates. The latter compounds were then separated by GC and their retention times, in comparison with retention times of authentic standards, led to establishing absolute configurations of monosaccharides they originated from.

Sapogenins in the post-hydrolysis mixture can be readily separated from carbohydrates of glycosides by liquid–liquid extraction. Similar approach, derivatization followed by GC combined with either mass spectrometric or flame ionization detector, can then be used to analyze and quantify them. Concentrations of sapogenins are frequently used as an estimate of concentrations of saponins. Recent example of such use is the work of Balestrazzi *et al*. (2011), who tested the biological activity of *M. sativa* saponins on white poplar (*Populus alba* L.) suspension cell cultures.

One drawback of the indirect analysis method is its dependence on a successful hydrolysis of all saponins in the sample. The reaction time, the type, and the concentration of the hydrolyzing agent and the medium in which saponins are dissolved are all critical and need to be carefully evaluated to ensure the completeness of the process. In some cases, even relatively gentle hydrolysis conditions can yield artifacts because of the aglycone decomposition and/or modification (Oleszek, 2002).

7.2 Liquid Chromatography–Mass Spectrometry

Chemical properties of saponins make them very well suited for atmospheric pressure ionization (API) methods, such as electrospray ionization (ESI) and atmospheric pressure chemical ionization (APCI). Both methods are considered as soft ionization techniques, because under normal conditions, they do not cause excessive fragmentation of the analytes. Instead, ions comprising intact molecules are formed. On the interface level, the two methods differ mainly in the way the electric charge is applied: via a sample capillary in ESI or via a corona needle in APCI. Typically, ESI is used for the analysis of polar and high molecular mass compounds, whereas APCI is usually applied for less polar and small molecular mass compounds. Both methods can be utilized for the analysis of saponins, even though applications of APCI for this purpose are relatively rare in comparison with ESI.

When API interfaces are operated in the positive ion mode, saponins can produce ions by direct protonation (forming $[M+H]^+$ ions), but molecules with low proton affinity also gain charge by cationization with a variety of metal cations from the mobile phase. Ions resulting from attachments of sodium $[M+Na]^+$ and potassium $[M+K]^+$ are typically observed, whereas ions with lithium can be generated by adding lithium salts to the mobile phase. These adduct ions are very useful for structural characterization and quantitative studies of molecules that do not ionize easily in the positive ion mode. In the negative ion mode, saponins gain charge predominantly by a loss of the proton ($[M-H]^-$ ions). Neutral molecules can, analogous to positive ion mode, be ionized by an anion attachment. Usually, these anions originate from the mobile phase additives, for example, from formic acid and acetic acid, but other attachments, such as $[M+Cl]^-$, are also often observed. Anion attachment ions are frequently used

in quantitative analyses and they can aid in the discrimination of the analyte ion from the matrix. Owing to their stability, they are not as useful for structural studies as adducts in the positive mode. Depending on the construction of the ion source, excessive formation of such adducts can usually be prevented by appropriate potential on the ion entrance (also called *sample cone*).

Although a sample can be supplied into the electrospray ion source with the effluent from a liquid chromatography column, for some applications, it is also possible to use the so-called direct infusion or direct injection mode. In this mode, a solution of the sample is introduced to the ion source using a syringe pump. This is often used to obtain a fingerprint of saponins in the sample, for example, *Medicago truncatula* Gaertn. triterpene saponins presented in Figure 5 (Kapusta *et al*., 2005), and in the fragmentation and structure elucidation studies (Bankefors *et al*., 2011). Such an analysis is usually very quick and relatively easy to perform, but it should be pointed out that, particularly for the samples of plant extracts, it is very susceptible to ion suppression matrix effects. Substances that ionize easily will produce strong signals and will frequently reduce and sometimes completely prevent ionization of other components of the mixture. Furthermore, on low resolution mass spectra, isobaric compounds (with the same nominal mass but different elemental compositions and exact masses) are recorded together as one peak of given mass-to-charge ratio. Because of these limitations,

the observed composition of the sample may not correspond very well to its actual composition.

8 MASS ANALYZERS USED IN LC–MS ANALYSIS OF SAPONINS

8.1 Quadruple Analyzers

Quadruple mass analyzers are universally used in saponin analysis because of their low cost, simplicity of operation, and small size. They separate ions based on their transmission through an oscillating electric field. Single quadruple devices offer low resolution selected ion monitoring (SIM) capabilities that can be used for quantitative analyses. Employing in-source fragmentation technique, they can be utilized for very limited range of structure determination studies. Detection limits of SIM-based methods are sometimes not sufficient for reliable quantitation, because of the difficulties in discriminating ions of analytes from other ions in the sample and/or in the mobile phase. Such problems can be solved using triple-quadruple systems, which are capable of second-stage mass analyses (MS/MS, MS2). Triple-quadruple detectors are very useful for structure determination, as they can be operated in daughter scan, parent scan, and neutral loss monitoring modes. The biggest advantages of triple-quadruple systems are, however, their excellent
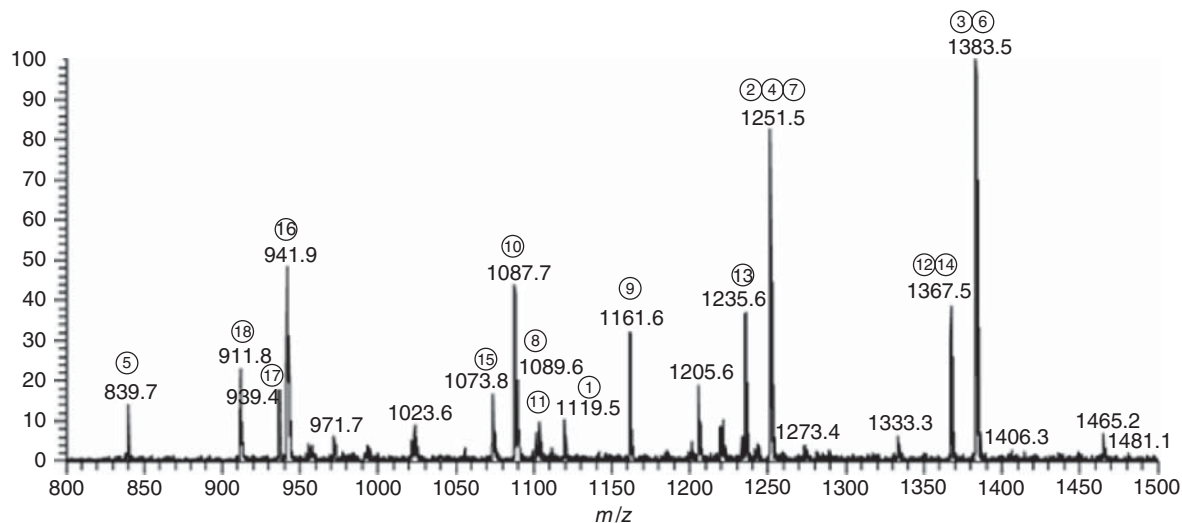
Figure 5 Direct injection MS fingerprint of *Medicago truncatula* Gaertn. saponins.

selectivity and sensitivity. They are most efficiently used in targeted quantitative analyses in selected reaction monitoring (SRM, sometimes also called *MRM*) mode.

An example of such use is the work of Munafo and Gianfagna (2011) on steroidal glycoalkaloids and furostanol saponins from Easter lily (*Lilium longiflorum* Thunb.). In this case, good sensitivity and selectivity of employed analytical method allowed for quantitation of three saponins and two glycoalkaloids on an organ level, in flower buds, stems, leaves, bulbs, and roots, and on a tissue level, in specific parts of a bulb. The analysis was performed in the positive ion mode, in which the analyzed furostanol saponins formed strong $[M - H_2O + H]^+$ ions (Munafo *et al.*, 2010). For each saponin, a fragmentation reaction, that is, the loss of a neutral dehydrohexose from the parent ion, was selected. One such reaction per analyzed compound is generally sufficient for reliable quantitation. However, in the case of co-eluting analytes forming similar product ions, additional qualifier reactions may be required. In MS methods based on SIM monitoring, additional qualifier ions also ensure dependability and aid in the discrimination of the analyte ions. Unfortunately, for both techniques, too many simultaneous reactions or ions to monitor can severely affect the overall sensitivity and accuracy of the method. This can be a serious problem in multicomponent analysis methods, where several compounds are quantitated in a single run. Good results can often be obtained using the segmented acquisition approach, in which only compounds eluting at a given retention time window are monitored. Alternatively, if enough of the sample is available, it can be reanalyzed multiple times with different sets of acquisition functions.

For SIM-based LC-MS quantitation of 16 ginsenosides in the rat plasma, Liang *et al.* (2010) selected $[M + Cl]^-$ adduct ions for monitoring and compared one-segment and multi-segmental acquisitions . As expected, lower limits of quantitation (LLOQ) were much higher (two- to fivefold increase) in one segment acquisition mode. The best results, LLOQs 4–10 times lower than that in one segment acquisition control, were obtained with five segments acquisition mode, in which no more than four ginsenosides were analyzed in each time window. The authors also evaluated five segments mode in terms of accuracy, precision, and matrix effects.

Foubert *et al.* (2010) developed the UPLC-MS/MS methodology for the quantitative analysis of 14 triterpene saponins from *Maesa lanceolata* Forssk. Saponins were quantitated in the positive ion mode, but sodium adduct ions were selected as precursors for SRM mode acquisition. The choice was dictated by intensity of their signals (the highest observed were for $[M + Na]^+$ and $[M - H]^-$ ions) and stability of their fragmentation. According to the authors and in agreement with the previous findings (Li *et al.*, 2005), for sodium adduct ions, fragmentation occurs mainly in the glycan part of a molecule. The authors were unable to obtain stable fragmentation from deprotonated ions; thus, the fragmentation reaction from appropriate sodium adduct ion, which was the same for all analyte fragment ions at *m/z* 348, was selected. The method was proven selective enough for very reliable quantitation and partially validated.

8.2   Low Resolution Ion Trap Analyzers

Ion trap mass analyzers operate on the basis of cyclical motion of ions in a magnetic or electro-dynamic field. Two designs of the analyzer, older 3D quadruple ion trap and newer 2D linear ion trap, are in common use. Both designs offer the possibility of multiple-stage mass analysis (MS^n), although 3D ion trap is slightly limited in this aspect – it cannot record product ions with *m/z* lower than approximately one-third of the precursor's *m/z* value. Typically, ion traps are operating as scanning mass spectrometers, even though they are usually fully capable of both single ion and single reaction monitoring. Their application for quantitative analyses is somewhat limited by the restricted dynamic range, which is dependent on the number of ions that can be trapped at the same time and by the length of a duty cycle (ion injection, trapping, and ejection), which can affect detection limits. Nevertheless, by a combination of multiple-stage MS with collision-induced dissociation (CID), ion traps provide excellent platforms for structure elucidation. In this capacity, they are ubiquitously used in saponin analysis.

As demonstrated by Kapusta *et al.* (2005), a 3D ion trap can be successfully used in both structural and quantitative studies. The authors identified 13 primary *M. truncatula* triterpenoid saponins by means of direct injection multiple-stage MS analysis. Following separation of saponin fraction by high pressure liquid chromatography, additional seven isobaric compounds were tentatively identified. All

compounds were subsequently quantitated in three *M. truncatula* subspecies with a SIM-based method.

A similar approach was employed by Zhang *et al.* (2010), who used the 3D ion trap to establish the structural characteristics of steroidal saponins in *Paris polyphylla* Sm. Successive negative ion mode ESI-MS[n] experiments were performed to investigate fragmentation patterns of 11 reference saponins with known structures. Knowledge gained in this experiment was then applied to structural elucidation of three unknown saponins. The study was completed by the development of quantitation method based on SRM using the triple-quadruple HPLC-MS/MS system.

Li *et al.* (2005) investigated the fragmentation pathways of triterpene saponins from *Symplocos chinensis* (Lour.) Druce using the quadruple linear ion-trap mass spectrometer. Quadruple linear trap devices combine regular quadruple analyzer, quadruple collision cell, and linear ion trap, which in addition to its trapping capabilities can function as a regular quadruple analyzer. Such a versatile combination allows performing all types of experiments characteristic for triple-quadruple systems: daughter scans, precursor scans and SRM, and the MS[n] experiments typical for the classic 3D ion traps.

Bankefors *et al.* (2011) used the 3D ion trap mass spectrometer to characterize 15 saponins from *Q. saponaria* Molina bark. On the basis of the fragmentation patterns and peak intensity ratios from MS1 to MS3 experiments, they established a stepwise flowchart-based procedure for the systematic identification of the structural elements in *Quillaja* saponins.

8.3 High Resolution Mass Analyzers

One of the key elements in characterization of compounds with natural origin is accurate mass measurement. On the basis of this, the elemental composition can be established and/or verified. Accurate mass measurements are often required for successful publication of newly discovered natural products. As a result, a majority of phytochemical research publications list masses and elemental compositions of compounds they describe. This, along with good spectra library, can significantly aid in the dereplication process and prevent the costly and time-consuming purification and identification of compounds that were already studied.

8.4 Time-of-Flight Analyzers

Three decades ago, accurate mass measurement of saponins usually involved using fast atom bombardment ionization (FAB: either in static form with sample on a probe or in continuous flow with HPLC to deliver the analyte) and rather expensive magnetic sector instruments. Although such measurements and such systems still occasionally make an appearance in research articles, the vast majority of accurate mass measurements are nowadays performed on much more affordable systems with time-of-flight (TOF) analyzers. The TOF analyzers derive m/z from the velocity of ions accelerated to the same kinetic energy. Using the so-called reflectron electrodes to increase the flight distance and focus ion beams, TOF analyzers are able to achieve 5-ppm mass accuracy [defined as $10^6 \times$ (theoretical m/z − observed m/z)/theoretical m/z]. Although TOF instruments are generally used as scanning mass spectrometers, employing in-source fragmentation by appropriate fragmentor electrode potential, they can also be very valuable tools for structure elucidation studies (Qi *et al.*, 2008; Guillarme *et al.*, 2010). Owing to their high scanning rates, TOF instruments can be easily paired with UPLC systems, which produce narrow chromatographic peaks (Guillarme *et al.*, 2010). This feature was extensively used for screening and multicomponent analyses (Ren *et al.*, 2008).

Perhaps the most versatile hybrid mass spectrometer is combination of quadruple and time-of-flight (QTOF) analyzers. Such devices are very commonly used in saponin analysis, because they offer high resolution quantitation and structure elucidation features. Numerous applications have been published in recent years. For example, Kang *et al.* (2012) used UPLC-QTOF-MS to analyze steroidal saponins in *P. polyphylla* with MS[E] data acquisition – a combination of full scan and "high energy" scan providing high resolution MS and MS/MS data in the single analysis. On the basis of the fragmentation patterns from 21 reference compounds, using high resolution MS, they were able to tentatively identify and propose structures for 98 saponins.

8.5 High Resolution Ion Trap Analyzers

Recently commercialized Orbitrap mass analyzers constrain ions between a central electrode and two

outer electrodes by a combination of electrostatic and centrifugal forces. Initial designs suffered from slow acquisition rates, but introduction of hybrid linear ion trap–Orbitrap instruments allowed to overcome this issue. Recent applications for multicomponent analysis of traditional Chinese medicines, Chaihu-Shu-Gan-San by Su *et al.* (2010) and Xin-Ke-Shu by Peng *et al.* (2011), indicate unmatched potential of this type of instrument for saponin analysis.

Fourier transform ion cyclotron resonance mass analyzer offers a very high resolving power, wide mass range, and a number of features derived from its ion-trap character, such as ion storage and MSn. Besides prohibitive costs of purchase, it requires a cryogenically cooled magnet and an operating vacuum in the nano Pa range. Applications of such devices to the analysis of saponins are quite rare. One recent example is the work of Pollier *et al.* (2011) on *M. truncatula*. The authors performed the profiling of the saponins in root hairs and identified 79 saponins, of which 61 were not detected before in *M. truncatula*. For each compound, high resolution MSn spectra were generated, which led to the identification of four new sapogenins.

9   RELATED ARTICLES

Extraction Methodologies: General Introduction; New Trends in Extraction of Natural Products: Microwave-Assisted Extraction and Pressurized Liquid Extraction; Solid-Phase Microextraction (SPME) and Its Application to Natural Products; Thin-layer Chromatography, with Chemical and Biological Detection Methods; HPLC and Ultra HPLC: Basic Concepts; LC and LC-MS: Techniques and Applications; NMR of Small Molecules.

REFERENCES

Avula, B., Wang, Y. H., Rumalla, C. S., *et al.* (2011) *J. Pharm. Biomed. Anal.*, **56**, 895–903.

Bai, C. C., Han, S. Y., Chai, X. Y., *et al.* (2009) *J. Liq. Chromatogr. Relat. Technol.*, **32**, 242–260.

Balestrazzi, A., Agoni, V., Tava, A., *et al.* (2011) *Physiol. Plant.*, **141**, 227–238.

Bankefors, J., Broberg, S., Nord, L. I., *et al.* (2011) *J. Mass Spectrom.*, **46**, 658–665.

Besso, H., Saruwatari, Y., Futamura, K., *et al.* (1979) *Planta Med.*, **37**, 226–333.

Bushway, R. J. and Storch, R. H. (1982) *J. Liq. Chromatogr.*, **5**, 731–742.

Chai, X. Y., Li, S. L. and Li, P. (2005) *J. Chromatogr. A*, **1070**, 43–48.

Eom, H. Y., Park, S. Y., Kim, M. K., *et al.* (2010) *J. Chromatogr. A*, **1217**, 4347–4354.

Foubert, K., Cuyckens, F., Vleeschouwer, K., *et al.* (2010) *Talanta*, **81**, 1258–1263.

Guan, J., Lai, C. M. and Li, S. P. (2007) *J. Pharm. Biomed. Anal.*, **44**, 996–1000.

Guillarme, D., Schappler, J., Rudaz, S., *et al.* (2010) Coupling ultra-high-pressure liquid chromatography with mass spectrometry. TrAC *Trends Anal. Chem.*, **29**(1), 15–27.

Güvenç, A., şatir, E. and Coşkun, M. (2007) *Chromatographia*, **66**, 141–145.

Ha, Y. W., Na, Y. C., Seo, J. J., *et al.* (2006) *J. Chromatogr. A*, **1335**, 27–35.

Hara, S., Okabe, H. and Mihashi, K. (1987) *Chem. Pharm. Bull.*, **35**, 501–506.

Higgins, J. W. (1976) *J. Chromatogr. A*, **121**, 329–334.

Hostettmann, K. and Marston, A. (1995) *Saponins*, Cambridge University Press, Cambridge, pp. 1–568.

Ireland, P. A. and Dziedzic, S. Z. (1985) *J. Chromatogr. A*, **325**, 275–281.

Ireland, P. A. and Dziedzic, S. Z. (1986) *J. Chromatogr. A*, **361**, 410–416.

Kang, L., Yu, K., Zhao, Y., *et al.* (2012) *J. Pharm. Biomed. Anal.*, **62**, 235–249.

Kapusta, I., Janda, B., Stochmal, A., *et al.* (2005) *J. Agric. Food Chem.*, **53**, 7654–7660.

Kesselmeier, J. and Strack, D. (1981) *Z Naturforsch*, **36c**, 1072–1074.

Lee, K. Y., Cho, Y. W., Park, J., *et al.* (2010) *Phytochem. Anal.*, **21**, 314–321.

Li, W. and Fitzloff, J. F. (2001) *J. Pharm. Pharmacol.*, **53**, 1637–1643.

Li, B., Abliz, Z., Fu, G., *et al.* (2005) *Rapid Commun. Mass Sp*, **19**, 381–390.

Liang, Y., Kang, A., Xie, T., *et al.* (2010) *J. Chromatogr. A*, **1217**, 4501–4506.

Lin, J. and Wang, C. (2004) *J. Food Sci.*, **69**, 456–462.

Luo, J. G., Chen, X. and Kong, L. Y. (2011) *Chem. Pharm. Bull.*, **59**, 518–521.

Marston, A. (2011) *J. Chromatogr. A*, **1218**, 2676–2683.

Marston, A. and Hostettmann, K. (2006) *J. Chromatogr. A*, **1112**, 181–194.

Massiot, G., Lavaud, C., Guillaume, D., *et al.* (1988) *J. Agric. Food Chem.*, **36**, 902–909.

Matsushita, A., Miyase, T., Noguchi, H., *et al.* (2004a) *J. Nat. Prod.*, **67**, 377–383.

Matsushita, A., Sasaki, Y., Warashina, T., *et al.* (2004b) *J. Nat. Prod.*, **67**, 384–388.

Melek, F. R., Miyase, T., Abdel-Khalik, S. M., *et al.* (2004) *Phytochemistry*, **65**, 3089–3095.

Munafo, J. P. and Gianfagna, T. J. (2011) *J. Agric. Food Chem.*, **59**, 995–1004.

Munafo, J. P., Ramanathan, A., Jimenez, L. S., *et al.* (2010) *J. Agric. Food Chem.*, **58**, 8806–8813.

Oleszek, W. (1988) *J. Sci. Food Agric.*, **44**, 43–49.

Oleszek, W. (2002) *J. Chromatogr. A*, **967**, 147–162.

Oleszek, W. and Biały, Z. (2006) *J. Chromatogr. A*, **1112**, 78–91.

Oleszek, W. and Hamed, A. (2010) Saponin based surfactants, in *Surfactants from Renewable Resources*, eds. M. Kjellin and I. Johansson, John Wiley & Sons, Ltd, Chichester, UK, pp. 239–249.

Oleszek, W. and Marston, A. (eds.), (2000) *Saponins in Food, Feedstuffs and Medicinal Plants*, Kluwer Academic Publishers, Dordrecht, pp. 1–291.

Oleszek, W. and Stochmal, A. (2010) High performance liquid chromatography of triterpenes (including saponins), in *High Performance Liquid Chromatography in Phytochemical Analysis*, eds. M. Waksmundzka-Hajnos and J. Sherma, Taylor & Francis Group, LLC, New York, pp. 639–657.

Oleszek, W., Jurzysta, M., Ploszynski, M., *et al.* (1992) *J. Agric. Food Chem.*, **40**, 191–196.

Oleszek, W., Sitek, M., Stochmal, A., *et al.* (2001) *J. Agric. Food Chem.*, **49**, 4392–4396.

Oleszek, W., Kapusta, I. and Stochmal, A. (2008) TLC of triterpenes (including saponins), in *Thin Layer Chromatography in Phytochemistry*, eds. M. Waksmundzka-Hajnos, J. Sherma and T. Kowalska, CRC Press, New York, pp. 519–541.

Patricia, Y.H., Aisyah, H.J., Reg, L., Kerry, P., William, K., James, J.D. (2008). Steroidal saponins from the roots of Asparagus racemosus. *Phytochemistry*, **69**, 796–804.

Peng, J. B., Jia, H. M., Liu, Y. T., *et al.* (2011) *J. Pharm. Biomed. Anal.*, **55**, 984–995.

Pollier, J., Morreel, K., Geelen, D., *et al.* (2011) *J. Nat. Prod.*, **74**, 1462–1476.

Qi, L. W., Cao, J., Li, P., *et al.* (2008) *J. Chromatogr. A*, **1203**, 27–35.

Ren, M. T., Chen, J., Song, Y., *et al.* (2008) *J. Pharm. Biomed. Anal.*, **48**, 1351–1360.

Saito, K., Horie, M., Hoshino, Y., *et al.* (1990) *J. Chromatogr. A*, **508**, 141–147.

Sticher, O. and Soldati, F. (1979) *Planta Med.*, **36**, 30–42.

Su, Z. H., Zou, G. A., Preiss, A., *et al.* (2010) *J. Pharm. Biomed. Anal.*, **53**, 454–461.

Vehovec, T. and Obreza, A. (2010) *J. Chromatogr. A*, **1217**, 1549–1556.

Wall, M. E., Eddy, C. R., McClennan, M. L., *et al.* (1952) *Anal. Chem.*, **24**, 1337–1341.

Wan, J. B., Yang, F. Q., Li, S. P., *et al.* (2006) *J. Pharm. Biomed. Anal.*, **41**, 1596–1601.

Wang, L., He, W. S., Yan, H. X., *et al.* (2009) *Chromatographia*, **70**, 603–608.

Xu, C. J. and Lin, J. T. (1985) *J. Liq. Chromatogr.*, **8**, 361–368.

Yeom, H., Suh, J. H., Youm, J. R., *et al.* (2010) *Bull. Korean Chem. Soc.*, **31**, 1159–1164.

Yoo, D. S., Choi, Y. H., Cha, M. R., *et al.* (2011) *Food Chem.*, **129**, 645–651.

Zhang, T., Liu, H., Liu, X. T., *et al.* (2010) *J. Pharm. Biomed. Anal.*, **51**, 114–124.

30

Cardiotonic Glycosides

Liselotte Krenn

Department of Pharmacognosy, University of Vienna, Vienna, Austria

1 INTRODUCTION

Cardiotonic steroids are highly active agents and have been proven in plants, animals, and as the so-called endogenous digitalis-like factor in humans. Due to their enhancing effect on heart contractility, the therapeutic use of plants containing cardiac glycosides has been common since ancient times. In the second half of the past century, isolated cardiotonic glycosides replaced herbal drugs for the treatment of congestive heart failure and related diseases. The anticancer potential and the fact that some members of this class of compounds act as steroid hormones in humans have intensely stimulated scientific research during the past two decades. Thus, numerous analytical methods for different purposes, for example, the standardization of plant material or preparations and the quantification of plasma levels for diagnostic or toxicological purposes, have been developed, which are discussed in this context.

1.1 Use of Cardiotonic Glycosides in Cardiac Diseases

Plants containing cardiotonic glycosides have been used medicinally since ancient times. First reports of the use of squill (*Urginea maritima* (L.) Baker syn. *Charybdis maritima* (L.) Speta) in dropsy date back to the Egyptian Papyrus Ebers (circa 1550 BC).

In ancient Greek, not only this drug had a continuous use (Krenn *et al.*, 2001) but also hellebore was employed in therapy. Around AD 1500, other species containing cardiac glycosides such as lily of the valley, foxglove, or pheasant's eye were mentioned for the first time in herbal books. During the eighteenth century, especially squill and foxglove became indispensable herbal drugs for the treatment of cardiac insufficiency. Toad venoms, as an animal source of cardiac steroids, have a long-standing use in traditional Chinese medicine (Teuscher and Lindequist, 2010). In the nineteenth century, several attempts to isolate the active principles in these plants were successful. Nevertheless, definite structure elucidation started in the first half of the twentieth century. In this era, the occurrence of such compounds in animal kingdom, namely in amphibia, was shown as well (Baumgarten, 1963; Ritz and Schoner, 2008). Around the 1950s, isolated cardiotonic glycosides from different plant sources, for example, *Digitalis* species, *Convallaria majalis* L., and *U. maritima*, were introduced into therapy and became the most important option for the treatment of congestive heart failure. Due to the introduction of new therapeutics into cardiology, the therapeutic relevance of cardiac glycosides in cardiac insufficiency, atrial fibrillation, or paroxysmal supraventricular tachycardia decreased during the 1990s despite a controversial debate on their effectiveness (Mordasini *et al.*, 2002; Prassas and Diamandis, 2008).

1.2   New Studies and Therapeutic Potential of Cardiotonic Steroids

In the 1990s, the anticancer potential of cardiotonic glycosides was deduced from retrospective epidemiological studies. The results pointed to a trend that in patients undergoing treatment with cardiotonic glycosides because of heart problems, lethality from cancer was lower compared to other populations (Prassas and Diamandis, 2008). As cardiotonic steroids act as specific inhibitors of sodium-potassium adenosine triphosphatase (Na^+/K^+-ATPase), and several therapy-resistant cancers such as glioblastome, melanoma, or colon cancers are known to overexpress subunits of Na^+/K^+-ATPase, the α-subunits of the sodium pump are discussed as a promising target for the treatment of cancers with poor prognosis (Mijatovic *et al.*, 2012). The initial observations stimulated intense research on the potential of cardiac glycosides against cancer. The supportive outcome of many *in vivo* and *in vitro* studies has resulted in first clinical trials (Prassas and Diamandis, 2008; Newman *et al.*, 2008).

The discovery that cardenolides and bufadienolides are synthesized in humans and sub-nanomolar concentrations influence multifaceted intracellular signaling pathways via the signalosome signal transduction system (Schoner and Scheiner-Bobis, 2009) has additionally stimulated research on this group of compounds.

1.3   Chemistry

Cardiotonic steroids are colorless, bitter substances and comprise two groups of polyhydroxy C-23 or C-24 derivatives with a 10,13-dimethylsterane core structure. Of the four rings of the core structure, A/B and C/D usually are cis fused, whereas B/C is trans fused with few exemptions of trans–trans–cis connection. The active compounds are characterized by a lactone ring in position C-17β, which is five-membered and monounsaturated in cardenolides (C-23) or six-membered and diunsaturated in bufadienolides (C-24) (Figure 1). In plant kingdom, the substances occur predominantly as glycosides, whereas from animals (toads or snakes), mainly unconjugated bufadienolides or esters have
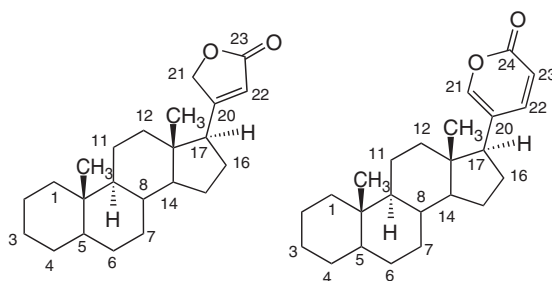
Figure 1   Basic structures of cardenolides and bufadienolides.

been isolated. In humans, only genins of both groups are known until now.

More than 100 genins and around 1000 cardiac glycosides have been isolated from plants and animals. The majority of them derive from the group of cardenolides (Krenn and Kopp, 1998; Grimm-Bursch, 1990; Mayerhofer, 1995). The increased interest for this class of compounds as an upcoming option for cancer treatment has recently again stimulated research focusing on the isolation from diverse sources and structure elucidation by spectroscopic and spectrometric methods. This resulted in the characterization of approximately 15 new cardenolides only in 2012.

The huge structural diversity of cardiotonic steroids is based on differing substitution patterns of the aglycone as well as on the diversity of the glycosidic part. In most of the aglycones, hydroxy groups are attached at C-3 and C-14. Additional hydroxy groups can occur at C-1, C-5, C-6, C-8, C-9, C-11, C-12, and C-16, and some of them might be esterified by acetic or formic acid. In several compounds, epoxy- or oxogroups are fused to the genine. The methylgroup at C-10 might be oxidized to a hydroxymethyl-, formyl-, or carboxyl group. Besides ubiquitous sugars such as D-glucose, L-rhamnose, L-arabinose, or D-xylose, the carbohydrate part of cardiac glycosides contains characteristic 6-deoxy- and 2,6-dideoxy-sugars such as D-digitoxose, D-cymarose, L-oleandrose, or L-thevetose. Usually linear chains of at most five sugars via $1 \rightarrow 4$ linkage are attached to the hydroxy group at C-3 or rarely at C-5 of the aglycone. Very few compounds with a branched sugar chain have been isolated as well (Krenn *et al.*, 1991a).

1.4 Occurrence

Cardiotonic glycosides have been proven in approximately 20 plant families, for example, Plantaginaceae/Scrophulariaceae, Apocynaceae, Ranunculaceae, Hyacinthaceae, or Convallariaceae. In animal kingdom, obviously only toads from the Bufonidae family, few snakes such as *Rhabdophis tigrinus*, or some insects, fire flies (*Photinus* species), or beetles from the genera *Chrysolina*, *Dlorochrysa*, and *Oreina* are able to autogenously synthesize such steroids. Some other insects enrich the compounds from their host plants for chemical defense (Teuscher and Lindequist, 2010).

Several cardiotonic steroids were proven in mammalians. In human plasma, urine, lenses, adrenal cortex, or hypothalamus, the cardenolides ouabain and digoxin as well as several bufadienolides, for example, 19-nor-bufalin, marinobufagin, telocinobufagin, or proscillaridin A have been detected as endogenous cardiotonic steroids. An increasing body of evidence is showing that these endogenous cardiotonic steroids are synthesized in the adrenal gland (Prassas and Diamandis, 2008; Gao *et al.*, 2011; Bagrov *et al.*, 2009).

2 ANALYSIS OF CARDIOTONIC GLYCOSIDES

Due to the clinical importance of cardiotonic steroids and their narrow therapeutic index, the precise quantification of these compounds in herbal drugs and preparations as well as the therapeutic monitoring in the control of the optimum therapeutic window was early recognized as essential. Until the end of the 1970s, the activity of herbal drugs was determined via the biological toxicity test in guinea pigs (Tittel, 1986). For drug monitoring, radioimmunoassays (RIA) were developed for digitoxin and digoxin at the end of the 1960s followed by the introduction of respective enzyme-multiplied immunoassays (EMIT®) and fluorescence-polarized immunoassays (FPIA) (Ahnhoff *et al.*, 1985).

Immediately after the introduction of high-performance liquid chromatography (HPLC), this method became the most important technique for the analysis of cardiotonic steroids in different sources as well as for drug monitoring (Castle, 1975; Enson and Seiber, 1978; Shimada *et al.*, 1976; Jurenitsch *et al.*, 1982).

2.1 Chromatographic Methods

The major analytical methods for the identification and quantification of cardiotonic glycosides are HPLC and thin-layer chromatography (TLC), whereas gas chromatography only plays a minor role.

2.1.1 High Performance Liquid Chromatography (HPLC)

Due to the rapid acceptance of HPLC for the analysis of cardenolides and bufadienolides in different materials (plants, insects, defensive secretions, plasma, etc.), a huge body of literature providing respective methods has been published since the early 1980s. Among the spectrum of different detection modes, the hyphenation with mass spectrometry (MS) and nuclear magnetic resonance (NMR) has further strengthened the role of HPLC in the research on cardiotonic steroids and their glycosides or esters. Because of the well-established and validated methods that have been developed and improved for the most important herbal drugs since almost three decades, ultra-performance liquid chromatography (UPLC) and hydrophilic interaction chromatography (HILIC) until now have not reached a significant role in the determination of this group of compounds.

Besides the analysis of plant material during recent years, several studies focused on methods for the examination of bufadienolides in toad venoms that are frequently used in traditional Chinese medicine. Thus, also respective HPLC systems for these animal-derived drugs are included in this chapter.

2.1.1.1 Sample Preparation
For cardiotonic glycosides from plant material as well as for bufadienolides and their esters from toad venom and insects, methanol (70–100%) and ethanol (70–95%) are the optimum solvents to extract these compounds as exhaustively as possible. In rare cases, 50% acetonitrile (Kwon *et al.*, 2011) or the mobile phase of the HPLC system (Song *et al.*, 2002) were used as extraction media. The extraction is usually performed by reflux or sonication. Due to the rapid and efficient extraction under avoidance of thermal stress, sonication has become the method of choice (Pellati *et al.*, 2009) although a recent study has again proven the thermostability of digitalis

glycosides during refluxing in 50% methanol and 50% acetonitrile for 1 h (Kwon *et al*., 2011).

The purification of the extracts used frequently until mid of the past decade was deduced from the clean-up applied for spectrophotometric determinations of the total content of cardiotonic glycosides. For the precipitation of accompanying substances, lead(II)acetate solution was added to the alcoholic extracts and excessive lead was removed by addition of disodium hydrogen phosphate solution. After centrifugation, the supernatant was further purified either by partition with organic solvents such as chloroform–iso-propanol (3 + 2) (Stuhlemmer *et al*., 1993) or on Extrelut-20® and elution with chloroform (El-Askary *et al*., 1995).

This "classical" way of enrichment of cardiac glycosides was improved by the introduction of solid phase extraction (SPE) on RP-18 materials. For such a clean-up step, the alcoholic extracts have to be evaporated and redissolved in aqueous acetonitrile. This solution is then applied to a C-18 cartridge. After respective washing steps with water or 15% aqueous acetonitrile, the analytes are eluted, for example, with 50% aqueous acetonitrile providing very good removal of phenolics, which are eluted in the washing steps, and of chlorophyll, which is retained at the cartridge (Pellati *et al*., 2009). Due to the avoidance of precipitation steps, this clean-up provides a much faster and convenient enrichment of the cardiotonic steroids and is also suitable for the study of cardenolides in insects (Abe *et al*., 1996). Thus, SPE purification of extracts is certainly more preferably than clean-up methods using numerous partition steps (Kanojiya and Madhusudanan, 2012).

In the analysis of toad venoms or Chinese medicines containing this ingredient, the extracts are directly injected after centrifugation (Gao *et al*., 2010a), whereas in studies of toad skin, purification with XAD-4 for removal of polar compounds might be recommended (Liu *et al*., 2010a, b). Direct analysis of alcoholic extracts is also possible for plant material with low amounts of accompanying substances, for example, hairy root cultures (Sun *et al*., 2012).

2.1.1.2 Stationary and Mobile Phases

For the separation of cardiotonic glycosides in HPLC, almost exclusively RP-18 phases are employed. The majority of analyses is still performed on columns with a dimension of 250 mm length and 4.6 mm diameter, although in several studies by miniaturization of the columns (50–150 mm length and 2 mm diameter), excellent results have been achieved under reasonable reduction of solvent consumption. The selectivity of different brands and column dimensions has recently been demonstrated impressively (Figure 2) (Pellati *et al*., 2009).

Also in UPLC, a C18 phase (Acquity HSS T3; 100 mm × 2.1 mm; 1.8 µm) was shown to be suitable for the fast separation of bufadienolides (Liu *et al*., 2010a, b).

The potential of HILIC by the diverse separation mode has been shown for a fraction from a toad venom (Figure 3) (Liu *et al*., 2010a, b).

In the selection of the mobile phases for the separation of cardenolides and bufadienolides, the mode of detection has to be considered. Most mobile phases are based on acetonitrile and water, which is an advantage in UV detection or diode array detection (DAD). The very low UV cutoff of acetonitrile at 190 nm makes it a preferable mobile phase component as compared to methanol with a UV cutoff of 205 nm especially in the analysis of cardenolides with a UV maximum at about 220 nm. In addition, the lower viscosity of acetonitrile is advantageous. Thus, the recommendation to use methanol instead of acetonitrile for a slightly better peakshape owing to the polar interactions between this solvent and these analytes (Wong *et al*., 2002) is hardly followed. The use of acidic modifiers such as 0.1% formic acid or 0.3% acetic acid is very common in the analysis of bufadienolides in toad venoms to improve the baseline, because the acidic eluent accelerates the elution of interfering alkaline substances in the venoms (Ye and Guo, 2005). Due to the volatility of these acids, no

Figure 2 Chromatograms of a standard mixture of cardiac glycosides under use of different stationary phases. Column: (a) LiChrospher RP-18 (125 mm × 4.0 mm I.D., 5 µm); (b) Zorbax SB-C18 (150 mm × 4.6 mm I.D., 5 µm); (c) Zorbax SB-Aq (150 mm × 4.6 mm I.D., 5 µm); (d) Symmetry C18 (75 mm × 4.6 mm I.D., 3.5 µm). Compound 1: digoxigenin; 2: deacetyllanatoside C; 3: digoxigenin-bis-digitoxoside; 4: gitoxigenin; 5: digoxin; 6: lanatoside C; 7: digitoxigenin; 8: *α*-acetyldigoxin; 9: *β*-acetyldigoxin; 10: lanatoside B; 11: gitoxin; 12: lanatoside A; 13: digitoxin. (*Source*: Reprinted from J Chromatogr A, 1216, Pellati F, Bruni R, Bellardi MG, Bertaccini A, Benvenuti S, Optimization and validation of a high-performance liquid chromatography method for the analysis of cardiac glycosides in Digitalis lanata, 3260–3269, Copyright (2009), with permission from Elsevier.)
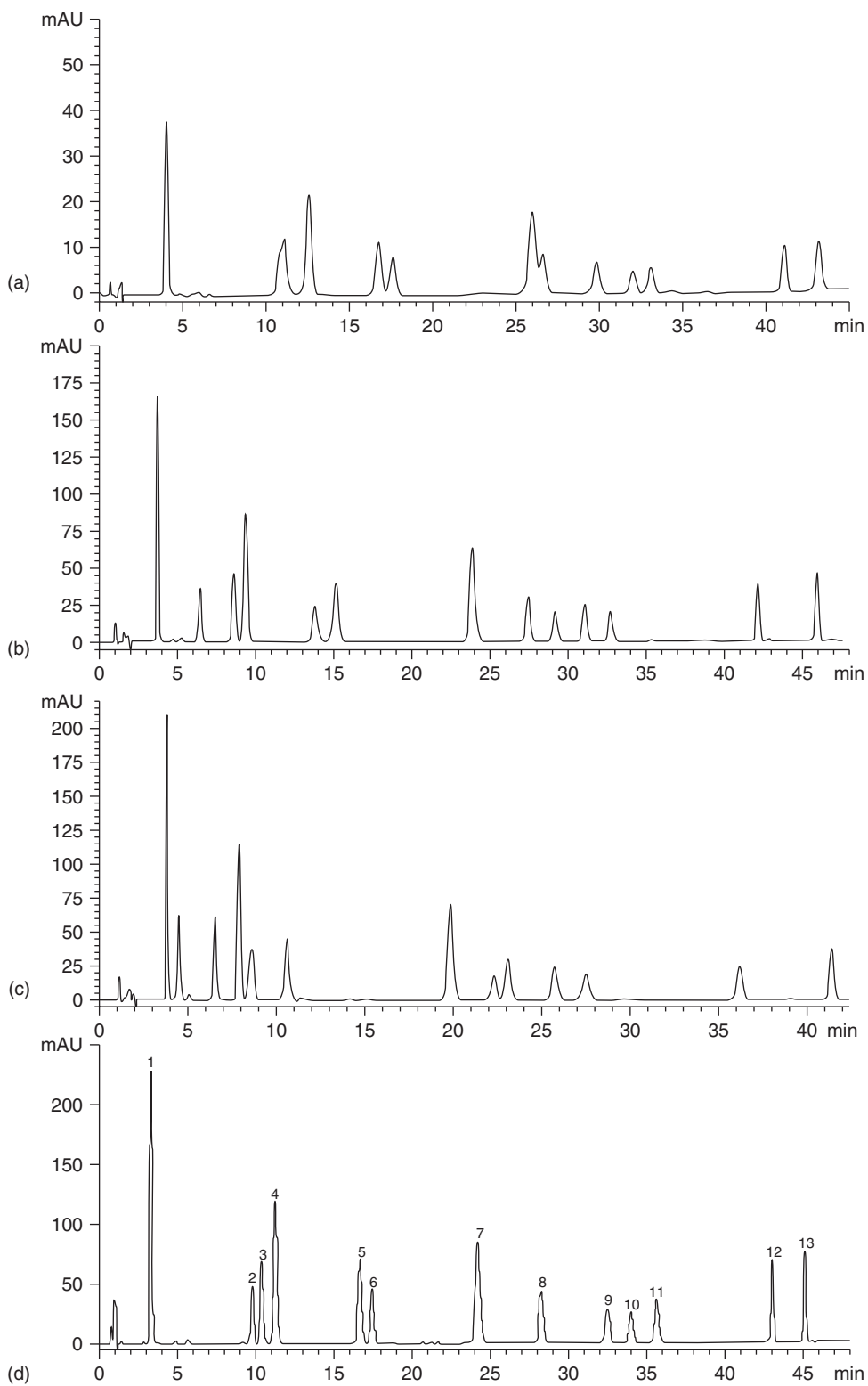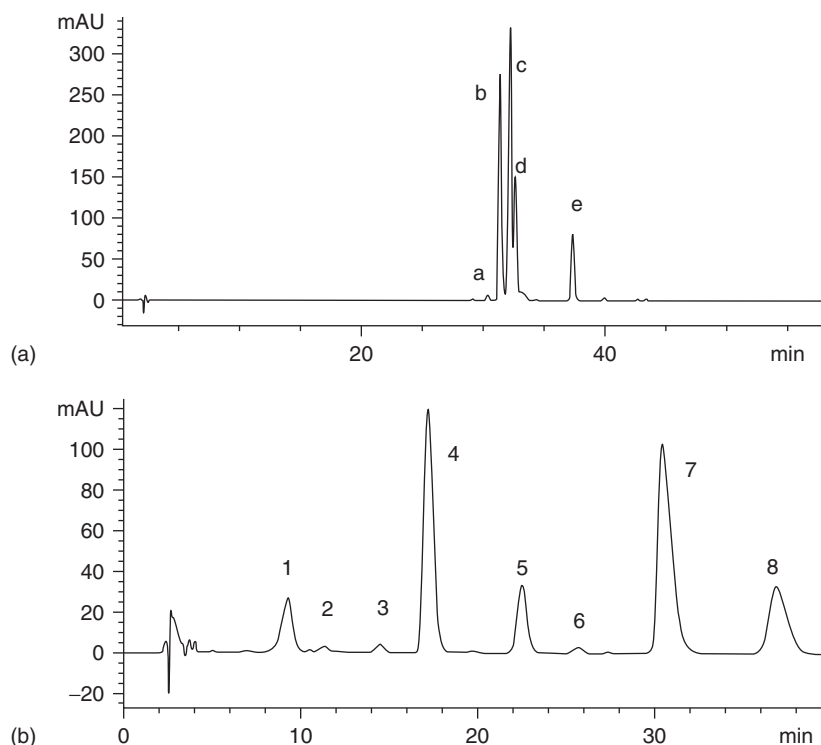
Figure 3 HPLC of a fraction from a toad skin extract. (a) Separation of the fraction into four insufficiently resolved peaks in RP-mode. (b) Separation of the fraction into eight highly resolved peaks in HILIC mode. LC-UV (300 nm) chromatograms on an Xterra MS C18 column (5 μm, 150 mm × 2.1 mm I.D.) under RPLC mode (a) and a Click-CD column (5 μm, 150 mm × 4.6 mm I.D., laboratory-made) under HILIC mode; column temperature: 30 °C. (*Source*: Reproduced with permission from Liu Y, Feng J, Xiao Y, Guo Z, Zhang J, Xue X, Ding J, Zhang X, Liang X, Journal of Separation Science. Copyright © 2010 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.)

interferences in MS detection occur, which was also shown for the use in UPLC (Liu *et al*., 2010a, b). Very rarely, buffered systems containing potassium dihydrogen phosphate or ammonium acetate have been employed (Kanojiya and Madhusudanan, 2012; Gao *et al*., 2010a) but are not applicable in MS detection because of their involatility.

2.1.1.3 *Analysis and Detection*
Depending on the complexity of the sample and of the efficiency of the clean-up, quite long separation times of 40–60 min are common in the analysis of plant and toad extracts, respectively (Figures 4 and 5). Only few methods provide satisfying faster separations (Figure 6) (Huang *et al*., 2009; Ye and Guo, 2005).

The analyses are usually performed at ambient temperature or at slightly increased temperatures up to 40 °C.

UV detection has been the predominant detection mode since the introduction of HPLC in the field of

cardiotonic steroids, and UV–DAD is still the most frequently applied method for the investigation of plant materials (Sun *et al*., 2012; Pellati *et al*., 2009; Roca-Pérez *et al*., 2004). The most well-established wavelengths of detection are 220 nm for cardenolides and 300 nm for bufadienolides. The symmetrical absorption of cardenolides between 217 and 222 nm can be used to identify unknown compounds of this class in complex extracts to support the assignment (Zehnder and Hunter, 2007). Thus, especially for quantifications of cardenolides and bufadienolides, UV detection is an indispensible and robust method (Ye *et al*., 2006).

The first application of liquid chromatography/mass spectrometry (LC/MS) for plant extracts containing cardiotonic glycosides investigated the potential of thermospray (TSP) MS and continuous flow fast atom bombardment (CF-FAB) MS for the dereplication of cardenolides in a root extract of *Nerium odorum* Aiton (Wolfender *et al*., 1995).
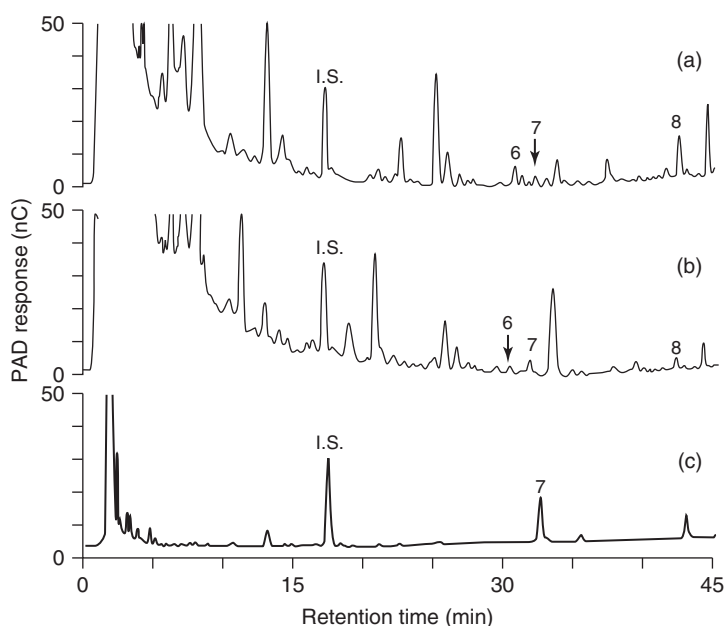
Figure 4 HPLC of *Digitalis purpurea* leaf extracts. (a) *D. purpurea* leaf dried at 60 °C; (b) *D. purpurea* leaf dried at RT; (c) *D. purpurea* seed. Peaks: 6, gitoxin; 7, digitonin; 8, digitoxin; I.S. internal standard. (*Source*: Reprinted from J Pharm Biomed Anal, 54, Kwon HJ, Sim HJ, Lee SI, Lee YM, Park YD, Hong SP, HPLC method validation for Digitalis and its analogue by pulsed amperometric detection, 217–221, Copyright (2011), with permission from Elsevier.)
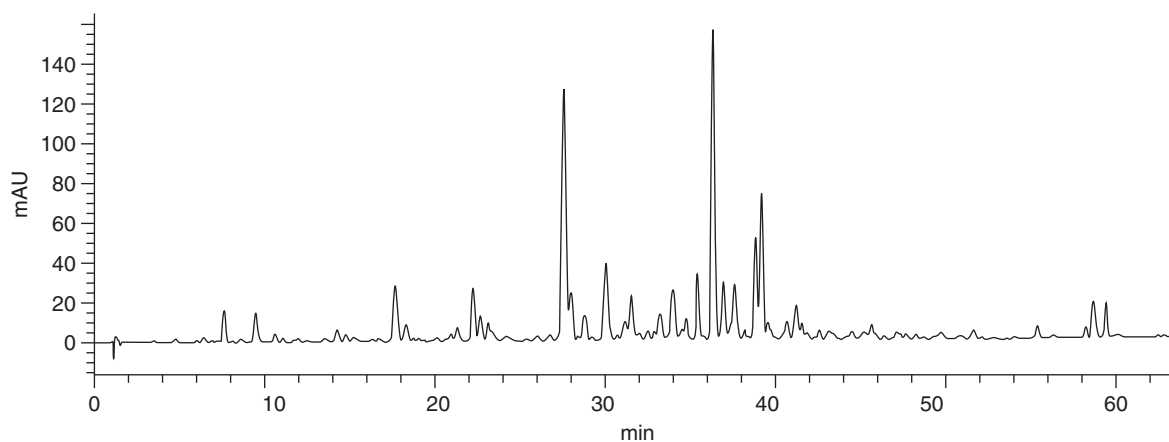
Figure 5 HPLC of an extract of toad skin. LC-UV (300 nm) chromatogram for the crude sample of toad skin on an XTerra MS C18 column (5 μm, 100 mm × 4.6 mm I.D.). (*Source*: Reproduced with permission from Liu Y, Feng J, Xiao Y, Guo Z, Zhang J, Xue X, Ding J, Zhang X, Liang X, Journal of Separation Science. Copyright © 2010 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.)

Despite this study, the hyphenation of LC/MS until now is mainly applied for bufadienolides from toad venoms and only very few applications for plant materials have been published.

In an LC/MS/MS study on the cardenolides in an aqueous extract from *Nerium oleander*, the four major components oleandrigenin, oleandrin, odorside, and neritaloside were tentatively identified by electrospray ionization (ESI) MS and the corresponding quadrupole-quadrupole-time-of-flight QqTOF product-ion spectra (Wang *et al.*, 2000). Calotropagenin glycosides in *Calotropis procera*
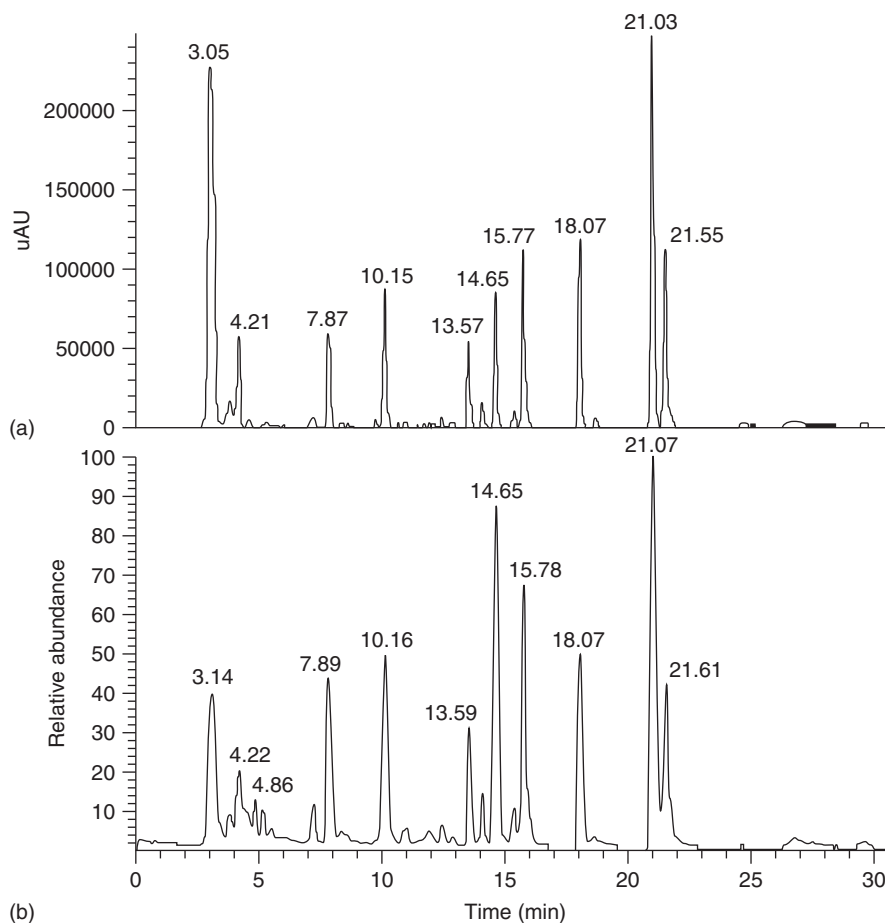
Figure 6 HPLC/DAD/APCI-MS/MS analysis of the methanolic extract of ChanSu crude drug. (a) HPLC/UV chromatogram monitored at 296 nm and (b) APCI-MS total ion current profile of ChanSu. (*Source*: Reproduced with permission from Ye M, Guo D, Rapid Commun Mass Spectrom. Copyright © 2005 John Wiley & Sons, Ltd.)

(Aiton) W.T. Aiton were studied by LC/ESI-MS involving cationization by post-column addition of alkali salts and on-line LC–ESI-MS/MS experiments. This way the genin calotropagenin was established in several glycosides (Kanojiya and Madhusudanan, 2012).

The reason for this rare use of LC/MS in the research of cardiotonic glycosides in plants can easily be explained by the complex pattern of numerous compounds containing the same genin and only differing in the glycoside part (e.g., Kopp *et al.*, 1996). As different hexoses or desoxyhexoses as well as the consecutive linkage of the single sugars in the carbohydrate chains of such compounds cannot be deduced unambiguously from MS data, the identification by LC/MS works for genins but remains tentative for glycosides (Wang *et al.*, 2000; Kanojiya and Madhusudanan, 2012).

In contrast, bufadienolides in toad venoms occur unconjugated as sulfates or suberoyl, suberoylarginine, and pimeloyl arginine esters for which much more unambiguous information can be deduced from LC/MS/MS experiments (Gao *et al.*, 2010a). Therefore, LC/MS hyphenation has achieved more importance for the analyses of bufadienolides. The potential for the identification of these analytes has impressively been shown by identifying 35 different bufadienolides in a methanolic extract of Chan Su by HPLC coupled with atmospheric pressure chemical ionization (APCI)-MS/MS (Ye and Guo, 2005) as well as in the comparison of 17 toad venoms of different origin by LC–DAD–MS/MS resulting in

the identification of 43 analytes (Gao *et al*., 2010a). In both studies, positive ion mode was found more suitable. In atmospheric-pressure chemical ionization (APCI), the positive ion mode was almost two orders of magnitude more sensitive than the negative ion mode (Figure 7) (Ye and Guo, 2005). However, differentiation between hydroxyl substitution at different positions or between a 14β-OH and a $14\beta,15\beta$-epoxy function was not possible without respective standard compounds (Ye and Guo, 2005).

In a similar investigation of toad venom, the potential of the hyphenation of UPLC with electrospray ionization quadrupole time-of-flight mass spectrometry (ESI-Q-TOF-MS) was recently proven resulting in the identification of 19 known bufadienolides and the characterization of 20 putative new compounds (Lui *et al*., 2010b).

LC/MS/MS has been applied not only for qualitative investigations of Chinese medicines (Wong *et al*., 2002) but also for the quantification of bufadienolides by triple quadrupole MS in a preparation used for injection (Wu *et al*., 2012).

The use of hyphenation of LC with NMR is a relatively new approach in the study of cardiac glycosides; until now, all studies combined LC/NMR with HPLC/DAD, high-performance liquid chromatography coupled with diode-array detection and electrospray ionization tandem mass spectrometry (HPLC/DAD/MS^n), and HPLC/HRMS (Clarkson *et al*., 2005; Li *et al*., 2010; Gao *et al*., 2010b). This way LC/NMR has been shown as an excellent tool for the rapid dereplication, identification, and targeted isolation of cardiotonic glycosides.

Combined methods as described were applied for the identification of cardenolides in fractions from the stems of *Periploca forrestii* Schltr. (Li *et al*., 2010) and in different plant parts from *Kanahia laniflora* (Forssk.) R. Br. (Clarkson *et al*., 2005) as well as of bufadienolide sulfates in the toad venom of *Bufo melanosticus* (Gao *et al*., 2010b). Online enrichment of the analytes by SPE after the HPLC separation is a suitable option to reach sufficient sample amounts for the NMR, which is lacking the outstanding sensitivity of MS.

Pulsed amperometric detection (PAD) has been suggested as a highly sensitive and selective method for the accurate and precise microanalysis of cardenolides. Nevertheless, this mode of detection has not gained a lot of interest, but served for the analysis of extracts from *Digitalis purpurea* leaves or seeds

prepared with 50% acetonitrile without further sample preparation (Kwon *et al*., 2011).

2.1.2 Thin-Layer Chromatography (TLC)

TLC and high-performance thin-layer chromatography (HPTLC) are convenient methods for fast screenings of extracts containing cardiotonic steroids. Both methods allow parallel analysis of several samples on one plate, are simple to perform, and are flexible, because different stationary phases and numerous solvent systems are available. By multiple detection, visual results of all components in the extracts are obtained, which is of importance in the identification of cardenolides and bufadienolides because of their complex patterns in plants. For TLC plates, the mean particle size of the stationary phase is 12 µm, whereas for HPTLC plates, usually 5 µm particles are used (Marston, 2007). Thus, much better and faster separations and higher reproducibility are achieved by HPTLC as compared to TLC. Automation in HPTLC is an additional option, which further increases reproducibility significantly.

Although especially automized HPTLC is an excellent tool for semiquantitative and quantitative determinations, hardly any applications for cardiotonic steroids have been published (Ikeda *et al*., 1996) because of the compelling role of HPLC for the quantification of these compounds. In contrast, for qualitative studies of cardiac glycosides in plant materials, the long-standing use of TLC has resulted in the development of standard procedures (Wagner and Bladt, 2009), which are broadly used and also applied in pharmacopeial monographs of the respective herbal drugs (Pharm. Eur., DAB).

2.1.2.1 Sample Preparation

Depending on the content of cardiac glycosides in the investigated plant, the adaption of the amount of drug for the extraction is recommended. Drugs containing considerable proportions of fat, for example, Strophanti semen or Hellebori radix should be defatted with light petroleum before the extraction (Wagner and Bladt, 2009; Holz, 2007).

Usually aqueous ethanol, aqueous methanol, or methanol is used for the extraction of cardenolides and bufadienolides. The removal of accompanying substances can be achieved by precipitation with lead(II)acetate solution. The cardiac glycosides
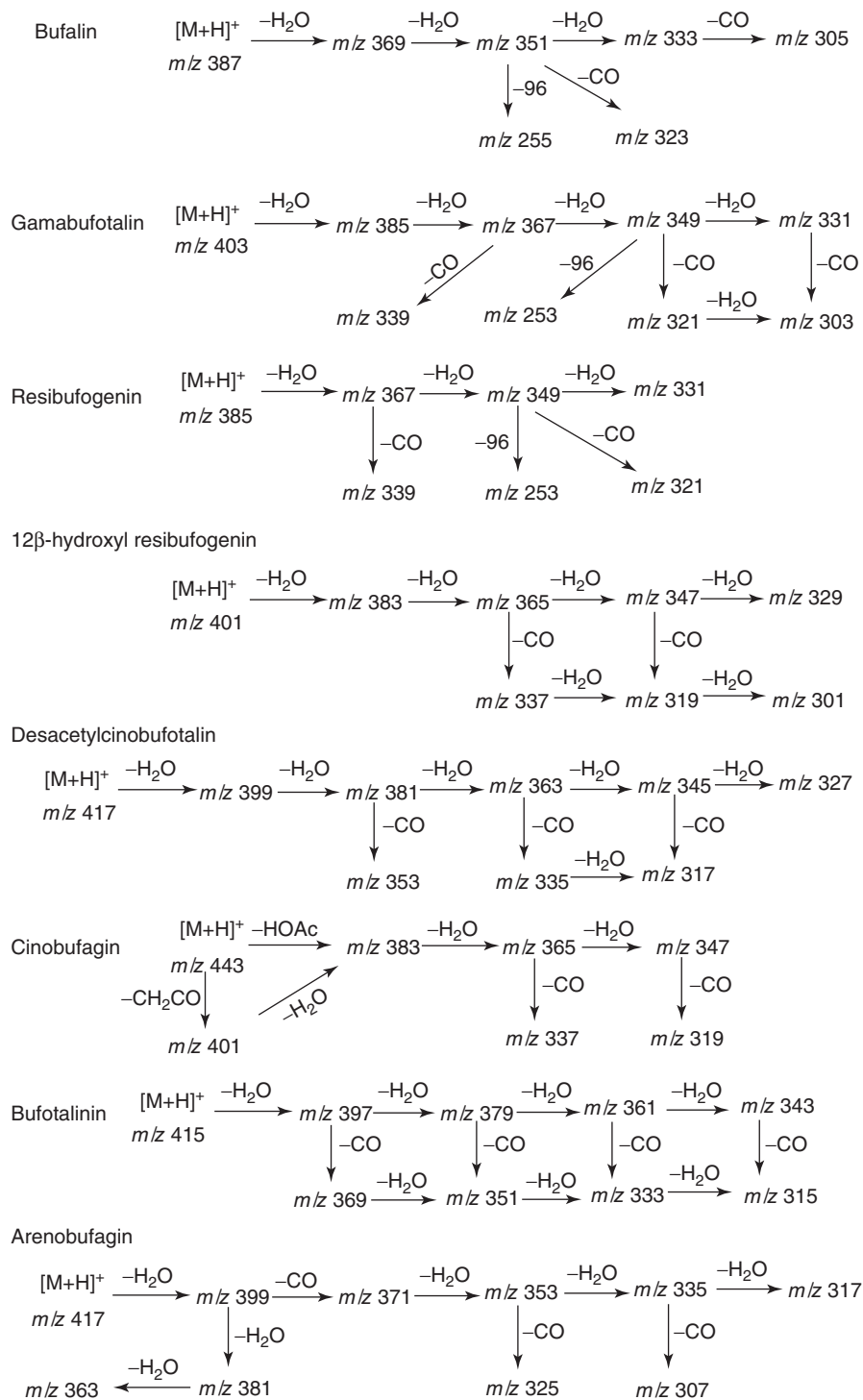
Figure 7 Major fragmentation pathways of protonated bufadienolides in APCI-MS. (*Source*: Reproduced with permission from Ye M, Guo D, Rapid Commun Mass Spectrom. Copyright © 2005 John Wiley & Sons, Ltd.)

are then enriched in organic solvents by partition of the aqueous extract with chloroform or chloroform–isopropanol mixtures (Wagner and Bladt, 2009; Krenn, 2007; Luckner and Diettrich, 2007).

2.1.2.2   Analysis

As stationary phase, almost exclusively silica gel F_{254} plates are in use for TLC and HPTLC of cardiac glycosides (Wagner and Bladt, 2009; Ferth *et al*., 1992; Mathen and Hardikar, 2010). HPTLC analyses on RP-2 F_{254} plates provide additional insight into complex mixtures of medium to high polarity but were performed in rare cases only (Krenn *et al*., 1991b; Kopp *et al*., 1992; Stuhlemmer *et al*., 1993; Ikeda *et al*., 1996). The mobile phases for normal-phase TLC consist mainly of ethylacetate–methanol–water or chloroform–methanol–water mixtures. Among the numerous slightly differing compositions, ethylacetate–methanol–water $(81 + 11 + 8)$ and chloroform–methanol–water $(70 + 22.5 + 3.5)$ can be recommended generally for cardiac glycosides of medium polarity. For highly polar or apolar analytes, the chloroform–methanol–water system can easily be amended according to Table 1. The ethylacetate–methanol–water system can be adapted to strongly polar compounds by increasing the proportion of methanol or adding ethanol (Wagner and Bladt, 2009).

For reversed phase high-performance thin-layer chromatography (RP-HPTLC), methanol–water mixtures containing more than 50% methanol or 0.5 M NaCl–acetonitrile $(1 + 1)$ are recommended. Under use of the latter mobile phase, a method for the quantification of cardenolides in *Digitalis lanata* Ehrh. by HPTLC was established (Ikeda *et al*., 1996).

Table 1   Amendments of TLC solvent system containing chloroform–methanol–water.

Solvent system	Parts chloroform	Parts methanol	Parts water	
1	95	1.5	0.10	apolar
2	90	3.5	0.20	
3	85	8	0.50	
4	80	10	1.0	
5	75	15	1.8	
6	70	22	3.3	
7	65	30	6.0	
8	60	40	10	polar

In a study focusing on the optimization of the separation of toad venom bufadienolides on a hydrophobic gel, for the selection of the mobile phase for preparative purposes, TLC on silica gel F_{254} plates was proven as a suitable tool. Numerous apolar solvent systems containing *n*-hexane or *n*-heptane in combination with dichloromethane, chloroform, or carbontetrachloride and different alcohols were tested and provided decision guidance for the preparative separation of the extract on Sephadex® LH20 (Kamano *et al*., 1999).

2.1.2.3   Detection

A reasonable number of spraying reagents for cardenolides and bufadienolides has been described and sulfuric acid reagents have been proven universal for both groups. Quite characteristic color reactions in daylight are obtained after spraying with vanillin–sulfuric acid reagent, anisaldehyde–sulfuric acid reagent, or 50% ethanolic sulfuric acid for cardenolides and bufadienolides, respectively. Characteristically fluorescent zones can be observed using chloramine–trichloroacetic acid reagent or antimony(III)chloride reagent (Wagner and Bladt, 2009).

For the detection of cardenolides via their butenolide moiety in complex mixtures with high amounts of accompanying substances, Kedde reagent (3% ethanolic 3,5-dinitrobenzoic acid–2 M NaOH $1 + 1$) is advantageous (Wagner and Bladt, 2009). This very specific reaction of the analytes, which results in violet zones under daylight, provides a reasonable simplification of the chromatograms and facilitates assignment.

2.1.3   Gas Chromatography (GC)

Because cardiac glycosides are not volatile, gas chromatography has never reached an important role in the analysis of these compounds in material from plant or animal origin. The method has been applied very rarely for the structure confirmation of cardenolide or bufadienolide genins under use of MS detection.

2.1.3.1   Sample Preparation

Hydrolysis of cardiac glycosides for gas chromatography/mass spectrometry (GC/MS) analysis is usually performed with sulfuric or hydrochloric

acid. Due to the occurrence of decomposition depending on the strength of the acid as well as on the temperature and duration applied in the process, the conditions for the hydrolysis have to be selected carefully. After acid hydrolysis of cardenolides from *D. lanata* with 0.1 M sulfuric acid for 3 h at 60 °C, the concentration of Δ14,15-unsaturated derivatives remained under 1% (Hensel and Kreis, 1997). In contrast, canarigenin from *Isoplexis canariensis* (L.) Lindl. Ex G. Don. was not stable when treated with concentrated hydrochloric acid for 12–14 h at 30 °C resulting in the formation of dianhydroperiplogenin (Schaller and Kreis, 2006). Treatment of thevetins from *Thevetia peruviana* K. Shum. with 10 M hydrochloric acid for 2 h at 70 °C resulted in several artifacts including Δ8,14- and Δ14,15-anhydrocardenolides (Kohls *et al.*, 2012). After hydrolysis, the genins have to be extracted into dichloromethane or chloroform–isopropanol mixtures.

For the derivatization of nonvolatile aglycones, acetylation or trimethylsilylation has been applied successfully (Hensel and Kreis, 1997; Yamaki *et al.*, 1999; Schaller and Kreis, 2006; Kohls *et al.*, 2012). In one study, the relatively apolar bufadienolides resibufogenin, bufalin, and cinobufagin from a West Indian aphrodisiac and the traditional Chinese medicine Chan Su prepared from the skin venom of Chinese toads were separated by GC without derivatization (Barry *et al.*, 1996).

2.1.3.2 *Analysis and Detection*
In the investigation of Barry *et al.*, a 100% dimethylpolysiloxane megabore column (DB-1; 15 m × 0.53 mm; 0.25 µm film thickness) had to be used because underivatized bufadienolides were not eluted from the respective capillary column.

In the most recent study, the separation of a mixture of six compounds from yellow oleander was achieved on a 5% diphenyl–95% dimethylsiloxane column (DB-5HT; 30 m × 0.25 mm; 0.1 µm film thickness; Kohls *et al.*, 2012), which was comparable to the column used for the separation of the four major cardenolides from *D. lanata* or for the identification of cardenolides in *I. canariensis* (HP-5MS; 30 m × 0.25 mm; Hensel and Kreis, 1997; Schaller and Kreis, 2006).

Mass spectrometry was the method of choice for the detection in all these investigations. The identification of the analytes may be performed by electron ionization-mass spectrometry (EI-MS) and comparison to reference compounds (Barry *et al.*, 1996; Schaller and Kreis, 2006). Detection in the SIM mode allows a very selective identification of cardenolides without interferences of the matrix (Hensel and Kreis, 1997; Yamaki *et al.*, 1999). Digitalis cardenolides can be identified focusing on the characteristic mass fragments for acetates of 12β-hydroxy cardenolides (digoxigenin, digitonatigenin) at *m/z* 201 and of digitoxigenin and gitoxigenin without a 12β-hydroxy group at *m/z* 203 (Figure 8).

Combination of EI-MS, EI-MS–MS, chemical ionization mass spectrometry (CI-MS), and high resolution mass spectrometry (HR-MS) was shown to be a powerful tool for the structure elucidation of cardenolides with 18,20-oxido-20,22-dihydro functionalities. By the formation of artifacts under rigorous hydrolytic conditions and detailed MS studies of the artifacts, the structures of thevetin C and acetylthevetin C were unambiguously elucidated (Kohls *et al.*, 2012).

2.2 Capillary Electrophoresis (CE)

Capillary electrophoresis (CE) with different modes of operation such as capillary zone electrophoresis (CZE), micellar electrokinetic chromatography (MEKC), microemulsion electrokinetic chromatography (MEEKC), capillary electrochromatography (CEC), or non-aqueous capillary electrophoresis (NACE) is widely recommended as complementary analytical tool for the analysis of plant secondary metabolites due to high separation efficiency, short separation times, very low consumption of sample and reagents, cheap columns, and low consumption or avoidance of organic solvents (Gotti, 2011; Chen *et al.*, 2012). The major disadvantage of low sensitivity in spectrophotometric detection could be resolved by coupling with, for example, MS detection (Ramautar *et al.*, 2011).

The versatility of CE is based on not only the different modes of separation but also the selection of capillaries, buffer type and concentration, pH of buffer, voltage, capillary temperature, micelle concentration and nature, addition of modifiers, and mode of detection. Thus, numerous methods for a broad range of secondary plant metabolites (hydrophilic, hydrophobic, chargeable, nonchargeable, etc.) have been published and reviewed (Ganzera, 2008; Unger, 2009; Yang *et al.*, 2010;
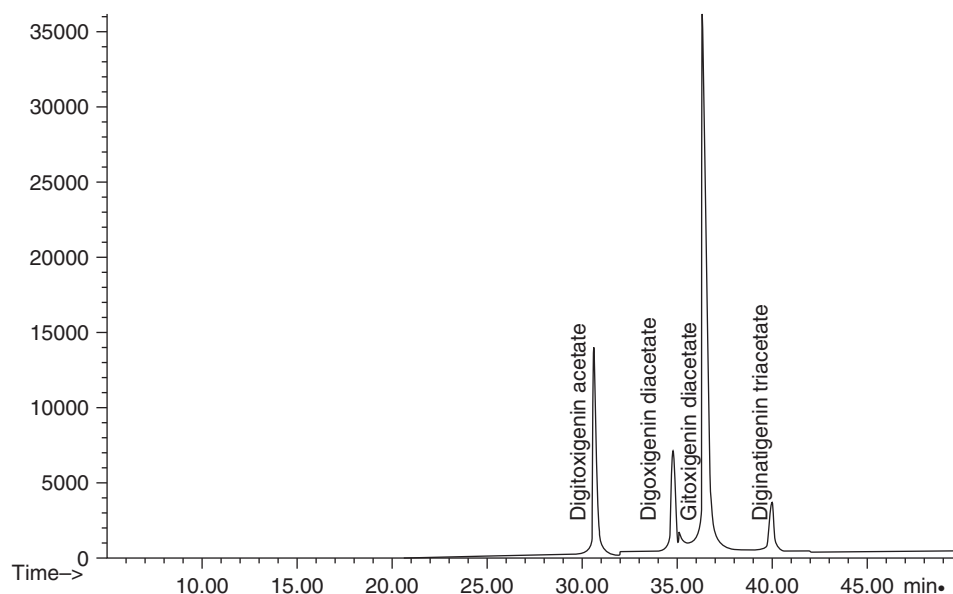
Figure 8 GLC–MS of four cardenolide acetates. (*Source*: Reprinted from Pharm Acta Helv, 72, Hensel A, Kreis W, GLC-MS investigations on cardenolide genins, 243–246, Copyright (1997), with permission from Elsevier.)

Gotti, 2011). Addition of chiral selectors such as cyclodextrins has proven the potential of CE for the separation of chiral plant compounds as well. Despite the wide applicability in the field of secondary plant metabolites, most published CE methods focus on the qualitative and quantitative analyses of alkaloids and phenolics. Only very few methods for the analysis of cardiotonic glycosides in plants have been described and all of those used UV detection.

By MEKC under use of a polyimide-clad fused-silica capillary (50 μm I.D.), primary and secondary cardenolides from *D. lanata* were separated. As buffer systems served 22.5 mM tetraborate (pH 9.3)–37.5 mM sodiumdodecylsulfate (SDS)–7 M urea for primary cardiac glycosides and 30 mM tetraborate (pH 9.3)–50 mM SDS–10 mM γ-cyclodextrin for secondary cardiac glycosides. The resolution of the two anomeric glycosides glucodigifucoside and glucodigiglucomethyloside was achieved under use of 30 mM tetraborate (pH 9.3)–50 mM SDS–20 mM γ-cyclodextrin (Gaus *et al*., 1993).

Deslanoside, digoxin, acetyldigoxin, and acetyldigitoxin were separated by MEKC on untreated fused-silica capillaries and 20 mM lithium tetraborate (pH 6)–35 mM SDS–7 M urea. In this study, the superiority of microemulsion electrokinetic chromatography (MEEKC) under use of microemulsions containing 0.8–1.7% SDS, 0.2–0.8% heptane, and 6.6% 1-butanol in 50 mM sodiumtetraborate (pH 9.2) for the fast separation of the four analytes was demonstrated (Figure 9) (Debusschére *et al*., 1997).

A very similar MEECK method was suggested for the determination of bufadienolides in chloroform extracts of toad venom and respective traditional Chinese medicines using 3.3% SDS, 0.8% heptane, and 6.6% 1-butanol in 10 mM sodiumtetraborate (pH 9.2) (Su *et al*., 2007).

2.3 Spectrophotometry

To replace the toxicity assay in guinea pigs, which was used for long for the determination of cardiotonic glycosides in herbal drugs, quantifications of these analytes by spectrophotometry were suggested. By the respective methods, the total content of cardenolides is usually determined after reaction with Baljet reagent or Kedde reagent and the one of bufadienolides after conversion with methanolic KOH. For *Convallariae herba*, a good correlation between the toxicity assay and the spectrophotometrically quantified total content of cardenolides was shown (Koehler *et al*., 2007). For the only drug
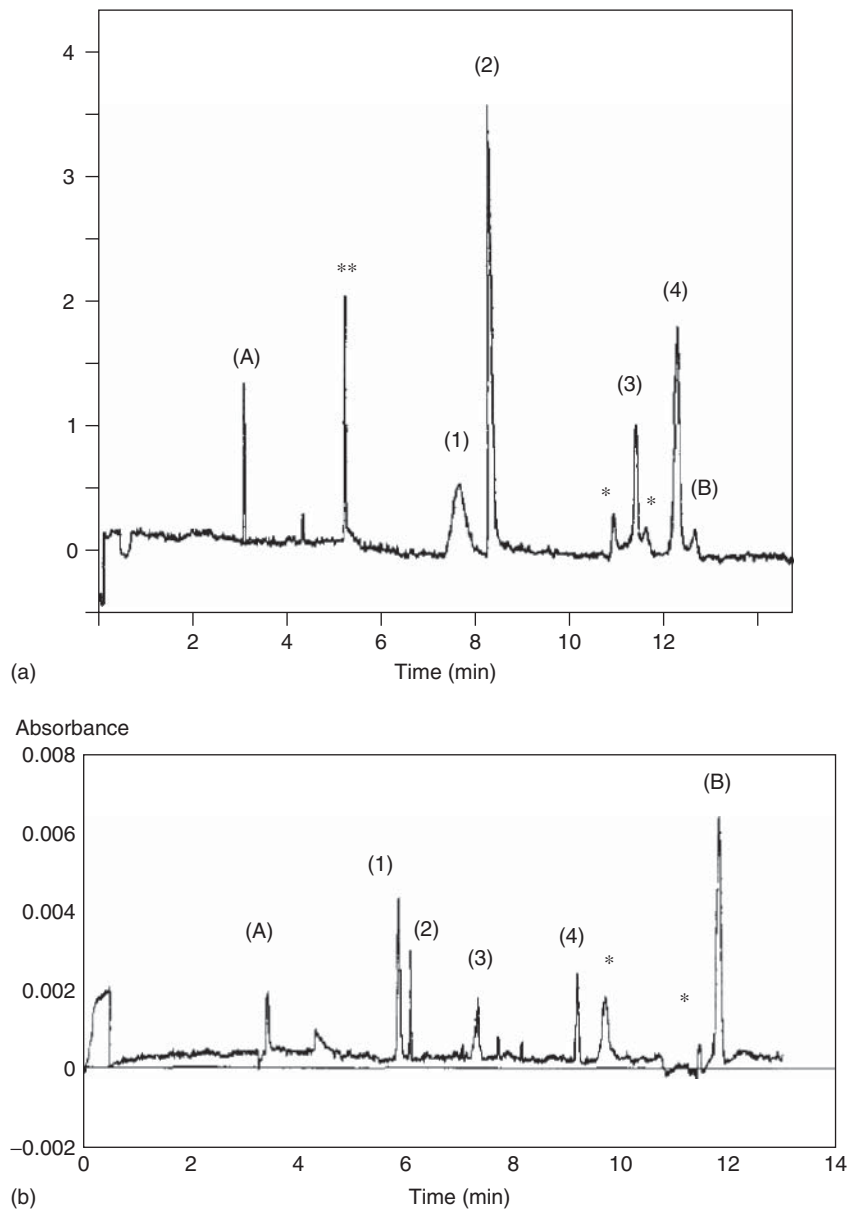
Figure 9 MEKC (above) and MEEKC (below) of four cardiac glycosides. Separation of four cardiac glycosides in MEKC in presence of urea. Electrolyte: $Li_2B_4O_7$ 0.02 mol l^{-1} pH = 6.0, SDS 0.035 mol l^{-1}, urea 7 mol l^{-1}. Fused-silica capillary: $L = 44$ cm, $l = 36$ cm, I.D. = 50 μm. $V = 30$ kV, $T = 25°C$, UV detection at 220 nm. (a) Methanol: electroosmotic flow marker; (b) Sudan III: micelle marker. (1) Digoxin, (2) deslanoside, (3) acetyldigoxin, (4) acetyldigitoxin. *, Impurities; **, Sudan III impurity. Separation of four cardiac glycosides in MEEKC. Microemulsion containing 92.14% borate buffer ($Na_2B_4O_7$ 0.05 mol l^{-1}, pH 9.2), 0.83% SDS, 0.42% heptane, and 6.61% 1-butanol; fused-silica capillary: $L = 47$ cm, $l = 40$ cm, I.D. = 50 μm. $V = 25$ kV, $T = 25°C$, UV detection at 220 nm. (a) Methanol; (b) dodecylbenzene; (1) deslanoside, (2) digoxin, (3) acetyldigoxin, (4) acetyldigitoxin. *, Impurities. (*Source*: Reprinted from J Chromatogr A, 779, Dubusschére L, Demesmay C, Rocca JL, Lachatre G, Lofti H, Separation of cardiac glycosides by micellar electrokinetic chromatography and microemulsion electrokinetic chromatography, 227–233, Copyright (1997), with permission from Elsevier.)

in the European Pharmacopoeia containing cardiac glycosides, namely digitalis leaf, such a spectrometric assay is required as well. The advantage of such methods is the simpler equipment needed as compared to respective HPLC methods. Nevertheless, the sample preparation for spectrophotometric quantifications is very laborious and time-consuming and because of numerous clean-up steps, the determination is prone to interferences and methodical errors.

2.3.1 Sample Preparation

Usually herbal drugs are extracted with aqueous ethanol or aqueous methanol for spectrophotometry. In contrast, in the quantification of digitalis leaf according to the European Pharmacopoeia, extraction with water is required. From extracts of above ground parts (Digitalis leaf, Convallaria herb, Oleander leaf, etc.) accompanying compounds have to be removed by precipitation with lead acetate. In several modifications of the assay, lead excess is eliminated with disodiumhydrogenphosphate. After filtration, whether the glycosides (Krenn *et al.*, 1996) or the genins after acid hydrolysis (Pharm. Eur. 7.5) are extracted from the aqueous phase with chloroform or chloroform–*n*-butanol.

For the determination of the total bufadienolide content in Squill bulbs, the sample preparation has to be modified because of the high amounts of polysaccharides and anthocyanins in the extracts: removal of polysaccharides from the aqueous methanolic extract can be achieved by precipitation with butanone–*n*-propanol (1 + 1). For the elimination of the phenolics, ion exchange chromatography with Dowex® MSA 1 is an option (Kopp *et al.*, 1990).

After evaporation, the purified cardenolide or bufadienolide concentrates are ready for indirect spectrophotometric determination.

The described mode of enrichment of cardiotonic glycosides is not only applicable to plant material but can as well be employed in investigations of insects (Dobler *et al.*, 1998), although ethanolic extracts of butterflies could be analyzed without purification as well (Moranz and Brower, 1998).

2.3.2 Derivatization for Spectrophotometry

For derivatization with respective reagents, the purified extracts are dissolved in methanol or 50% ethanol.

In the determination of cardenolides, almost exclusively alkaline solutions of picric acid or 3,5-dinitrobenzoic acid are used, and both reagents react with the butenolide ring to colored Meisenheimer complexes. The reaction is based on the activated methylene group in the ring. The use of 2,2′,4,4′-tetranitrophenyl for the spectrophotometry of cardiotonic glycosides (Harry-O'kuru and Abbot, 1997) has not been accepted as an alternative.

As bufadienolides do not contain an activated methylene group, these cardiotonic glycosides are quantified after reaction with methanolic KOH (Kopp *et al.*, 1990). By addition of the reagent, the cumalinring is cleaved under generation of the respective methylester. The UV maximum of the methylester shows a bathochromic shift from 300 to 355 nm and the absorption is increased to approximately the eightfold as compared to the uncleaved bufadienolide.

Due to the cumbersome sample preparation and derivatization, the spectrophotometric quantification of the total cardiac steroid content in herbal drugs has almost completely been replaced by HPLC methods determining the amounts of single components.

2.4 Immunoassays

The cardenolides digoxin and digitoxin are the most important cardiotonic glycosides in the therapy of congestive heart failure. Their narrow therapeutic index between 0.8–2.0 ng ml^{-1} and 8–30 ng ml^{-1}, respectively (Hallbach, 2009), requires a close-mesh drug monitoring in which chromatographic techniques or immunoassays can be applied.

On the other hand, the important role of endogenous cardiotonic steroids such as ouabain and marinobufagenin in the control of blood pressure, salt metabolism, as well as in affecting the proliferation and differentiation of heart and smooth muscle cells among other functions has been confirmed in recent years (Schoner and Scheiner-Bobis, 2007). For the detailed analysis of these compounds and their fate in the body, sensitive and precise methods are an indispensable prerequisite as well.

Although HPLC–MS is the most important confirmatory method in the quantification of cardiac glycosides in biological fluids, in clinical routine, immunoassays are an indispensable well-established, rapid, and cost-effective option. These assays can be

run fully automated, which is a significant advantage in daily routine work. While for the cardenolides digoxin and digitoxin numerous kits are available, commercial assays for bufadienolides are still missing. Only recently, a patent application for respective immunoassays for marinobufagenin has been filed (Puschett *et al*., 2011).

Radioimmunoassays were the first immunoassays employed for the determination of cardiotonic steroids in drug monitoring and proved of value for sensitivity and speed. Due to the necessary special precautions in the use of radiolabeled compounds, radioimmunoassays (RIA) have been completely replaced by enzyme immunoassays (EIA). Several systems such as homogeneous competitive ones, for example, (EMIT®) or heterogeneous noncompetitive ones such as enzyme-linked immunosorbent assays (ELISA®) are available. Other options are homogeneous competitive FPIA as well as chemiluminescence (CLIA), recombinant enzyme (CEDIA®), or particle enhanced turbidimetric inhibition (TIA) kits.

Besides their essential role in medicinal chemistry, such immunoassays have rarely been applied for studies of plant material but have never reached the importance of chromatographic methods in this field.

In a study of plant-induced poisoning of livestock, an FPIA was successfully applied for the determination of cardiac glycosides in the suspected plants as well. In different members from the plant families Iridaceae, Hyacinthaceae, Crassulaceae, Santalaceae, and Apocynaceae, bufadienolides and cardenolides were quantified with this digoxin kit. From the cross-check with plants not containing cardiac glycosides, the authors deduced the suitability of the assay. Although the sample preparation for immunoassays usually is simpler as compared to chromatographic methods, in this examination, a quite laborious method including hydrolysis of the glycosides and partition was applied (Schultz *et al*., 2005). A similar assay was applied for Australian poisonous plants under very simple sample preparation. On the basis of the detection of cardiac steroids in 27 different species, this assay was suggested as a suitable tool in the fast confirmation of cardiac glycoside poisoning of stock (Radford *et al*., 1994).

Samples of *D. lanata* from *in vitro* cultures (available only at very small amounts and partly containing very low amounts of cardiac glycosides) were analyzed by an ELISA after very simple sample preparation. Hairy roots were extracted with 70% ethanol by sonification. After centrifugation and dilution with phosphate-buffered saline containing 0.1% casein to a concentration of less than 10% ethanol, the amount of cardiac glycosides was measured (Yoshimatsu *et al*., 1995).

The most important limitation of immunoassays is the vulnerability to interferences that have been proven in studies performed in the field of therapeutic drug monitoring (Dasgupta, 2012).

Therefore, several studies were also performed with plant material to determine interferences of immunoassays with herbal drugs or botanicals.

The interference of *Uzara* preparations containing uzarigenin (the 5α-epimer of digitoxigenin) was shown not only by spiking blank serum with the preparation but also in healthy volunteers under use of a recombinant enzyme assay for digitoxin and a turbidimetric monoclonal assay for digoxin (Thürmann *et al*., 2004). Such an approach allows an adequate estimation of interferences, whereas the very simple way of testing plant extracts added to serum samples applied in some other investigations has to be questioned:

In a recent examination of a new luminescent oxygen channeling digoxin assay for interferences with herbal supplements, different hawthorn or Indian ginseng (*Withania somnifera* (L.) Dunal Prodromus; ashwagandha) preparations were added to drug-free serum as well as to pooled serum of patients receiving digoxin. For one hawthorn and one ashwagandha brand, statistically but not clinically significant interferences were seen with the used assay (Dasgupta *et al*., 2012), which was in contrast to two earlier studies detecting clinically significant interferences with preparations of the two plants (Dasgupta *et al*., 2008, 2010). In similar investigations by the same group, the influence of extracts of oleander or different ginseng species has been compared in various kits (FPIA, MEIA, CLIA, and TIA) (Dasgupta and Datte, 2004; Dasgupta and Reyes, 2005). Although the approach to add plant extracts to drug-free serum or pooled serum of patients receiving digoxin allows to compare between different assays, the major disadvantage of such studies is that the experimental setup does not consider the metabolism of the studied herbal supplements at all. In case of plants containing cardiac glycosides, a similar fate of the active compounds in the organism as compared to digoxin can be assumed. However, to deduce interferences with digoxin assays after oral intake of plants such as

ginseng, hawthorn, or ashwagandha are not justified under such an experimental setup.

3 CONCLUSION

For qualitative and quantitative determinations of the complex patterns of cardiotonic glycosides in plants, the application of chromatographic methods is the very best choice. For qualitative screenings, TLC provides fast results, and under use of different spraying reagents and available standard compounds, hints on the structures of the analytes can be collected as well.

For quantitative purposes, HPLC is the optimum method. Owing to its excellent selectivity, complex patterns can be separated sufficiently, and UV detection/DAD detection allows the precise and robust quantification of the single analytes. For both TLC and HPLC, several standard procedures have been established, which facilitate the selection of stationary and mobile phases and the development of a respective method.

Alternative methods such as GC or CE under the use of different separation modes might be applied but are not competitive. The quantification of the total content of cardiotonic glycosides by spectrophotometry becomes more and more outdated but might be an option in case HPLC analyses are not possible and only simple equipment is available. The use of immunoassays does not seem suitable for a broader role in plant research.

Hyphenated techniques such as LC/MS/MS or LC/SPE/NMR are excellent tools for fast structure elucidation and dereplication of cardiotonic glycosides.

REFERENCES

Abe, F., Yamauchi, T., and Minato, K. (1996) *Phytochemistry*, **42**, 45–49.

Ahnhoff, M., Magnar, E., Lagerström, P. O., *et al*. (1985) *J. Chromatogr.*, **340**, 73–138.

Bagrov, A. Y., Shapiro, J. I., and Fedorova, O. V. (2009) *Pharmacol. Rev.*, **61**, 9–38.

Barry, T. L., Petzinger, G., and Zito, S. W. (1996) *J. Forensic Sci.*, **41**, 1068–1073.

Baumgarten, G. (1963) *Die herzwirksamen Glykoside*, VEB Thieme, Leipzig, vol. **1–9**.

Castle, M. C. (1975) *J. Chromatogr.*, **115**, 437–445.

Chen, X., Zhao, J., Wang, Y., *et al*. (2012) *Electrophoresis*, **33**, 168–179.

Clarkson, C., Staerk, D., Hansen, S. H., *et al*. (2005) *Anal. Chem.*, **77**, 3547–3553.

Dasgupta, A., and Datte, P. (2004) *Ther. Drug Monit.*, **26**, 658–663.

Dasgupta, A., Johnson, M. J., and Wahed, A. (2012) *J. Clin. Lab. Anal.*, **26**, 227–231.

Dasgupta, A., Kidd, L., Poindexter, B., *et al*. (2010) *Arch. Pathol. Lab. Med.*, **134**, 1188–1192.

Dasgupta, A., Tso, G., and Wells, A. (2008). *J. Clin. Lab. Anal.*, **22**, 295–301.

Dasgupta, A., and Reyes, M. (2005) *Am. J. Clin. Pathol.*, **124**, 229–236.

Dasgupta, A. (2012) *Resolving Erroneous Reports in Toxicology and Therapeutic Drug Monitoring: A Comprehensive Guide*, John Wiley & Sons, Inc., New York, pp. 237–263.

Debusschére, L., Demesmay, C., Rocca, J. L., *et al*. (1997) *J. Chromatogr.*, **779**, 227–233.

Dobler, S., Daloze, D., and Pasteels, J. M. (1998) *Chemoecology*, **8**, 111–118.

El-Askary, H., Hölzl, J., Hilal, S., *et al*. (1995) *Phytochemistry*, **38**, 1181–1184.

Enson, J. M., and Seiber, J. N. (1978) *J. Chromatogr.*, **148**, 521–527.

Ferth, R., Baumann, A., Robien, W., *et al*. (1992) *Z. Naturforsch. B*, **47**, 1444–1458.

Ganzera, M. (2008) *Electrophoresis*, **29**, 3489–3503.

Gao, H., Popescu, R., Kopp, B., *et al*. (2011) *Nat. Prod. Rep.*, **28**, 953–969.

Gao, H., Zehl, M., Leitner, A., *et al*. (2010a) *J. Ethnopharmacol.*, **131**, 368–376.

Gao, H., Zehl, M., Kählig, H., *et al*. (2010b) *J. Nat. Prod.*, **73**, 603–608.

Gaus, H. J., Treumann, A., Kreis, W., *et al*. (1993) *J. Chromatogr.*, **635**, 319–327.

Gotti, R. (2011) *J. Pharm. Biomed. Anal.*, **55**, 775–801.

Grimm-Bursch, S. (1990) *Cardenolides: structures of C-23 glycosides and genins*. Diploma Thesis. University of Vienna.

Hallbach, J. (2009) Cardiac glycosides, in *Clinical Toxicological Analysis: Procedures, Results, Interpretation*, ed. W. R. Külpmann, Wiley-VCH Verlag GmbH & Co., Weinheim, 327–338.

Harry-O'kuru, R. E., and Abbot, T. P. (1997) *Ind. Crops Prod.*, **7**, 53–58.

Hensel, A., and Kreis, W. (1997) *Pharm. Acta Helv.*, **72**, 243–246.

Holz, W. (2007) Helleborus, in *Hagers Enzyklopädie der Drogen und Arzneistoffe*. 6th edn. Vol **8** Gle-Iot. ed. W. Blaschek, S. Ebel, E. Hackenthal, U. Holzgrabe, K. Keller, J. Reichling, V. Schulz, Wiss VerlagsgesmbH, Stuttgart, 387–400.

Huang, H., Liang, M., Luo, H., *et al*. (2009) *Chem. Res. Chin. Univ.*, **25**, 801–806.

Ikeda, Y., Fujii, Y., Umemura, M., *et al*. (1996) *J. Chromatogr.*, **746**, 255–260.

Jurenitsch, J., Kopp, B., Bamberg-Kubelka, E., *et al*. (1982) *J. Chromatogr.*, **240**, 235–242.

Kamano, Y., Nogawa, T., Kotake, A., *et al*. (1999) *J. Liq. Chromatogr. Rel. Technol.*, **22**, 2455–2465.

Kanojiya, S., and Madhusudanan, K. P. (2012) *Phytochem. Anal.*, **23**, 117–125.

Koehler, H., Kopp, B., and Loew, D. (2007) Convallaria, in *Hagers Enzyklopädie der Drogen und Arzneistoffe*. 6th edn. Vol. **5** Cod-Dig. (ed. W. Blaschek, S. Ebel, E. Hackenthal, U. Holzgrabe, K. Keller, J. Reichling, V. Schulz), Wiss VerlagsgesmbH, Stuttgart, 134–150.

Kohls, S., Scholz-Böttcher, B. M., Teske, J., *et al*. (2012) *Phytochemistry*, **75**, 114–127.

Kopp, B., Krenn, L., Draxler, M., *et al*. (1996) *Phytochemistry*, **42**, 513–522.

Kopp, B., Krenn, L., and Jurenitsch, J. (1990) *DAZ*, **130**, 2175–2180.

Kopp, B., Krenn, L., Kubelka, E., *et al*. (1992) *Phytochemistry*, **31**, 3195–3198.

Krenn, L. (2007) Urginea, in *Hagers Enzyklopädie der Drogen und Arzneistoffe*. 6th edn. Vol **16** Tri-Zyg. (ed. W. Blaschek, S. Ebel, E. Hackenthal, U. Holzgrabe, K. Keller, J. Reichling, V. Schulz), Wiss VerlagsgesmbH, Stuttgart, 331–356.

Krenn, L., Kopp, B., Wallner, E., *et al*. (1991a) *Planta Med.*, **57**(Suppl.) A682.

Krenn, L., Ferth, R., Robien, W., *et al*. (1991b) *Planta Med.*, **57**, 560–565.

Krenn, L., Kopp, B., Bamberger, M., *et al*. (1993) *Nat. Prod. Lett.*, **3**, 139–143.

Krenn, L., Kopp, B., Speta, F., *et al*. (2001) *Stapfia*, **75**, 101–120.

Krenn, L., and Kopp, B. (1998) *Phytochemistry*, **48**, 1–29.

Krenn, L., Schlifelner, L., Stimpfl, T., *et al*. (1996) *Pharmazie*, **51**, 906–909.

Kwon, H. J., Sim, H. J., Lee, S. I., *et al*. (2011) *J. Pharm. Biomed. Anal.*, **54**, 217–221.

Li, Y., Wu, X., Li, J., *et al*. (2010) *J. Chromatogr. B*, **878**, 381–390.

Liu, Y., Feng, J., Xiao, Y., *et al*. (2010a) *J. Sep. Sci.*, **33**, 1487–1494.

Liu, Y., Xiao, Y., Xue, X., *et al*. (2010b) *Rapid Commun. Mass Spectrom.*, **24**, 667–678.

Luckner, M., Diettrich, B. (2007) Digitalis, in *Hagers Enzyklopädie der Drogen und Arzneistoffe*. 6th edn. Vol. **5** Cod-Dig. (ed. W. Blaschek, S. Ebel, E. Hackenthal, U. Holzgrabe, K. Keller, J. Reichling, V. Schulz), Wiss VerlagsgesmbH, Stuttgart, 1025–1051.

Marston, A. (2007) *Phytochemistry*, **68**, 2785–2797.

Mathen, C., and Hardikar, B. (2010) *J. Exp. Ther. Oncol.*, **8**, 177–185.

Mayerhofer, L. (1995) *Bufadienolides and cardenolides: review concerning structures of glycosides and genins*. Diploma Thesis. University of Vienna.

Mijatovic, T., Dufrasne, F., and Kiss, R. (2012) *Curr. Med. Chem.*, **19**, 627–646.

Moranz, R., and Brower, L. P. (1998) *J. Chem. Ecol.*, **24**, 905–932.

Mordasini, M. R., Krähenbühl, S., and Schlienger, R. G. (2002) *Swiss Med. Wkly*, **132**, 506–512.

Newman, R. A., Yang, P., Pawlus, A. D., *et al*. (2008) *Mol. Interv.*, **8**, 36–49.

Pellati, F., Bruni, R., Bellardi, M. G., *et al*. (2009) *J. Chromatogr. A*, **1216**, 3260–3269.

Prassas, I., and Diamandis, E. P. (2008) *Nat. Rev.*, **7**, 926–935.

Puschett, J. B., Romo, D., Berghman, L. R., Abi-Ghanem, D. A., Lai, X. (2011) *Method for determination of marinobufagin levels and compounds employable in such method*. U.S. Pat. Appl. Publ. US 20110207154 A1 20110825.

Radford, D. J., Cheung, K., Urech, R., *et al*. (1994) *Austr. Vet. J.*, **71**, 236–238.

Ramautar, R., Mayboroda, O. A., Somsen, G. W., *et al*. (2011) *Electrophoresis*, **32**, 52–65.

Ritz, E., and Schoner, W. (2008) *Dtsch Med Wochenschr*, **133**, 2690–2694.

Roca-Pérez, L., Boluda, R., Gavidia, I., *et al*. (2004) *Phytochemistry*, **65**, 1869–1878.

Schaller, F., and Kreis, W. (2006) *Planta Med.*, **72**, 1149–1156.

Schoner, W., and Scheiner-Bobis, G. (2007) *Am. J. Physiol. Cell Physiol.*, **293**, 509–536.

Schoner, W., and Scheiner-Bobis, G. (2009) *Dtsch Med Wochenschr*, **134**, 632–636.

Schultz, R. A., Kellerman, T. S., and van den Berg, H. (2005) *Onderstepoort J. Vet. Res.*, **72**, 189–201.

Shimada, K., Hasegawa, M., Hasebe, K., *et al*. (1976) *Chem. Pharm. Bull.*, **24**, 2995–3000.

Song, S. K., Tsui, S. K., and Kwan, S. Y. (2002) *J. Pharm. Biomed. Anal.*, **30**, 161–170.

Stuhlemmer, U., Kreis, W., Eisenbeiss, M., *et al*. (1993) *Planta Med.*, **59**, 539–545.

Su, Q., Wang, Y., and Luo, X. (2007) *Bejing Zhongyiyao Daxue Xuebbao*, **30**, 203–205.

Sun, J., Xiao, J., Wang, X., *et al*. (2012) *Biotechnol. Lett.*, **34**, 563–569.

Teuscher, E., Lindequist, U. (2010) *Biogene Gifte*. 3rd edn. Wiss VerlagsgesmbH, Stuttgart, 230–266.

Thürmann, P. A., Neff, A., and Fleisch, J. (2004) *Int. J. Clin. Pharmacol. Ther.*, **42**, 281–284.

Tittel, G. (1986) Application of HPLC–DAD instead of the guinea-pig test for the quantitative determination of cardiac glycosides, in *Cardiac Glycosides* 1785–1985, Biochemistry-Pharmacology-Clinical Relevance, (ed. E. Erdmann, K. Greef, J. C. Skou), Springer-Verlag, New York, 315–312.

Unger, M. (2009) *Planta Med.*, **75**, 735–745.

Wagner, H, Bladt, S (2009) Plant Drug Analysis. A Thin Layer Chromatography Atlas. 2nd ed. Springer: Berlin; 99–102.

Wang, X., Plomley, J. B., Newman, R. A., *et al*. (2000) *Anal. Chem.*, **72**, 3547–3552.

Wolfender, J. L., Hostettmann, K., Abe, F., *et al*. (1995) *J. Chromatogr. A*, **712**, 155–168.

Wong, S. K., Tsui, S. K., and Kwan, S. Y. (2002) *J. Pharm. Biomed. Anal.*, **30**, 161–170.

Wu, X., Zhao, H., Wang, H., *et al*. (2012). *J. Sep. Sci.*, **35**, 1893–1898.

Yamaki, H., Nonaka, A., Tani, M., *et al*. (1999) *Nippon Shokuhin Kagaku Kogaku Kaishi*, **46**, 255–261.

Yang, F., Zhao, J., and Li, S. (2010) *Electrophoresis*, **31**, 260–277.

Ye, M., and Guo, D. (2005) *Rapid Commun. Mass Spectrom.*, **29**, 1881–1892.

Ye, M., Guo, H., Guo, H., *et al*. (2006) *J. Chromatogr. B*, **838**, 86–95.

Yoshimatsu, K., Sawada, J., Jaziri, M., Shimomura, K. (1995) Detection of cardenolides by ELISA in plant sciences, in *Studies in Natural Products Chemistry* Vol. **15**, (ed. Atta-ur-Rahman), Elsevier Science B.V., Amsterdam, 361–380.

Zehnder, C. B., and Hunter, M. D. (2007) *J. Chem. Ecol.*, **33**, 2044–2053.

Plant Steroids: Occurrence, Biological Significance, and Their Analysis

G. M. Kamal B. Gunaherath[1] and A. A. Leslie Gunatilaka[2]

[1]*The Open University of Sri Lanka, Nugegoda, Sri Lanka* and [2]*The University of Arizona, Tucson, AZ, USA*

1 INTRODUCTION

Steroids constitute a group of biologically important secondary metabolites produced by plants, animals, and microorganisms. They typically contain the tetracyclic 1,2-cyclopentenoperhydrophenanthrene (5α- or 5β-gonane) carbon skeleton, usually bearing methyl substituents at C-10 and C-13 and an alkyl substituent (side-chain) at C-17 (Figure 1). The immense chemical diversity exhibited by steroids arises because of various oxidation states of the carbons of its tetracyclic core and methyl groups, and the structure of the side-chain.

Biosynthetically, all steroids, irrespective of their source, are derived from *S*-squalene-2,3-epoxide, an intermediate of the acetate-mevalonate pathway. The key steps of biosynthetic pathways leading to plant, animal, and fungal steroids are depicted in Figure 2. In plants, *S*-squalene-2,3-epoxide cyclizes leading to cycloartenol (**1**) (Figure 2), which further undergoes enzymatic transformations to yield phytosterols such as β-sitosterol (**2**), campesterol (**3**) (Figure 2), and stigmasterol (**4**) (Figure 3). During these transformations, three methyl groups are lost and the cyclopropane ring of the cycloartenol opens up to form the methyl group at C-10. Biosynthesis of sterols in animals and fungi takes place via lanosterol

(**5**) (Figure 2), which is also a cyclization product of *S*-squalene-2,3-epoxide. The major sterols (steroidal alcohols) present in animals and fungi are cholesterol (**6**) and ergosterol (**7**) (Figure 2), respectively. It is noteworthy that the overall biosynthetic pathway leading to sterols has been elucidated and the respective genes/enzymes have been fully characterized (http://www.genome.jp/dbget-bin/get_pathway?org_name=map&mapno=00100).

Despite their origin from two different cyclic precursors, some plants are known to produce small amounts of animal steroids such as cholesterol (**6**) (Festucci-Buselli *et al.*, 2008). It is noteworthy that relatively high levels of cholesterol (approximately 15–20% of total sterols) have been reported from some Solanaceous plants such as potato (*Solanum tuberosum* L.) and tobacco (*Nicotiana tabacum* L.) (Arnqvist *et al.*, 2003).

1.1 Major Groups of Plant Steroids

Plant steroids can be classified into seven major groups based on taxonomic considerations and their biological functions and/or structures (Table 1) (Kreis and Müller-Uri, 2010).
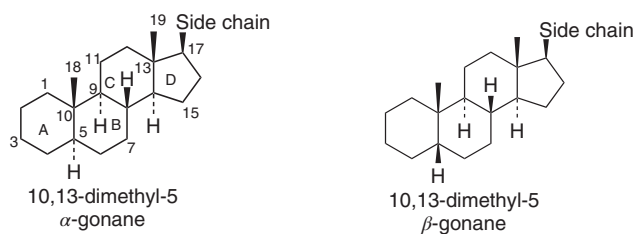
Figure 1 Basic tetracyclic core of steroids showing the numbering of C atoms and labeling of rings.
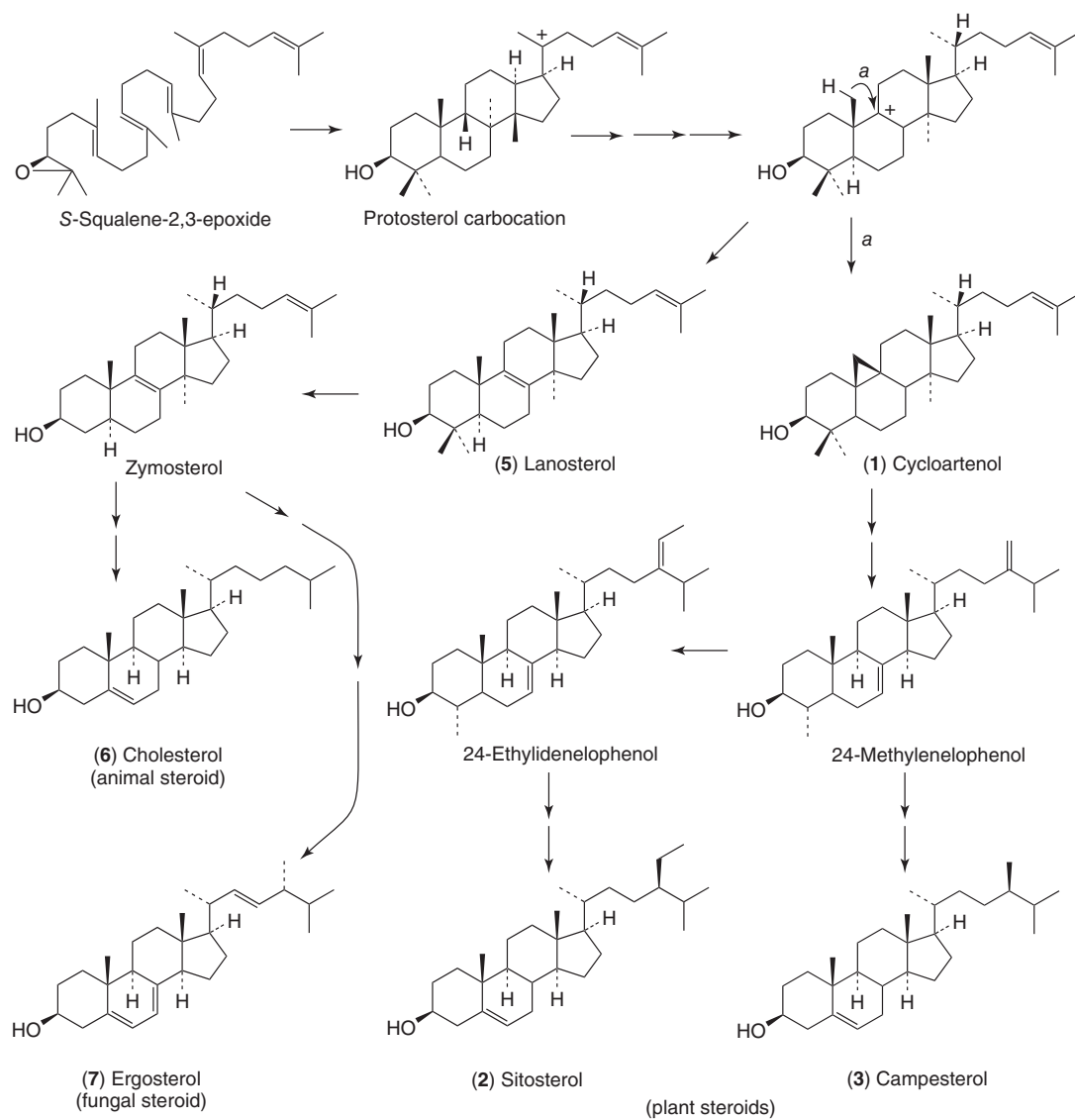
Figure 2 Key steps in the biosynthesis of plant, animal, and fungal steroids.
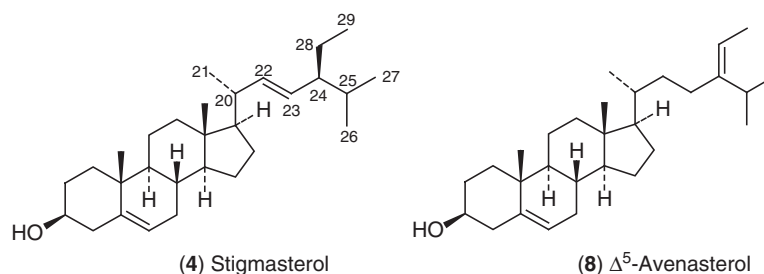
Figure 3   Some common phytosterols.

Table 1   Major groups of plant steroids.

Group	Some common examples
Phytosterols	β-Sitosterol (**2**), campesterol (**3**), stigmasterol (**4**)
Withanolides	Withaferin A (**9**), withanolide A (**10**)
Brassinosteroids	Brassinolide (**12**), castasterone (**13**)
Phytoecdysteroids	20-Hydroxyecdysone (**15**)
Steroidal alkaloids	Jervine (**21**), α-solanine (**17**), conessine (**24**)
Mammalian steroidal hormones	Androstenone (**29**), progesterone (**30**)
Steroidal saponins[a]	—
Cardiac glycosides[a]	—

[a] Not discussed in this chapter. For detailed discussion of these, see Analysis of plant saponins; Cardiotonic Glycosides.

1.2   Natural Occurrence of Plant Steroids

Steroids play important functional roles in plants. Phytosterols are integral components of the plant cell membrane lipid bilayer that controls the membrane fluidity and permeability (Kreis and Müller-Uri, 2010). Withanolides are steroidal lactones whose function in plants has not been well understood but they show significant biological activities in mammals.

Some brassinosteroids act as plant hormones and have essential function in plant development (Bajguz and Tretyn, 2003). Phytoecdysteroids are structurally similar to the insect hormones, ecdysteroids. Steroidal alkaloids are known to act as insect deterrents and chemical barriers against pathogens (Valkonen *et al.*, 1996).

1.2.1   Occurrence of Phytosterols

Phytosterols are steroidal alcohols that are ubiquitous in the plant kingdom. Recently, phytosterols

have attracted considerable attention because of the claims of their hypocholesterolemic activity (Moghadasian, 2000). Common phytosterols found in plants are β-sitosterol (**2**), campesterol (**3**), (Figure 2), stigmasterol (**4**), and Δ^5-avenasterol (**8**) (Figure 3). They differ from cholesterol (**6**) by having an alkyl substituent at C-24.

1.2.2   Occurrence of Withanolides

Withanolides are a group of steroidal lactones built on an unmodified or a modified ergostane skeleton in which C-22 and C-26 are appropriately oxidized to form a δ-lactone or a potential lactone ring. They mainly occur in plants of the family Solanaceae belonging to the genera *Acnistus, Datura, Deprea, Dunalis, Discopodium, Exodeconus, Hyoscyamus, Iochroma, Jaborosa, Larnax, Lycium, Nicandra, Physalis, Salpichroa, Trechonaetes, Tubocapsicum, Vassobia*, and *Withania* (Chen, He, and Qiu, 2011). Withanolides are not only restricted to the family Solanaceae but they also occur in plants of the families Fabaceae (*Cassia siamea* Lam.), Lamiaceae (*Ajuga parviflora* Benth. and *Ajuga bracteosa* Benth.), Myrtaceae (*Eucalyptus globulus* Labill.), and Taccaceae [*Tacca plantaginea* (Hance) Drenth and *Tacca chantrieri* André] (Chen, He, and Qiu, 2011).

Following the simultaneous discovery of antitumor withanolide, withaferin A (**9**), from *Withania somnifera* (L.) Dunal (Lavie, Glotter, and Shvo, 1965) and from *Acnistus arborescens* Schltdl. (Kupchan *et al.*, 1965), investigations have led to the isolation and identification of over 200 withanolides in which the carbocyclic skeleton, the side chain, or both are modified. Withanolides have been the subject of several recent reviews (Mirjalili *et al.*, 2009; Chen, He, and Qiu, 2011; Misico *et al.*, 2011). Other important
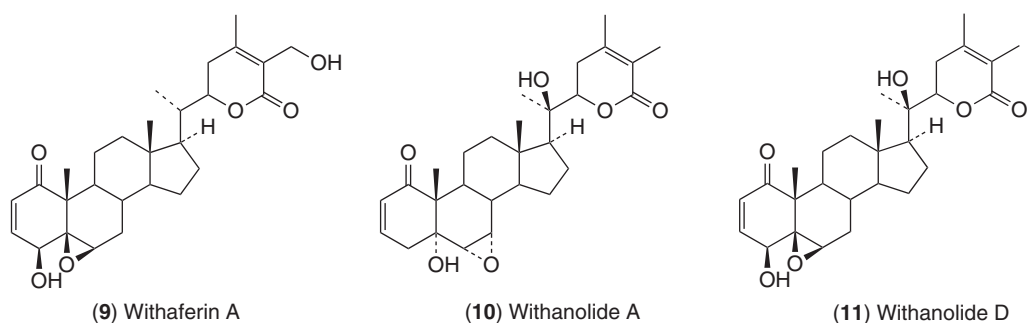
Figure 4   Some commonly encountered withanolides.

biologically active withanolides include withanolide A (**10**) and withanolide D (**11**) (Figure 4).

1.2.3   Occurrence of Brassinosteroids

Brassinosteroids are a small group of steroidal plant growth hormones consisting of over 70 members (Bajguz and Tretyn, 2003; Kreis and Müller-Uri, 2010). They can be classified as C_{27}-, C_{28}-, and C_{29}-brassinosteroids based on the number of carbon atoms in the side chain at C-17. Their structural diversity originates as a result of the type and positions of the substituents on A/B rings and in the side chain. Brassinolide (**12**) (Figure 5), the first brassinosteroid to be encountered, has been obtained from the pollen of rapeseed (*Brassica napus* L.) and is a steroidal lactone (Grove *et al.*, 1979) in which ring B has undergone expansion resulting in a lactone. Brassinolide (**12**), castasterone (**13**), and 24-*epi*brassinolide (**14**) are three representative examples of brassinosteroids (Figure 5). Brassinolide (**12**) and castasterone (**13**) have been

frequently encountered in plants and are considered to be the most important brassinosteroids in the plant kingdom (Kim *et al.*, 2005).

1.2.4   Occurrence of Phytoecdysteroids

Ecdysteroids are steroidal hormones of all classes of arthropods and probably of other invertebrates (Dinan, 2001). They are important in all developmental stages of these insects. The early reports of occurrence of ecdysteroids in plants include 20-hydroxyecdysone (**15**) (Figure 6) in the fern, *Polypodium vulgare* L. (Polypodiaceae), and few others from *Achyranthes fauriei* H. Lév. & Vaniot (Amaranthaceae) and *Podocarpus nakaii* Hayta (Podocarpaceae) (Dinan, 2001). Since then, over 400 ecdysteroids have been reported from plant kingdom (Lafont *et al.*, 2002). Phytoecdysteroids contain C_{27}, C_{28}, or C_{29} steroidal skeleton possessing 7-en-6-one motif with *cis* fusion of A/B rings (5β-H). Most phytoecdysteroids possess hydroxyl group at C-14 (α-oriented), while they are usually polyhydroxylated
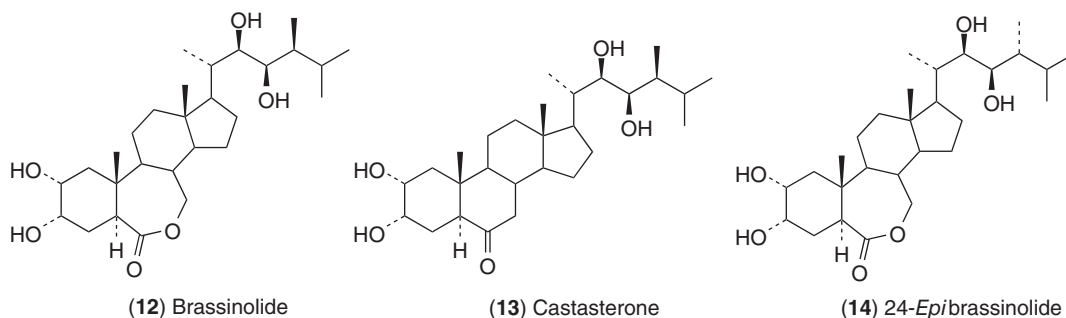
Figure 5   Some commonly encountered brassinosteroids.

(**15**) 20-Hydroxyecdysone   (**16**) Polypodine B
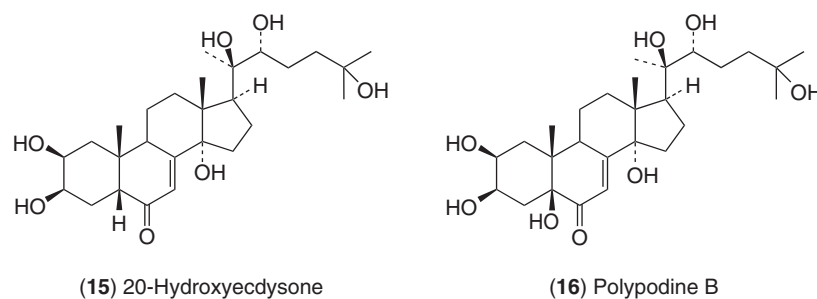
Figure 6 Some frequently encountered phytoecdysteroids.

at C-2, C-3, C-20, and C-22. The most predominant phytoecdysteroids are 20-hydroxyecdysone (**15**) and polypodine B (**16**) (Dinan, Harmatha, and Lafont, 2011) (Figure 6).

1.2.5 Occurrence of Steroidal Alkaloids

Steroidal alkaloids are a highly diverse class of plant steroids. These are usually found in plants of the families, Apocynaceae, Asclepiadaceae, Buxaceae, Liliaceae, and Solanaceae. Steroidal alkaloids consist of modified or unmodified steroid skeleton with nitrogen atom(s) integrated to a ring or attached

to the side chain or as a substituent of the tetracyclic core. Although there are several subgroups of steroidal alkaloids based on the position of nitrogen atom(s) and the steroidal skeleton and its modifications (Li, Jiang, and Li, 2006; Dinan, Harmatha, and Lafont, 2011), they can be broadly classified into two groups, based on the steroidal skeleton. These are steroidal alkaloids having modified or unmodified C_{27} steroidal skeleton and those containing a pregnane carbon skeleton.

α-Solanine (**17**), α-tomatine (**18**), dehydrotomatine (**19**), and verazine (**20**) (Figure 7) are some selected examples of steroidal alkaloids having an unmodified C_{27} steroidal skeleton. Many plants of the family Solanaceae are rich in steroidal alkaloids

(**17**) α-Solanine   (**18**) α-Tomatine

(**19**) Dehydrotomatine   (**20**) Verazine
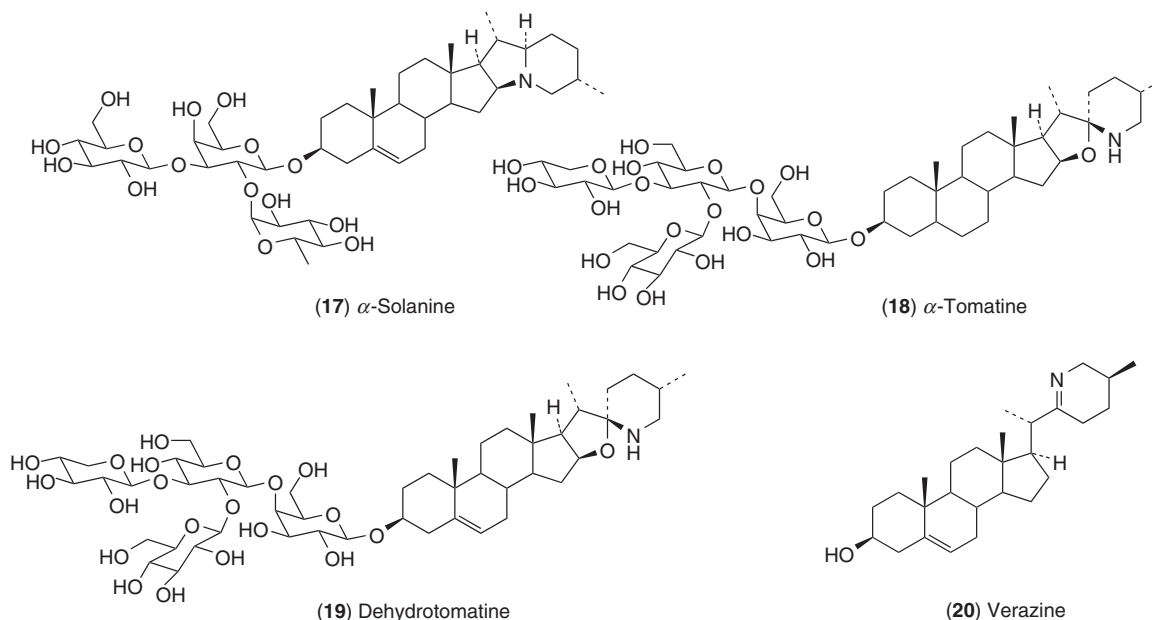
Figure 7 Some selected steroidal alkaloids having an unmodified C_{27} steroidal skeleton.
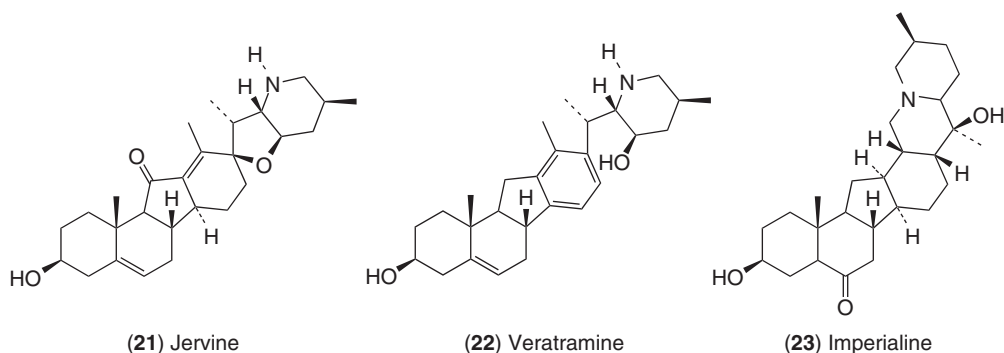
Figure 8   Some selected steroidal alkaloids with modified C_{27} steroid skeleton.

that occur as their glycosides (glycoalkaloids) (Arnqvist *et al*., 2003). These steroidal glycoalkaloids are reported to be derived from cholesterol with the introduction of a nitrogen atom (from the amino acid L-arginine) into the side chain and glycosylation usually at 3-OH (Kreis and Müller-Uri, 2010). The chemical diversity of these glycoalkaloids is largely due to the type of skeleton (solanidine or spirosolane), diverse nature of the glycoside moiety, substitution and/or stereochemistry of the heterocyclic ring, and the presence or absence of the Δ^5 double bond (Ginzberg, Tokuhisa, and Veilleux, 2009). A major solanidine-type glycoalkaloid present in potato (*S. tuberosum* L.) (Solanaceae) is α-solanine (**17**), whereas α-tomatine (**18**) and dehydrotomatine (**19**) are the major spirosolane glycoalkaloids in tomato (*Solanum lycopersicum* L.) (Solanaceae) (Figure 7).

Jervine (**21**), veratramine (**22**), and imperialine (**23**) (Figure 8) are representative examples of steroidal alkaloids having a modified C_{27} steroidal skeleton. They occur mainly in plants of the family Liliaceae. The C_{27} steroid skeleton of these steroidal alkaloids has undergone a rearrangement to form a five-membered C ring and a six-membered D ring (C-*nor*-D-*homo* steroid ring system) (Li, Jiang, and Li, 2006).

The structural diversity of steroidal alkaloids containing the pregnane skeleton is exemplified by conessine (**24**) (Yang *et al*., 2012), cyclomicrobuxamine (**25**), irehine (**26**) (Ata and Andersh, 2008), paravallarine (**27**) (Phi *et al*., 2011), and terminamine A (**28**) (Zhai *et al*., 2012) (Figure 9). Plants belonging to the families Asclepiadaceae and Buxaceae are rich in these steroidal alkaloids (Deng,

Liao, and Chen, 2005; Devkota *et al*., 2010; Phi *et al*., 2011).

1.2.6 Occurrence of Plant-Derived Mammalian Steroidal Hormones

Mammalian steroidal hormones such as estrogens, androgens, and progestogens are classified based on their functions and the number of carbons present in their molecules. The presence of mammalian steroidal hormones in plants was first detected in 1926 but because of the methods used for their detection were not reliable at that time, the results were considered to be only preliminary (Janeczko and Skoczowski, 2005). Since then, there were many investigations and with the use of improved analytical methods, the qualitative and quantitative estimations of mammalian steroidal hormones in plants have proven to be more reliable. A radioimmunoassay (RIA)-based study of 128 plant species revealed that more than 80% of the species investigated contained androstenone (**29**) and progesterone (**30**) (Figure 10), whereas 70% showed the presence of androgens [testosterone (**31**) and/or dihydrotestosterone (**32**)] and 50% showed the presence of estrogens [estrone (**33**) and/or 17β-estradiol (**34**)] (Figure 10) (Simons and Grinwich, 1989). The occurrence and activity of progesterone (**30**) in plants have recently been reviewed (Janeczko, 2012). Although there were many reports of detection of mammalian steroidal hormones using various advanced techniques (Janeczko and Skoczowski, 2005; Simerský *et al*., 2009), it is noteworthy that the isolation and characterization of progesterone (**30**) from a plant [*Adonis aleppica* Boiss. (Ranunculaceae)] has
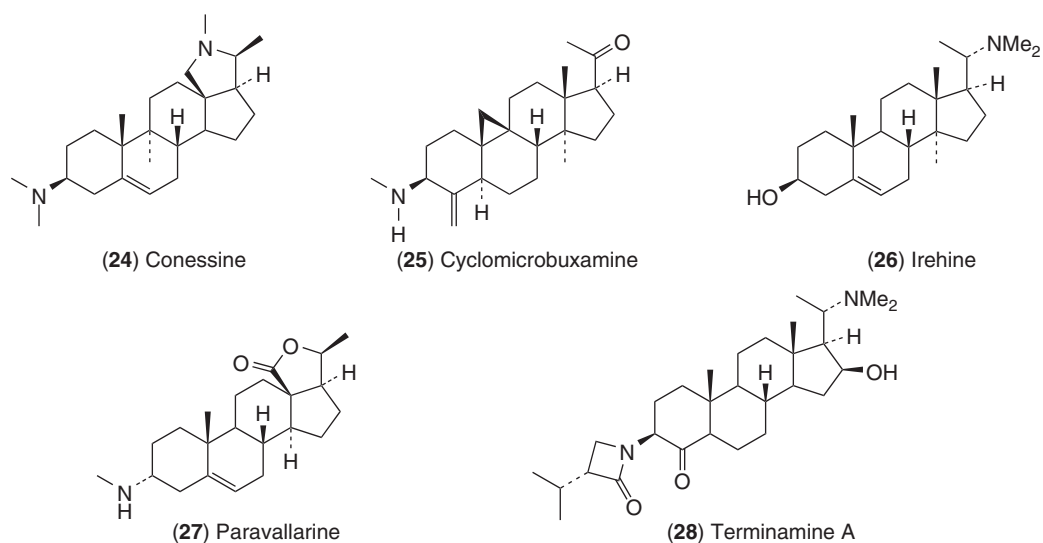
Figure 9 Some selected steroidal alkaloids having pregnane skeleton.

(**24**) Conessine

(**25**) Cyclomicrobuxamine

(**26**) Irehine

(**27**) Paravallarine

(**28**) Terminamine A
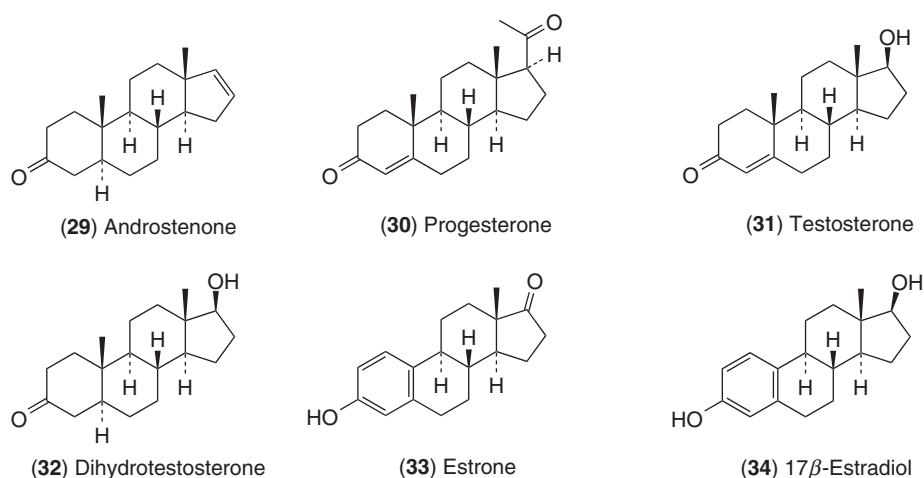
Figure 10 Some mammalian steroidal hormones encountered in plants.

(**29**) Androstenone

(**30**) Progesterone

(**31**) Testosterone

(**32**) Dihydrotestosterone

(**33**) Estrone

(**34**) 17β-Estradiol

been reported recently for the first time (Pauli *et al.*, 2010).

2 BIOLOGICAL SIGNIFICANCE OF PLANT STEROIDS

Plants and plant parts are major sources of food and constituents of traditional medicines. The active ingredients of these plant-derived traditional medicines, practiced in many countries, are the secondary metabolites present in them. As steroids are ubiquitous in plants, they play a major role in pharmacological activities of these medicines. With the advancement of high throughput screening methods and introduction of new bioassays, these plants have been subjected to various chemical and biological investigations. These investigations have led to the identification of steroids with a variety of biological activities.

2.1   Phytosterols

All vegetable foods are known to contain appreciable quantities of phytosterols. The most abundant sources of phytosterols are vegetable oils. Except for highly refined carbohydrates and animal products, nearly all foods contribute considerably to phytosterol intake by humans (Ostlund, 2002). The major sterols present in edible fats of plant origin are β-sitosterol (**2**) (∼65% of sterol content) and campesterol (**3**) (∼30% of sterol content) (Figure 2). Phytosterols are reported to interfere with the intestinal absorption of cholesterol and lower the blood cholesterol levels. The mechanism of action has been attributed to the bulkiness of phytosterols compared to cholesterol (Segura *et al*., 2006). Recent work comparing phytosterol-free corn oil and commercial corn oil in single test meal on human subjects reported that cholesterol absorption increased when corn oil was free of phytosterols (Ostlund *et al*., 2002). A reduction in cholesterol absorption has been observed when phytosterols were added back to the phytosterol-free corn oil, confirming that this effect was due to the presence of phytosterols in natural corn oil. Many studies have been carried out on this subject, and it was found that on average, 13% reduction of blood low density lipoprotein-cholesterol (LDL-cholesterol) and 10% reduction of total blood cholesterol occurred without affecting the high density lipoprotein-cholesterol (HDL-cholesterol) and total lipid content in humans (Moghadasian and Frohlich, 1999). Sitostanol, the saturated analog of β-sitosterol (**2**) (Figure 2), which is present in minute amounts in plants, is found to be more potent in inhibition of intestinal cholesterol absorption (Jones *et al*., 1997). Stanols (the saturated sterols) can be obtained by catalytic hydrogenation of sterols. Subsequent investigation by Jones *et al*. (1999) has revealed that a mixture of β-sitosterol (**2**) and sitostanol was as effective as sitostanol alone in inhibiting cholesterol absorption.

Apart from the hypocholesterolemic activity, phytosterols show *in vitro* anticancer properties and are claimed to inhibit the development of various cancers in humans (Woyengo, Ramprasath, and Jones, 2009). It has been reported that β-sitosterol is effective in the symptomatic treatment of benign prostatic hyperplasia (BPH) (enlargement of the prostate), although the prostatic size remained unchanged during the treatment period (Berges, Kassen, and Senge, 2000). Further, it has been shown that β-sitosterol exhibits antihelminthic and antimutagenic activities, whereas both β-sitosterol and β-sitosteryl-β-D-glucoside encountered in the leaves of *Mentha cordifolia* Opiz (Lamiaceae) have exhibited analgesic activities (Villaseñor *et al*., 2002).

2.2   Withanolides

Biological activities of withanolides have been extensively reviewed (Chen, He, and Qiu, 2011). Plants containing withanolides as major constituents such as *W. somnifera* (L.) Dunal, *Physalis angulata* L., and other *Physalis* species are widely used in traditional medicines in various countries. *W. somnifera* (L.) Dunal, commonly known as *Ashwagandha* or *Indian ginseng*, is an important medicinal plant employed in Ayurvedic medicine for over 3000 years. Roots of *W. somnifera* (L.) Dunal are a constituent of over 200 formulations in Ayruveda, Siddha, and Unani medicines (Mirjalili *et al*., 2009). Hence, these plants have been subjected to extensive biological and chemical investigations and constituent withanolides have been found to show antitumor, immunomodulatory, anti-inflammatory, anti-leishmanial, and antimicrobial activities (Chen, He, and Qiu, 2011). Significant biological activities of a few selected withanolides (Figures 4 and 11) are summarized in Table 2.

2.3   Brassinosteroids

Brassinosteroids are plant growth regulators occurring in very low concentrations in plants and are required for the normal plant growth and development. In addition, brassinosteroids have the ability to protect plants from various environmental stresses, including drought, extreme temperatures, heavy metals, herbicidal injury, and salinity (Clouse and Sasse, 1998; Sasse, 1997). It has been shown that 24-*epi*brassinolide (**14**) (Figure 5) increases thermotolerance of tomato plants by inducing the expression of mitochondrial small heat-shock proteins (Singh and Shono, 2005). It is noteworthy that 24-*epi*brassinolide (**14**) has shown antigenotoxicity effect in a chromosomal aberration assay using *Allium cepa* L. (Sondhi *et al*., 2008) and antioxidant and neuroprotective properties in a mammalian neuronal cell culture model (Carange *et al*., 2011).

(35) Withangulatin A   **(36)** Physagulin H   **(37)** 4-Deoxywithaperuvin

(38) Bracteosin A   **(39)** Viscosalactone B   **(40)** Witharistatin
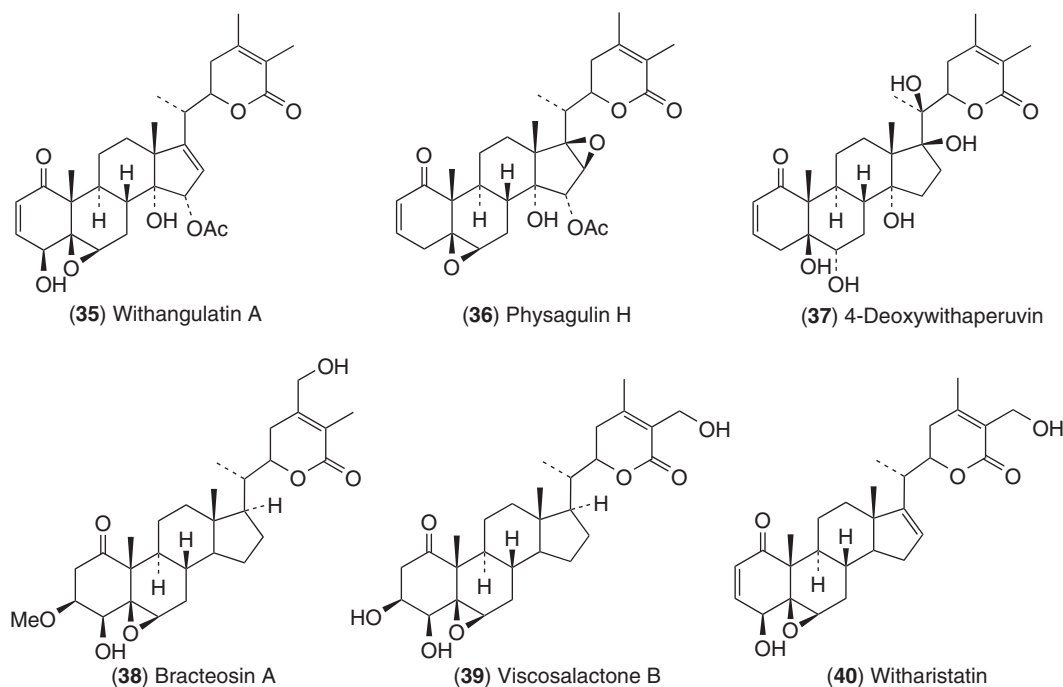
Figure 11 Some selected biologically active withanolides.

Table 2 Biological activities of some selected withanolides.

Withanolide	Biological activity	References
Withaferin A (**9**)	Anticancer, cytotoxic	Misra *et al*. (2008), Samadi *et al*. (2010)
	Immunosuppresive	Furmanowa *et al*. (2001)
	Cholinesterase inhibitory	Chen, He, and Qiu (2011)
	Anti-inflammatory	Jayaprakasam and Nair (2003)
Withanolide A (**10**)	Cholinesterase inhibitory	Chen, He, and Qiu (2011)
Withanolide D (**11**)	Antitumor, antimetastatic	Leyon and Kuttan (2004)
Withangulatin A (**35**)	Immunosuppresive	Sun *et al*. (2011)
Physagulin H (**36**)	Trypanocidal	Chen, He, and Qiu (2011)
4-Deoxywithaperuvin (**37**)	Antimicrobial	Chen, He, and Qiu (2011)
Bracteosin A (**38**)	Cholinesterase inhibitory	Chen, He, and Qiu (2011)
Viscosalactone B (**39**)	Anti-inflammatory	Jayaprakasam and Nair (2003)
Witharistatin (**40**)	Diuretic	Benjumea *et al*. (2009)

2.4 Phytoecdysteroids

Plants containing phytoecdysteroids are used in folk medicines in many countries, while some herbal preparations containing these plants have been approved by certain countries as tonics in phytomedicine. Some of the preparations containing phytoecdysteroids are also used for the purpose of body building by athletes (Báthori and Pongrácz, 2005). Among the natural phytoecdysteroids, 20-hydroxyecdysone (**15**) (Figure 6) has shown anabolic, hypoglycemic, antiarrhythmic, immunostimulant, and hepatoprotective activities on animal test models (Báthori and Pongrácz, 2005). It is noteworthy that the anabolic effects exhibited by phytoecdysteroids are not associated with androgenic, antigonadotropic, and thymolytic (causing destruction of thymic tissue) side effects.

(41) Cyasterone

	R$_1$	R$_2$	R$_3$	R$_4$	R$_5$	
(42)	H	OH	H	H	H	2-Deoxyecdysone
(43)	H	OH	H	OH	H	2-Deoxy-20-hydroxyecdysone
(44)	OH	OH	H	H	H	Ecdysone
(45)	OH	H	OH	OH	H	Pterosterone
(46)	OH	OH	H	OH	OH	Turkesterone
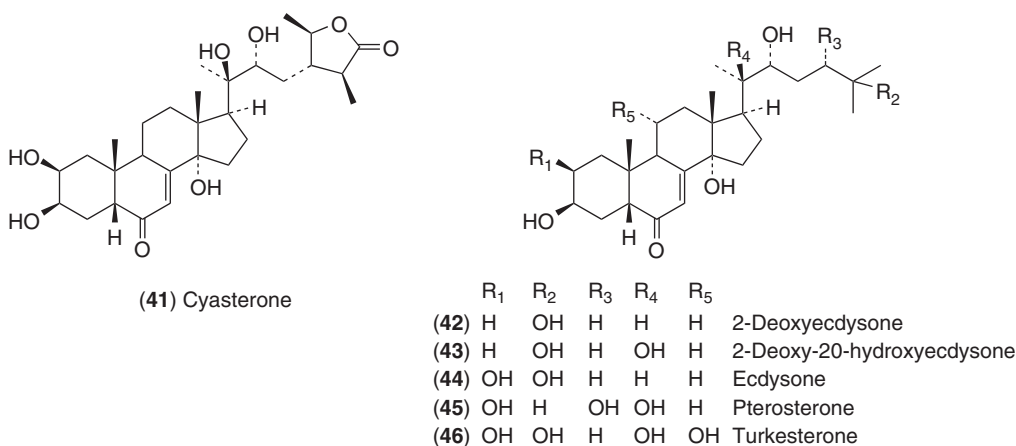
Figure 12   Some selected biologically active phytoecdysteroids.

Table 3   Pharmacological effects shown by some phytoecdysteroids (Báthori and Pongrácz, 2005).

Phytoecdysteroid	Pharmacological activity/activities
20-Hydroxyecdysone (**15**)	Anabolic, antiarrhythmic, antioxidant, hypoglycemic hepatoprotective, hypoazatemic,[a] and immunostimulant
Polypodine B (**16**)	Hypoazatemic
Cyasterone (**41**)	Anabolic, hepatoprotective, and hypoazatemic
2-Deoxyecdysone (**42**)	Immunostimulant
2-Deoxy-20-hydroxyecdysone (**43**)	Immunostimulant
Ecdysone (**44**)	Hypocholesterolemic
Pterosterone (**45**)	Anabolic
Turkesterone (**46**)	Anabolic, hypoazatemic

[a] Lowering of urea and other nitrogenous constituents in the blood.

Pharmacological effects of some selected phytoecdysteroids (Figures 6 and 12) are summarized in Table 3.

2.5   Steroidal Alkaloids

Steroidal glycoalkaloids of solanidine and spirosolane types are known to possess antitumor, antifungal, antiviral, and embryotoxic activities (Torres *et al.*, 2011; Wang *et al.*, 2011). Jervine (**21**) (Figure 8) and 11-deoxyjervine (cyclopamine, **47**) (Figure 13) are known to be natural teratogens that induce holoprosencephaly (a cephalic disorder in which the forebrain of the embryo fails to develop into two hemispheres) in animals (Li, Jiang, and Li, 2006). Structure–activity relationship (SAR) studies of these C$_{27}$ steroidal alkaloids have revealed

that the presence of Δ^5-double bond is a critical structural feature for their teratogenicity (Li, Jiang, and Li, 2006). Some selected C$_{27}$ steroidal alkaloids (Figure 13) with significant biological activities are summarized in Table 4.

Steroidal alkaloids bearing pregnane skeleton show antibacterial, antifungal, antiviral, antimalarial, glutathione *S*-transferase inhibitory, and cholinesterase inhibitory activities. Glutathione *S*-transferases are detoxification isozymes that can form adducts with electrophilic substances. Anticancer drugs with electrophilic centers readily form adducts with glutathione *S*-transferase, which are easily excreted from the body, thus lowering the efficiency of the drug. Hence, glutathione *S*-transferase inhibitors are important in cancer chemotherapy (Ata and Andersh, 2008). Cholinesterases including acetyl- and butyl-cholinesterases are found to be

(47) Cyclopamine

(48) Lycioside A

(49) 22α,23α-Epoxysolanida-1,4,9-trien-3-one

(50) Ebeinone

(51) Ebeiensine

(52) Jervinone

(53) Stenophylline B
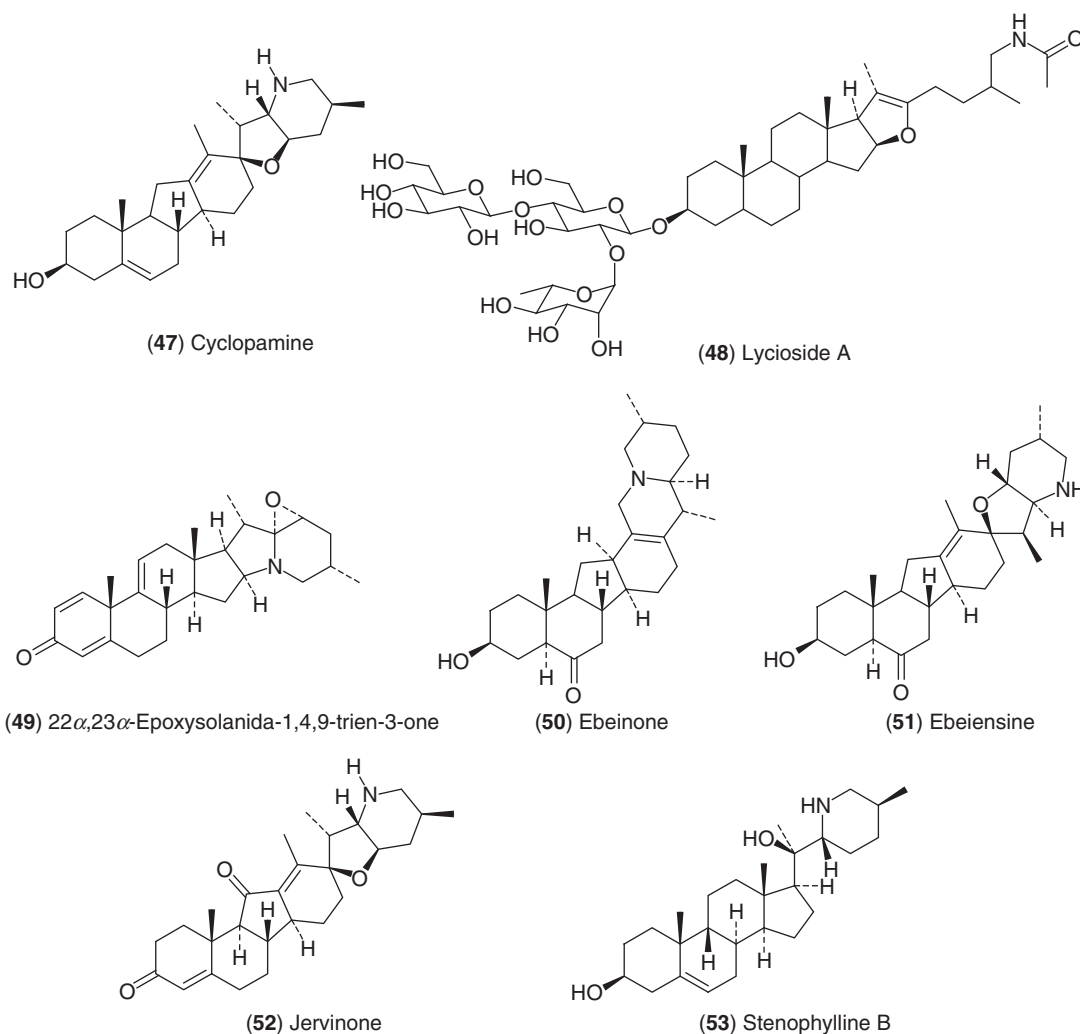
Figure 13　Some selected biologically active C$_{27}$ steroidal alkaloids.

responsible for the hydrolysis of acetylcholine. Deficiency of acetylcholine in the brain has been identified as a cause for Alzheimer's disease. Hence, one of the treatment strategies for Alzheimer's disease is to use potent cholinesterase inhibitors (Ata and Andersh, 2008).

Plants belonging to families, Asclepiadaceae, Apocynaceae, and Buxaceae, that contain pregnane alkaloids are used in traditional medicines in many parts of the world (Devkota *et al*., 2010; Phi *et al*., 2011; Zhai *et al*., 2012). For example, *Holarrhena antidysenterica* Wall. ex A.DC.

(Apocynaceae) is a plant used in traditional medicine in the eastern Asian countries as a remedy for diseases of liver, amoebic dysentery, diarrhea, and other intestinal ailments as well as asthma, bronchopneumonia, and malaria (Kumar *et al*., 2007; Yang *et al*., 2012). A recent study of a methanol extract of *H. antidysenterica* Wall. ex A.DC. as well as its alkaloid constituents has shown strong acetylcholinesterase inhibitory activity (Yang *et al*., 2012). Significant biological activities of some selected steroidal alkaloids having pregnane skeleton (Figures 9 and 14) are given in Table 5.

Table 4   Pharmacological effects shown by some selected C_{27} steroidal alkaloids.

Steroidal alkaloid	Biological activity	References
Lycioside A (**48**)	α-Glucosidase inhibitor	Wang *et al*. (2011)
22α,23α-Epoxysolanida-1,4,9-trien-3-one (**49**)	Antiophidic[a]	Torres *et al*. (2011)
Ebeinone (**50**)	Cholinesterase inhibitor	Li, Jiang, and Li (2006)
Ebeiensine (**51**)	Anti-asthmatic	Li, Jiang, and Li (2006)
Jervinone (**52**)	Antihypertensive	Li, Jiang, and Li (2006)
Stenophylline B (**53**)	Antifungal	Li, Jiang, and Li (2006)

[a] Ability to inhibit or relieve the effects of snakebite.

(54) Saligenamide E   **(55)** Sarcovagine D   **(56)** Hookerianamide J
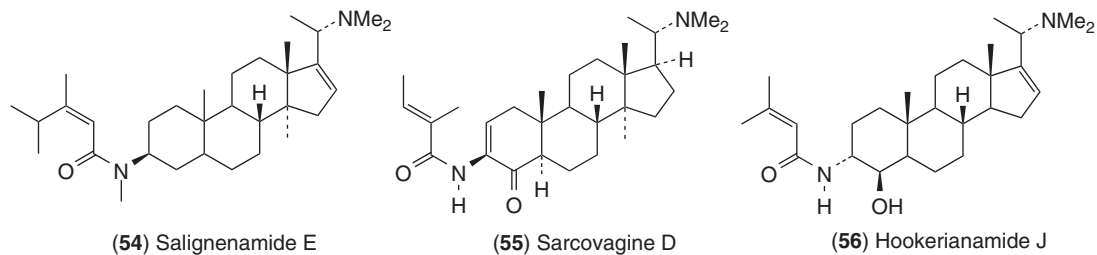
Figure 14   Some biologically active steroidal alkaloids with pregnane skeleton.

Table 5   Significant biological activities shown by some selected pregnane type steroidal alkaloids.

Steroidal alkaloid	Biological activity	References
Conessine (**24**)	Cholinesterase inhibitor	Yang *et al*. (2012)
Irehine (**26**)	Glutathione *S*-transferase inhibitor	Ata and Andersh (2008)
Paravallarine (**27**)	Cytotoxic	Phi *et al*. (2011)
Terminamine A (**28**)	Antimetastasis	Zhai *et al*. (2012)
Saligenamide E (**54**)	Cholinesterase inhibitor	Atta-ur-Rahman *et al*. (2002)
Sarcovagine D (**55**)	Cytotoxic	Yan *et al*. (2011)
Hookerianamide J (**56**)	Cholinesterase inhibitor	Devkota *et al*. 2008

3   ANALYSIS OF PLANT STEROIDS

World Health Organization (WHO) has estimated that about 78% of the world's population depends on traditional medicines employing medicinal plants for their primary health care. Chemical and pharmacological analyses of many of these plants have revealed that they contain biologically active constituents. Hence, knowledge on plants used in the traditional medicines has led to many herbal supplements in the global market. With the increasing global demand for these herbal supplements, they are manufactured in industrial scale. Hence, it is important that these preparations and their constituent plants should undergo standardization procedures to maintain their safety and efficacy.

Numerous plants used as foods and in traditional medicines are known to contain steroids, many of which may exert physiological activities. Therefore, it is necessary to have simple and routinely employable analytical methods that can detect and quantify these steroids in plants and herbal preparations. For this purpose, methods have been developed using chromatographic techniques such as thin-layer chromatography (TLC), high performance liquid chromatography (HPLC), ultra high performance liquid chromatography (UHPLC), and gas liquid chromatography (GLC). In addition to these techniques, mass spectrometry (MS) and tandem mass spectrometry (MS/MS) are also used as hyphenated techniques with HPLC, UHPLC, and GLC. Application of these techniques for detection and quantification of plant steroids is discussed in the following sections.

3.1   Thin-Layer Chromatography (TLC)

TLC continues to be an important analytical method in plant steroid analysis. It has many advantages though there are some disadvantages that could be

overcome with method development. It is fast and hence many samples can be analyzed simultaneously and rapidly. Other advantages are that it does not need sophisticated instruments and many different methods of detection could be applied. Hence, TLC is the method of choice in many instances where rapid and reasonably accurate routine analyses of samples are required. In addition, TLC is an ideal classic tool for the initial phytochemical analysis and for monitoring of column chromatographic fractions during purification of plant steroids. Application of TLC in the analysis of steroids has been reviewed recently (Bhawani *et al.*, 2010). Plant steroids including phytosterols, withanolides, phytoecdysteroids, and steroidal glycoalkaloids have been successfully analyzed qualitatively and quantitatively by TLC. Quantitative measurements on TLC are carried out usually by TLC densitometry. Densitometry is the quantitative measurement of the optical density of a light-sensitive substance, which can be used to quantitate a spot on a thin-layer chromatogram. However, TLC has limited applications in the analyses of brassinosteroids and mammalian steroidal hormones in plants, which are present in trace amounts. (See **Thin-layer Chromatography, with Chemical and Biological Detection Methods** for a detailed description of TLC.)

3.1.1 TLC Analysis of Phytosterols

Clerodendrum phlomidis Linn. (Verbenaceae) is a constituent plant in the well-known ayurvedic formulation, Dashamoola, used in the treatment of fatigue of women after child birth. Determination of the marker compound, 24*S*-ethylcholesta-5,22*E*,25-trien-3β-ol (**57**) (Figure 15), present in the aerial parts of *C. phlomidis* Linn. was achieved by employing high performance thin-layer chromatography (HPTLC) on silica gel 60 F_{254}, using $CHCl_3$: MeOH (98.5 : 1.5) as the eluent (Shanker *et al.*, 2008). The quantification of **57** was carried out using TLC densitometry at 650 nm after treatment with the anisaldehyde reagent. A precise and accurate quantification of this compound has been achieved at concentrations of 150–400 ng per band. The method has been validated for peak purities, precision, robustness, limit of detection (LOD), and limit of quantification (LOQ) as per ICH (International Conference on Harmonization) guidelines.

(**57**) 24*S*-Ethylcholesta-5,22*E*,25-triene-3-β-ol
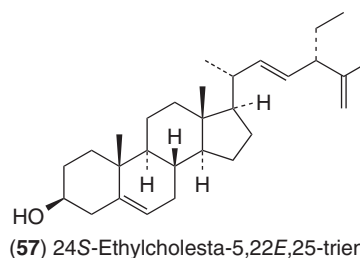
Figure 15 An example of a phytosterol analyzed by TLC.

The method was reported to be useful in the determination of this sterol in *C. phlomidis* Linn. extract and herbal preparations containing this plant.

3.1.2 TLC Analysis of Withanolides

Separation and simultaneous quantification of four withanolides including withaferin A (**9**) and withanolide A (**10**) (Figure 4) in a root extract of *W. somnifera* L. Dunal by HPTLC densitometry has recently been described (Devkar *et al.*, 2012). Separation of these withanolides and their tentative identification utilizing R_f (retardation factor) values has been achieved on silica gel G60 F_{254} HPTLC eluting with CH_2Cl_2 : toluene : MeOH : acetone : Et_2O (7.5 : 7.5 : 3 : 1 : 1 v/v). The quantification has been achieved by TLC densitometry employing absorption reflection mode at 235 nm. Quantification of withanolide A (**10**) (Figure 4) in an ayurvedic medicine containing *W. somnifera* L. Dunal by HPTLC has also been described (Shinde *et al.*, 2011). In this determination, withanolide A (**10**) has been separated on silica gel G60 F_{254} HPTLC eluting with EtOAc : MeOH : toluene : H_2O (4 : 1 : 1 : 0.5 v/v) and was detected at 320 nm by densitometry.

3.1.3 TLC Analysis of Phytoecdysteroids

A comprehensive account on the application of TLC techniques for the determination of phytoecdysteroids has appeared (Báthori *et al.*, 2003). Several mobile phases with normal phase (NP) silica gel TLC as well as reversed phase (RP) TLC with aminopropyl and cyano-propyl bonded phases, as well as paraffin-impregnated TLC-silica plates, have been used. RP-TLC is usually
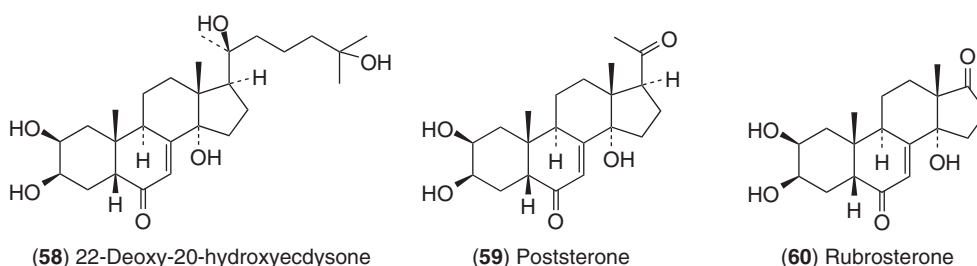
Figure 16   Some examples of phytoecdysteroids analyzed by TLC.

employed to complement NP-TLC analysis. It is reported that some poorly resolved pairs of phytoecdysteroids such as 20-hydroxyecdysone (**15**) (Figure 6)/22-deoxy-20-hydroxyecdysone (**58**) and poststerone (**59**)/rubrosterone (**60**) (Figure 16) could be conveniently separated by RP-TLC, although resolution between 5β-H and 5β-OH analogs was found to be poor.

Visualization of phytoecdysteroids on TLC could easily be done by UV (ultra violet) at 254 nm as they contain αβ-unsaturated carbonyl motif. Usually, the detection of phytoecdysteroids in a plant extract is achieved by a triple detection method that is specific to ecdysteroids (Báthori *et al.*, 2003). The method consists of (i) detection at 254 nm for spots that quench the UV radiation, (ii) spraying with vanillin-sulfuric acid reagent and observing the spots under day light (spots are colored from turquoise green through yellow, orange, and brown to violet), and (iii) viewing the induced fluorescence (FL) under UV at 354 nm after spraying with vanillin-sulfuric acid reagent.

Dietary supplements comprising plants containing 20-hydroxyecdysone (**15**) (Figure 6) are marketed as it is claimed that this steroid exhibits anabolic effects without adverse consequences. *Sida rhombifolia* L. (Malvaceae), a constituent of some herbal supplements, is one of the plant sources that contains 20-hydroxyecdysone (**15**). A TLC method to quantify this phytoecdysteroid in *S. rhombifolia* L. has been developed (Jadhav *et al.*, 2007) using silica gel G60 F$_{254}$ HPTLC eluting with MeOH : CHCl$_3$ (2 : 8). Detection and quantification of 20-hydroxyecdysone (**15**) has been carried out with TLC densitometry at 250 nm. Several *Sida* species and samples of dietary supplements containing *Sida* species have been analyzed for 20-hydroxyecdysone (**15**) using this method.

3.1.4   *TLC Analysis of Steroidal Alkaloids*

Steroidal alkaloids including glycoalkaloids are polar compounds and their detection in biological samples is a challenging problem as the majority of them do not contain a chromophore. Quantification of three steroidal glycoalkaloids by HPTLC densitometry has recently been reported (Shanker *et al.*, 2011). In this method, a semi-purified fraction containing steroidal glycoalkaloids from the medicinal plant, *Solanum xanthocarpum* Schrad. & J. C. Wendl. (Solanaceae) were separated by HPTLC (silica gel G60 F$_{254}$) eluting with CHCl$_3$: MeOH : H$_2$O (70 : 30 : 10). Detection and quantification has been achieved by spraying with Dragendorff's reagent followed by TLC densitometry scanning at 520 nm. Simultaneous quantification of solasonine (**61**), solamargine (**62**), and khasianine (β-2-solamargine) (**63**) (Figure 17) has been achieved by application of this method.

TLC hyphenated with immunostaining offers a powerful method of analyzing plant steroids. Thus, the presence of steroidal glycoalkaloids has been determined using a TLC immunostaining technique (Tanaka *et al.*, 1997). In this method, the steroidal glycoalkaloids solasonine (**61**) and solamargine (**62**) were separated on TLC using silica gel G60 eluting with CHCl$_3$: MeOH : NH$_4$OH (7 : 2.5 : 1) and the separated compounds were transferred to a polyvinylidene difluoride (PVDF) membrane by pressing the TLC plate against PVDF membrane at 130°C for 45 s. The membrane was removed, treated with NaIO$_4$ solution followed by bovine serum albumin (BSA), and immunostained with anti-solamargine monoclonal antibody (MAb). This method was found to be superior to detection with Dragendorff's or sulfuric acid reagents.
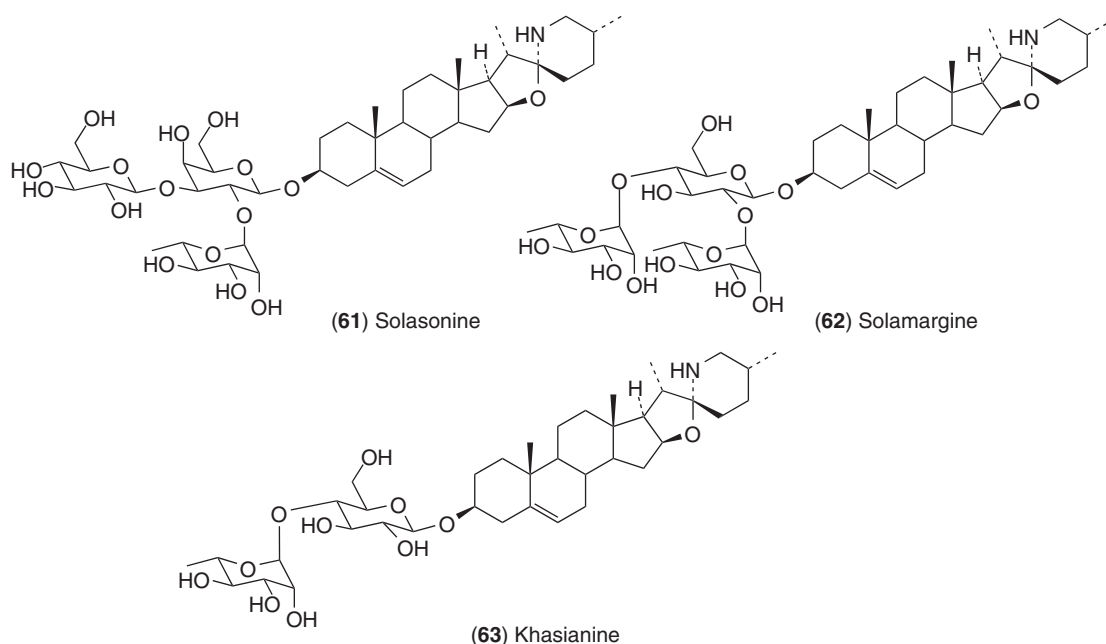
Figure 17 Some examples of steroidal alkaloids analyzed by TLC.

3.2 High Performance Liquid Chromatography (HPLC)

HPLC is an efficient technique for separating a mixture of organic compounds for the purpose of identifying, quantifying, and purifying the individual components. As for all chromatographic techniques, HPLC also consists of a stationary phase and a mobile phase. Compared to open column liquid chromatography, HPLC utilizes columns consisting of densely packed stationary phase having small porous particles (5 μm) and mobile phases under elevated pressures (400–600 bar), provided by a suitable pump. The porous particles in the column usually have a chemically bonded phase on their surface, which interacts with the components of the sample mixture as they pass through the column to separate them from one another causing each component to elute at different retention times (t_R, the time taken for the compound to elute from the column after injection). These features bring about improved separation of mixtures of compounds. Basic components of a typical HPLC instrument are shown in Figure 18. In general, HPLC separation of plant steroids is carried out using the RP and NP chromatographic modes. Some important

characteristics of these two methods are summarized in Table 6.

HPLC systems used for plant steroid analysis are usually equipped with, UV–vis, photodiode array (PDA; DAD, diode array detector), refractive index (RI), evaporative light scattering (ELS), and FL detectors. Additionally, MS detectors are often used to detect and identify the compounds present in a sample. Quantification of the individual components is achieved by the measurement of peak areas in comparison to a known amount of the standard sample. The most common method of identification (ID) of individual compounds in the sample is to use its retention time (t_R). Depending on the detector used, compound identification could be based on molecular parameters such as the chemical structure or molecular weight. (See **HPLC and Ultra HPLC: Basic Concepts** for a detailed description of HPLC.)

Polarities of plant steroids span in a wide range, from relatively less-polar sterols to high polar steroidal glycoalkaloids, through medium polar steroids including polyhydroxylated withanolides, brassinosteroids, and phytoecdysteroids. HPLC has become the method of choice for the efficient separation, identification, and quantification of each of these groups of steroids as it offers the possibility of using a range of columns with different bonded
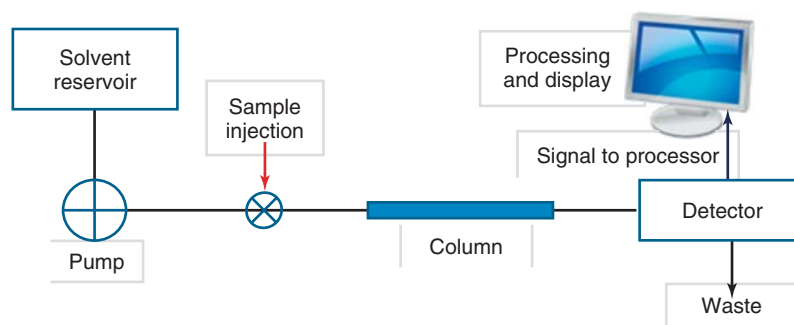
Figure 18 Basic components of a typical HPLC system.

Table 6 Comparison of RP- and NP-HPLC methods.

Mode	Stationary phase	Mobile phase	Applications
Reversed phase (RP)	Non polar Bonded phases of C_{18}, C_8, C_3, phenyl, etc.	H_2O and H_2O miscible solvents (e.g., MeCN, MeOH)	Nonpolar, polar, ionizable, and ionic molecules including many classes of steroids
Normal phase (NP)	Polar SiO_2, bonded phases of cyanopropyl, amino, etc.	Organic solvents (e.g., hexane, *iso*-octane, CH_2Cl_2, EtOAc, MeOH)	Water-sensitive compounds. Useful in the separation of stereoisomers

phases and sizes and also mobile phases with a variety of modifiers. Most steroids have been detected by PDA and ELS detectors. The use of HPLC in steroid analysis has recently been reviewed (Dinan, Harmatha, and Lafont, 2011).

3.2.1 HPLC Analysis of Phytosterols

Application of RP-HPLC hyphenated with atmospheric pressure chemical ionization-tandem mass spectrometry (APCI-MS/MS) in the analysis of sterols with high sensitivity has recently been reported (Igarashi *et al*., 2011). Application of this analytical technique has enabled the detection of β-sitosterol (**2**), campesterol (**3**), stigmasterol (**4**), cholesterol (**6**), ergosterol (**7**) (Figures 2 and 3), 7-dehydrocholesterol (**64**), and desmosterol (**65**) (Figure 19) in silkworm, mulberry leaves, and a silkworm artificial diet at femtomole (10^{-15} moles) quantities. To establish the detection method of the above-mentioned sterols and 3,4-$^{13}C_2$-cholesterol, the standard samples of them were first analyzed by APCI-MS and APCI-MS/MS and detected as $[M + H - H_2O]^+$ ions by selected reaction monitoring (SRM) mode of APCI-MS. In the

APCI-MS/MS, each dehydrated molecular ion was subjected to collisionally activated dissociation (CAD) to generate fragment ions that were then analyzed by comparing the fragmentation pattern of 3,4-$^{13}C_2$-cholesterol. The sterol mixtures have been separated using RP-C_{18} column eluting with acetonitrile (MeCN) and detected by APCI-MS as $[M + H - H_2O]^+$ ions and identified by APCI-MS/MS using the results obtained for standard samples.

3.2.2 HPLC Analysis of Withanolides

Biologically active constituents of *W. somnifera* (L.) Dunal, a plant used in Ayurvedic medicine and for the preparation of the herbal supplement, Ashwagandha or Indian ginseng, have long been recognized as a group of steroidal lactones collectively known as *withanolides*. Hence, there exist extensive literature reports with regard to the analysis of withanolides in *W. somnifera* (L.) Dunal (Dinan, Harmatha, and Lafont, 2011). Analytical techniques used for standardization of *W. somnifera* (L.) Dunal and commercial products containing this plant is limited

(64) 7-Dehydrocholesterol   **(65)** Desmosterol
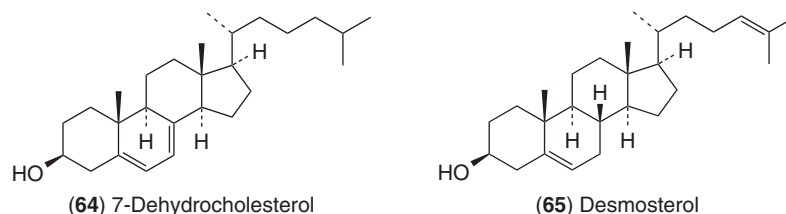
Figure 19 Some examples of phytosterols analyzed by HPLC.
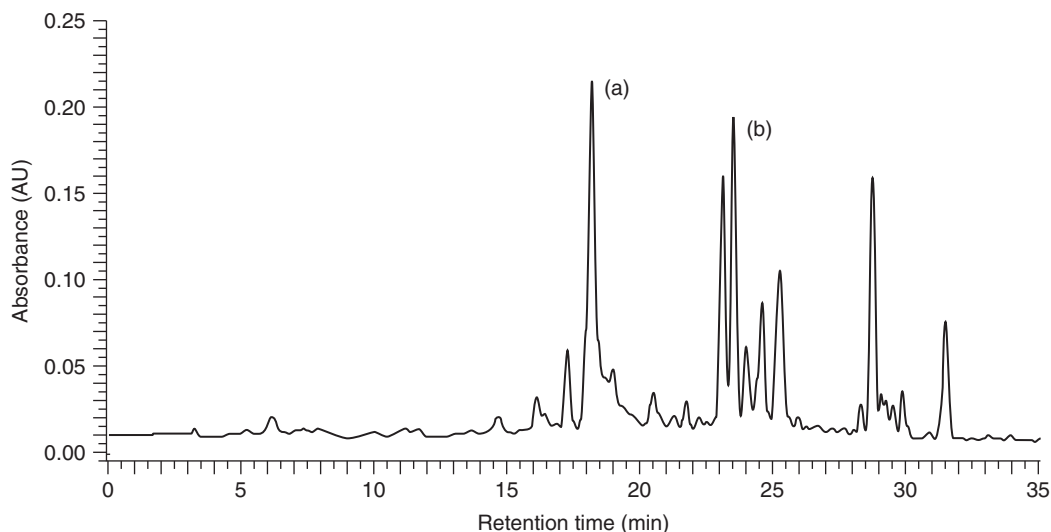
Figure 20 An HPLC trace of a methanol extract of *W. somnifera* (L.) Dunal root. Peak (a) 2,3-dihydrowithaferin A-3β-*O*-sulfate (**66**), peak (b) withaferin A (**9**). Conditions: column, Phenomenex, C_{18}, 5 μm, 4.6 × 250 mm; mobile phase, a gradient from 40% to 100% MeOH in 30 min, flow rate 0.7 mL min^{-1}; detection, UV at 250 nm.

to the determination of major withanolides, withaferin A (**9**), and/or withanolide D (**11**) (Figure 4) as they are claimed to be the bioactive constituents of *W. somnifera* (L.) Dunal. The occurrence of another bioactive constituent and a prodrug of withaferin A, 2,3-dihydrowithaferin A-3β-*O*-sulfate (**66**) (Figure 21), in aerophonically grown *W. somnifera* (L.) Dunal and its conversion into withaferin A (**9**) under physiological conditions has recently been reported (Xu *et al*., 2011). An HPLC trace of a methanol extract of *W. somnifera* (L.) Dunal root is depicted in Figure 20.

An HPLC analysis of nine withanolides, withaferin A (**9**), withanolide A (**10**), withanolide D (**11**) (Figure 4), 27-deoxywithaferin A (**67**), 17-hydroxy-withaferin A (**68**), 17-hydroxy-27-deoxywithaferin A (**69**), withanone (**70**), 27-hydroxywithanone (**71**), and 27-hydroxywithanolide B (**72**) (Figure 21), in *W. somnifera* (L.) Dunal has also been reported

(Chaurasiya *et al*., 2008). This HPLC analysis has been performed on a RP column with a binary mobile phase consisting of 0.1% HOAc in H_2O and 0.1% HOAc in MeOH using a gradient elution. All withanolides have been detected using both PDA (at 227 nm) and ELS detectors. This method has been applied to the analysis of leaf and root extracts of *W. somnifera* (L.) Dunal for identification and quantification of constituent withanolides present.

3.2.3 HPLC Analysis of Brassinosteroids

Separation and quantification of brassinosteroids using HPLC hyphenated with ESI-MS (electrospray ionization mass spectrometry) has been reported during an investigation of uptake of exogenous 24-*epi*brassinolide (**14**) (Figure 5) and its effect on

(66) 2,3-Dihydrowithaferin A-3 β-O-sulfate

	R₁	R₂	
(67)	H	H	27-Deoxywithaferin A
(68)	OH	OH	17-Hydroxywithaferin A
(69)	OH	H	17-Hydroxy-27-deoxywithaferin A

	R₁	R₂	R₃	
(70)	OH	H	H	Withanone
(71)	OH	H	OH	27-Hydroxywithanone
(72)	H	H	OH	27-Hydroxywithanolide B
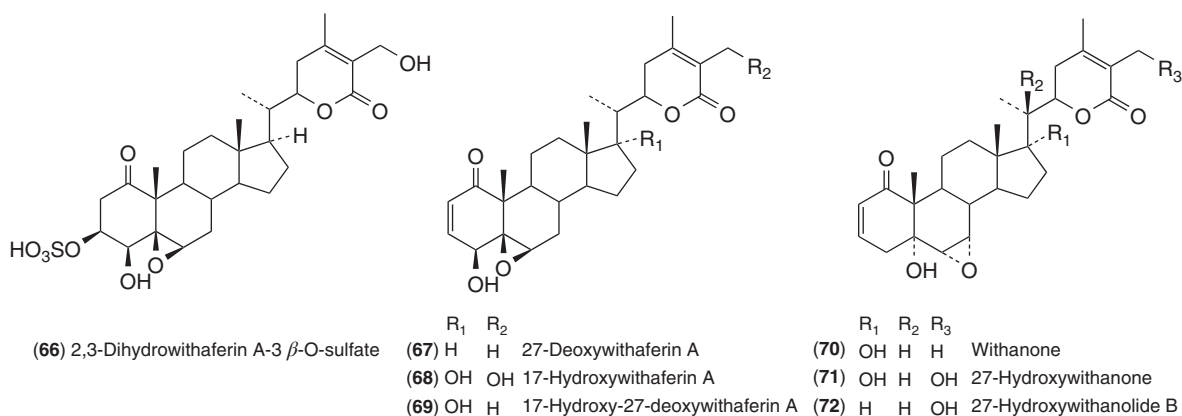
Figure 21 Some examples of withanolides analyzed by HPLC.

R = Alkyl or Aryl
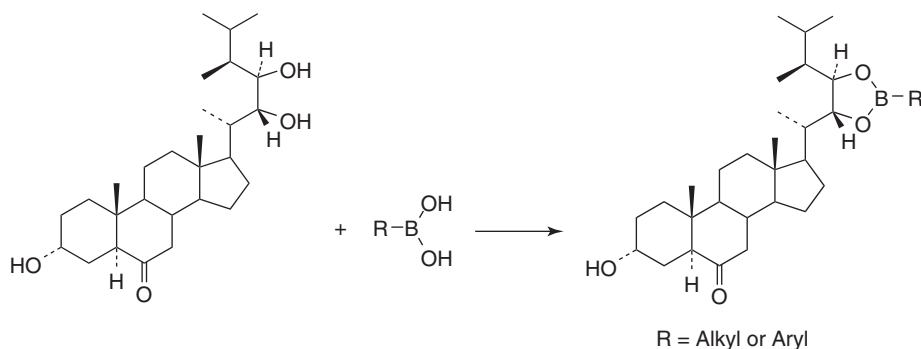
Figure 22 Formation of cyclic boronates from brassinosteroids.

the endogenous brassinosteroid content of wheat seedlings (Janeczko and Swaczynová, 2010). In this study, a semi-purified fraction containing brassinosteroids has been analyzed, using an RP C_{18} column and a mobile phase of MeOH containing 5 mM aqueous HCO_2H, which revealed the presence of brassinolide (**12**), castasterone (**13**), and 24-*epi*brassinolide (**14**) (Figure 5) in wheat leaves. A PDA detector at 210–400 nm and ESI-MS have been employed for the detection of these brassinosteroids. Quantification of the brassinosteroids were carried out by comparing the areas of m/z $[M + Na]^+$ peaks with those of isotopically labeled and authentic brassinosteroids in the selective ion monitoring (SIM) mode.

Brassinosteroids are known to occur in plant tissues in very low concentrations and as complex mixtures. Thus, they present a considerable challenge in their detection, quantification, and identification. Analysis of brassinosteroids as their boronate derivatives by LC/MS has been developed using either APCI or ESI

techniques. It has been reported that brassinosteroids with vicinal diol functionality in their side chains form stable five-membered ring boronates when reacted with organic boronic acids in the presence of pyridine (Gamoh *et al*., 1996; Svatoš, Antonchick, and Schneider, 2004) (Figure 22). Brassinolide (**12**), castasterone (**13**) (Figure 5), teasterone (**73**), and typhasterol (**74**) (Figure 23) have been analyzed as their naphthaleneboronates using LC/APCI-MS (Gamoh *et al*., 1996). In this analysis, the separation of the brassinosteroids has been achieved on a RP C_{18} column and elution with a gradient of MeCN and H_2O. Derivatization of these brassinosteroids to their naphthaleneboronates provided the advantage of detecting them using a UV detector (at 280 nm).

In a study directed toward understanding the biosynthesis of brassinosteroids in *Arabidopsis thaliana* (L.) Heynh., a highly sensitive and selective LC/ESI-MS method has been developed for the analysis of brassinolide (**12**), castasterone (**13**)

(73) R = α-H, β-OH   Teasterone
(74) R = α-OH, β-H   Typhasterol

(75) 24-*Epi*castasterone
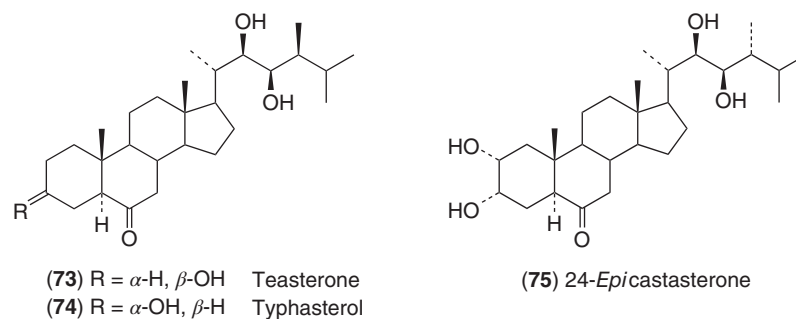
Figure 23   Some examples of brassinosteroids analyzed by HPLC.

(Figure 5), 24-*epi*castasterone (**75**) (Figure 23), and the deuterium-labeled brassinosteroids, [26,28-^{2}H$_6$]castasterone, [26,28-^{2}H$_6$]brassinolide, [26-^{2}H$_3$]3-*epi*castasterone, and [26-^{2}H$_3$]3-*epi*brassinolide in the plant extract as their dansyl-3-aminophenylboronates (Svatoš, Antonchick, Schneider, 2004). In this analysis, derivatization of brassinosteroids with dansyl-3-aminophenylboronic acid and the use of a microbore RP C$_{18}$ column have allowed achieving much lower LOD of 10 ng mL^{-1}.

3.2.4   HPLC Analysis of Phytoecdysteroids

A recent review by Dinan, Harmatha, and Lafont (2011) provides a comprehensive account of the analyses of phytoecdysteroids by HPLC. Polarities of phytoecdysteroids vary over a wide range because they usually contain 4–9 hydroxyl groups and often occur as their glycosides or esters. Both NP- and RP-HPLC have been used for the analysis of phytoecdysteroids. Unlike many other steroids, they can be easily detected by UV because of the αβ-unsaturated ketone chromophore present in them. Phytoecdysteroid profiles of seven plant species of the genus *Silene* have been identified by a combination of several RP and NP-HPLC methods (Zibareva *et al.*, 2009). This study has led to the identification of 20-hydroxyecdysone (**15**), polypodine B (**16**) (Figure 6), 2-deoxyecdysone (**42**), 2-deoxy-20-hydroxyecdysone (**43**) (Figure 12), integristerone A (**76**), 26-hydroxypolypodine B (**77**), and 20,26-dihydroxyecdysone (**78**) (Figure 24) by co-chromatography with authentic compounds, whereas the two new phytoecdysteroids

26-hydroxyintegristerone A (**79**) and 2-deoxy-20-hydroxyecdysone-25-β-D-glucoside (**80**) (Figure 24) have been identified by isolation and spectroscopic analysis. It was evident that one HPLC method alone was not adequate to identify these phytoecdysteroids with certainty, because they could not only elute with non-phytoecdysteroids at the same t_R but also overlap with their isomers or analogs.

Determination of three major phytoecdysteroids, 20-hydroxyecdysone (**15**) (Figure 6), 2-deoxy-20-hydroxyecdysone (**43**) (Figure 12), and 2-deoxy-20-hydroxyecdysone 3-*O*-β-D-glucopyranoside (**81**) (Figure 24), along with five non-phytoecdysteroids belonging to different classes of natural products present in a traditional Chinese medicine was achieved by HPLC coupled with PDA and ESI-MS/MS detection (Shi *et al.*, 2007). This analysis was carried out using an RP C$_{18}$ column eluting with a gradient solvent system consisting of 20 mM NH$_4$OAc in 0.2% HCO$_2$H/H$_2$O and MeOH/MeCN (1 : 1) and detecting at wavelengths specific for different classes of compounds (248 nm for phytoecdysteroids). The ESI-MS spectra were acquired in positive mode to produce [M + Na]$^+$ or [M + H]$^+$ ions and the analysis of MS/MS spectra were used to identify individual compounds.

3.2.5   HPLC Analysis of Steroidal Alkaloids

Steroidal alkaloids often show similar polarities because of their structural similarities hindering their separation by TLC and open column chromatography. HPLC has become the method of choice for the analyses of steroidal alkaloids. Although the

	R$_1$	R$_2$	R$_3$	
(**76**)	OH	H	H	Integristerone A
(**77**)	H	OH	OH	26-Hydroxypolypodine B
(**78**)	H	H	OH	20,26-Dihydroxyecdysone
(**79**)	OH	H	OH	26-Hydroxyintegristerone A

(**80**) 2-Deoxy-20-hydroxyecdysone-25-β-D-glucoside

(**81**) 2-Deoxy-20-hydroxyecdysone 3-O-β-D-glucopyranoside
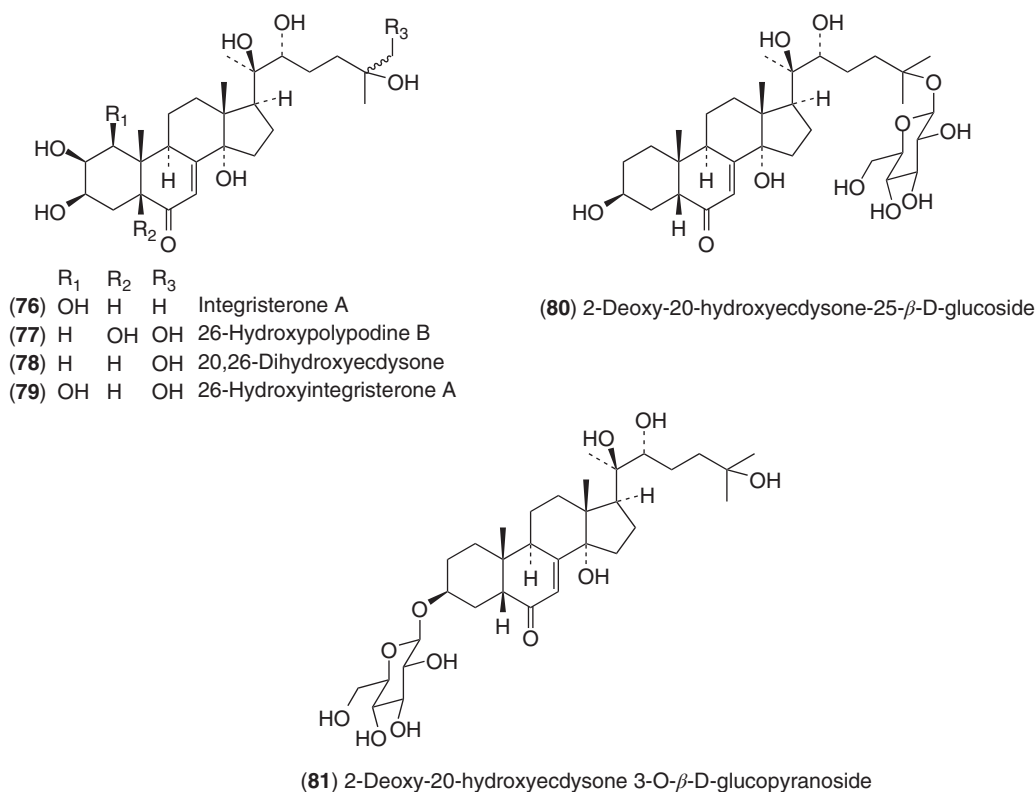
Figure 24   Some examples of phytoecdysteroids analyzed by HPLC.

absence of a chromophore in these compounds is a major drawback for their routine analysis, this problem is usually overcome by the use of either ELSD or MS for their detection. Separation and identification of 41 steroidal alkaloids (including some steroidal glycoalkaloids) of eight species of the genus *Fritillaria* has been achieved by RP-HPLC hyphenated with ESI-MS/MS detection (Zhou *et al.*, 2010). The HPLC separation utilized a C$_{18}$ column and a binary mobile phase consisting of 10 mM aqueous HCO$_2$NH$_4$ containing 0.1% HCO$_2$H and MeCN and their detection has been carried out by ESI-MS using a positive mode. Identification of 26 of the 41 steroidal alkaloids has been achieved by direct comparison of the HPLC retention times (t_R) and the fragmentation patterns in the MS/MS spectra with those of known steroidal alkaloids isolated from several *Fritillaria* species. The remaining 15 were identified by the analysis of their MS/MS data.

A similar application of HPLC/ESI-MS/MS characterization of 23 pregnane-type steroidal alkaloids present in the medicinal herb *Sarcococca coriacea* Hook. f (Buxaceae) has been reported (Musharraf *et al.*, 2012). Separation of the steroidal alkaloids in the extract of this plant utilized a C$_{18}$ capillary column and elution with a binary mobile phase consisting of 0.1% aqueous HCO$_2$H and MeCN. All 23 steroidal alkaloids have been detected by MS as [M + H]$^+$ ions and identified by analyzing the MS/MS data.

An LC/MS method has been developed for the fingerprinting of *S. xanthocarpum* Schrad. & J. C. Wendl. (Solanaceae) extracts and quantification of the constituent steroidal glycoalkaloids, solasonine (**61**), solamargine (**62**), and khasianine (β-2-solamargine) (**63**) (Figure 17) (Paul, Vir, and Bhutani, 2008). Optimum separation of these compounds has been achieved on a RP C$_{18}$ column using 0.5% aqueous HCO$_2$H and MeCN : 2-propanol : HCO$_2$H (94.5 : 5.0 : 0.5) as the mobile phase and the detection was carried out by ESI-MS in positive mode.

3.3 Ultra High Performance Liquid Chromatography (UHPLC)

UHPLC is an advanced technological breakthrough that does not compromise its resolving power even at increased flow rates. It uses chromatographic columns containing packing materials with decreased particle size (1.7 µm compared to 5 µm in HPLC) that provides improved peak capacity (number of peaks resolved per unit time in gradient separations) and greater resolution and sensitivity in comparison to conventional HPLC (Swartz, 2005; Zhao *et al.*, 2006). Hence, the use of UHPLC is advantageous when analytes are present in minute quantities and in large numbers. There are several notable examples of the use of UHPLC in the analyses of plant steroids. (See Chapter **HPLC and Ultra HPLC: Basic Concepts** for a detailed description of UHPLC.)

3.3.1 UHPLC Analysis of Mammalian Steroidal Hormones Present in Plants

Some plants contain mammalian steroidal hormones in trace quantities (Section 1.2.6). Therefore, the analysis of plants for these compounds requires separation and detection techniques of high sensitivity. Identification and quantification of mammalian steroidal hormones in *Digitalis purpurea* L., *N. tabacum* L., and *Inula helenium* L. leaf extracts by UHPLC/ESI-MS/MS has been reported (Simerský *et al.*, 2009). During the method development, 16 steroid standards including 4 with isotopic labels were separated by UHPLC using a short RP C_{18} column and eluting with a binary solvent gradient of MeOH and 10 mM aqueous HCO_2H. Constituent steroids were identified by detecting the transition of $[M+H]^+$ ions into

the appropriate product ions by MS/MS using multiple reaction monitoring (MRM) mode. This UHPLC/ESI-MS/MS method has been used to detect and quantify the mammalian steroidal hormones in above-mentioned plant extracts, which were subjected to a pre-immunoaffinity chromatographic purification step. The occurrence of progesterone (**30**) (Figure 10), 17α-hydroxyprogesterone (**82**), 16-dehydroprogesterone (**83**), and androstenedione (**84**) (Figure 25) in these three plant extracts together with their concentrations are summarized in Table 7.

3.4 Gas Liquid Chromatography (GLC)

GLC or gas chromatography (GC) is a technique for separation and analysis of compounds that could be vaporized without decomposition. GC can be used to verify the purity of a compound or separate components of a mixture of compounds with simultaneous quantification. As for all chromatographic techniques, GC also consists of a mobile phase (helium or nitrogen gas) and a stationary phase (a microscopic layer of a liquid or polymer on an inert solid support, inside a glass or metal column). As the gaseous mixture of compounds passes through the column, they are differentially partitioned between the carrier gas and the liquid stationary phase, causing each component to elute at a different time, known as the retention time (t_R) of the compound. Flame ionization detectors (FIDs) and thermal conductivity detectors (TCDs) are the most common detectors used with GC. In GC/MS, a mass spectrometer acts as the detector. Even at very low sample concentrations, GC/MS is highly sensitive and effective in analysis of components of a volatile mixture. (See **Headspace Sampling and Gas Chromatography: A Successful Combination to Study the Composition of a Plant Volatile Fraction** for a comprehensive discussion of GC.)

(**82**) 17α-Hydroxyprogesterone   (**83**) 16-Dehydroprogesterone   (**84**) Androstenedione
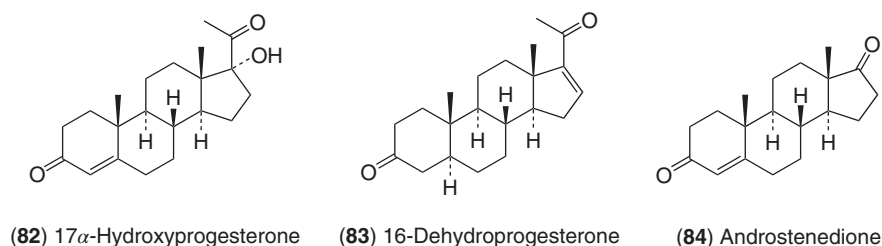
Figure 25   Some examples of mammalian steroidal hormones analyzed by UHPLC.

Table 7   Occurrence of mammalian steroid hormones in leaf extracts of *Digitalis purpurea* L., *Nicotiana tabacum* L., and *Inula helenium* L. (Simerský *et al.*, 2009)[a].

Plant	**30** (pmol g^{-1})	**82** (pmol g^{-1})	**83** (pmol g^{-1})	**84** (pmol g^{-1})
D. purpurea	58.92 ± 5.77	173.53 ± 20.70	28.94 ± 3.99	—
N. tabacum	55.46 ± 3.89	—	—	7.69 ± 2.13
I. helenium	2.10 ± 0.62	—	—	11.18 ± 2.98

[a] Concentrations are expressed with respect to fresh weight plant material.

Source: J Plant Growth Regul, 28, 2009, 125–136, Identification and Quantification of Several Mammalian Steroid Hormones in Plants by UPLC-MS/MS, Simerský R., Novák O., Morris D.A., Pouzar V. and Strnad M. With kind permission from Springer Science and Business Media.

(85) Δ^7-Avenasterol   **(86)** Brassicasterol
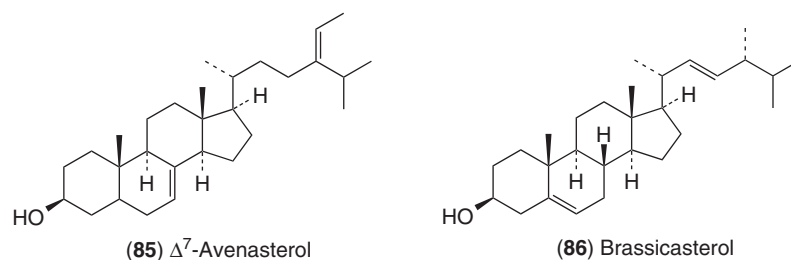
Figure 26   Same examples of phytosterols analyzed by GC-MS.

3.4.1   GC/MS Analysis of Phytosterols

Application of GC/MS analysis has been reported for the determination of total and individual sterol contents in 35 commonly available vegetables, fruits, and berries purchased from retail stores or market places in Helsinki, Finland (Piironen *et al.*, 2003). In this analysis, each of the fresh or freeze-dried samples was homogenized, an internal standard of cholesterol or dihydrocholesterol added, and subjected to hydrolysis to liberate free sterols from their glycosides and esters. Each sterol fraction thus obtained was purified by solid-phase extraction and reacted with *N,O*-bis(trimethylsilyl) trifluoroacetamide and trimethylchlorosilane (BSTFA/TMCS) to convert constituent phytosterols and the internal standard into their corresponding trimethylsilyl (TMS) ether derivatives. The samples were then analyzed by GC for the presence of β-sitosterol (**2**), stigmasterol (**3**), campesterol (**4**), Δ^5-avenasterol (**8**) (Figures 2 and 3), Δ^7-avenasterol (**85**), and brassicasterol (**86**) (Figure 26) as their TMS ether derivatives and were quantified by using internal standards and calibration curves of β-sitosterol, stigmasterol, campesterol, and sitostanol. Comparison of the GC retention times of the TMS ether derivatives of the constituent sterols with those of commercially available sterols (β-sitosterol, campesterol, stigmasterol, and sitostanol), mass spectral analysis, and comparison of these data with those reported in the literature allowed their identification.

3.4.2   GC/MS Analysis of Brassinosteroids

The presence of six brassinosteroids in *Camellia sinensis* (L.) Kuntze (tea plant) shoots has been established by the application of GC/MS analysis (Gupta *et al.*, 2004). In this analysis, fresh tea shoots were extracted with MeOH and fractionated by solvent–solvent partitioning and HPLC. The fractions containing brassinosteroids were then reacted with methaneboronic acid in pyridine to convert the constituent brassinosteroids into their volatile methylboronates (Figure 22, R = Me) and analyzed by GC/MS. Comparison of GC retention times and MS data of the methylboronates obtained above with those of methylboronates of known brassinosteroids revealed the presence of 24-*epi*brassinolide (**14**), typhasterol (**74**) (Figures 5 and 23), 6-deoxocastasterone (**87**), 3-dehydroteasterone (**88**), 6-deoxotyphasterol (**89**), and 28-homodolicholide (**90**) (Figure 27) in *C. sinensis* (L.) Kuntze.

	R₁	R₂	R₃	
(87)	α-OH, β-H	α-OH, β-H	H$_2$	6-Deoxocastasterone
(88)	H$_2$	O	O	3-Dehydroteasterone
(89)	H$_2$	α-OH, β-H	H$_2$	6-Deoxotyphasterol
(91)	α-OH, β-H	α-H, β-OH	O	3-*Epi*castasterone
(92)	α-H, β-OH	α-H, β-OH	O	2,3-Di*epi*castasterone

(90) 28-Homodolicholide
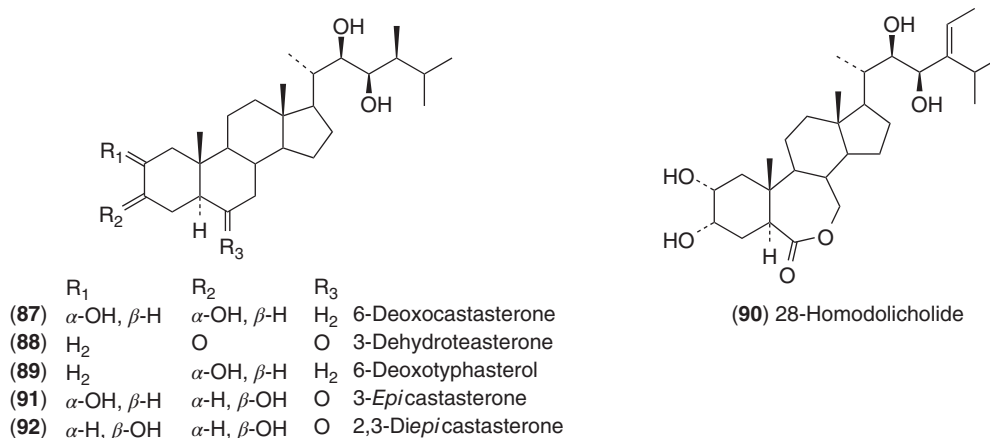
Figure 27 Some examples of brassinosteroids analyzed by GC-MS and HPLC.
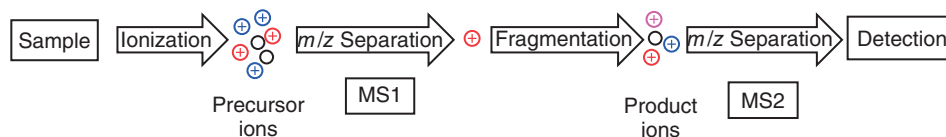
Figure 28 Schematics of a tandem mass spectrometer.

3.5 Tandem Mass Spectrometry (MS/MS)

MS/MS involves multiple stages of mass spectroscopic analysis with fragmentation of precursor ions into product ions between each of the stages. Mass spectrometry involving two stages is called MS/MS (Figure 28).

HPLC hyphenated with ESI-MS/MS is becoming an increasingly important and powerful technique in the analysis and identification of natural products. It offers several advantages over other analytical techniques including high sensitivity and short run time while providing considerable structural information rendering online identification of compounds (Musharraf *et al.*, 2012; Zhou *et al.*, 2010).

3.5.1 Application of Tandem Mass Spectrometry in the Analysis of Steroidal Alkaloids

During an analysis of steroidal alkaloids in eight species of plants belonging to the genus *Fritillaria* (Section 3.2.5) (Zhou *et al.*, 2010), MS/MS data

were obtained for 26 standards of C$_{27}$ steroidal alkaloids and their fragmentation pathways have been predicted. In this analysis, the [M + H]$^+$ ions were used as the precursor ions to generate the product ions, of which five most abundant ions were used for the structure elucidation. The optimum collision energy (CE) required for the fragmentation of the precursor ions providing maximum structural information has been found to be 70 eV. At this CE, the loss of substituent groups and cleavage of the steroid skeleton has resulted in the product ions. Identification of 41 steroidal alkaloids in these eight *Fritillaria* species has been achieved by direct comparison of fragmentation patterns with those of 26 standards and detailed analysis of the MS/MS data of the 15 remaining steroidal alkaloids.

3.6 Immunoassay and Immunoaffinity Chromatography of Plant Steroids

The basic principle of this assay is the ability of binding a steroid as an antigen to an antibody specifically raised against it. Detection

and quantification by immunoassay-based analysis have been reported only for a few groups of plant steroids. Commonly reported immunoassays include RIAs and enzyme-linked immunosorbent assays (ELISAs). Several RIA-based analysis of phytoecdysteroids and mammalian steroidal hormones in plants have been reported (Janeczko and Skoczowski 2005; Zibareva *et al.*, 2003). In immunoaffinity chromatography, the antibody raised against a steroid is immobilized on a rigid solid support to yield an active immunosorbent. When a mixture containing the steroid of interest is passed through a column of immunosorbent, it binds to the antibody, allowing its separation from other compounds. The bound steroid is then recovered by eluting the column with a suitable buffer solution.

3.6.1 Enzyme-Linked Immunosorbent Assay (ELISA) of Brassinosteroids

ELISAs have been developed for the quantification of 24-*epi*castasterone (**75**) (Figure 23) and related brassinolide analogs with the detection range of 0.005–50 pmol (Swaczynová *et al.*, 2007). Polyclonal antibodies used in this assay have been raised against 24-*epi*castasterone carboxymethyloxime-BSA conjugates and found to be highly specific to 24-*epi*castasterone (**75**) while cross reactivity has been shown to 24-*epi*brassinolide (**14**) and brassinolide (**12**) (Figure 5). It is noteworthy that these antibodies did not show cross reactivity with brassinosteroids that are epimeric at C-2, C-3, C-22, and C-23 of 24-*epi*castasterone (**75**) and with non-brassinosteroids.

3.6.2 Immunoaffinity Purification of Mammalian Steroidal Hormones from Plant Extracts

Purification of fractions containing mammalian steroidal hormones in several plants has been carried out using immunoaffinity chromatography (Simerský *et al.*, 2009). The polyclonal antibodies used in this study were raised against 4-androsten-3-one-17-carboxymethyloxime linked to BSA at C-17. These antibodies were found to be highly cross reactive with Δ^4-3-ketosteroids.

3.7 Detection of Plant Steroids by Biological Assays

Detection and analysis of plant steroids present in trace quantities, such as brassinosteroids, is a challenging task and could not be done by the application of chromatographic techniques alone. As brassinosteroids are capable of enhancing elongation, division, and expansion of plant cells (Clouse and Sasse, 1998), biological assays have been developed based on these physiological effects to detect brassinosteroids during the fractionation of plant extracts containing these compounds. One of the most frequently used biological assays for this purpose is rice lamina inclination assay (Clouse and Sasse, 1998). In a recent study involving characterization of 3-*epi*castasterone (**91**) and 2,3-di*epi*castasterone (**92**) (Figure 27), this assay has been used to detect brassinosteroids in the HPLC fractions of a partially purified seed extract of *Phaseolus vulgaris* L. (Kidney bean; Fabaceae) (Lee, Joo, and Kim, 2011). Other commonly used biological assays for the detection of brassinosteroids are radish hypocotyl bioassay (Gupta *et al.*, 2004; Sondhi *et al.*, 2008) and wheat leaf unrolling assay (Isidro *et al.*, 2012).

4 CONCLUSION

Plant steroids are a diverse group of secondary metabolites that can be classified into several groups based on taxonomic considerations and their functions and/or structures. They play important physiological functions within plants and exhibit pharmacological activities beneficial to the mankind. Plants containing steroids are being used in traditional medicines in various countries. On the basis of these traditional medical practices and with the increasing knowledge on the pharmacological activities of their constituent plant steroids, these medicinal plants are used as herbal supplements in many countries. Hence, it is important that plant raw materials as well as the final products are routinely analyzed for the detection and quantification of specific plant steroids.

TLC continues to be the method of choice, when rapid and reasonably accurate analysis of samples containing plant steroids is required. An improved separation of plant steroids and their quantification can be achieved with the use of HPTLC coupled

with TLC densitometry. Application of TLC is limited when the analytes are present in trace quantities. The possibility of using a range of columns with different bonded phases and sizes and mobile phases with varying modifiers, HPLC offers efficient separation, identification, and quantification of each group of steroids. Detection and quantification are commonly achieved by the use of PDA and ELS detectors. UHPLC can be used to achieve separation of a complex mixture of plant steroids when they are present in trace amounts and in large numbers. GLC can be employed to analyze plant steroids after converting them into volatile derivatives such as TMS ethers. Gas chromatography hyphenated with mass spectrometry (GC/MS) is a powerful analytical method for analyzing some groups of plant steroids. HPLC coupled to tandem mass spectrometry (LC/MS/MS) offers highly sensitive and relatively rapid technique for the analysis of plant steroids, requiring small amounts of sample. It provides considerable structural information, sometimes allowing online identification of compounds. Apart from these chromatography-based analytical tools, immunoassay methods and biological assay-based methods are also used for the analysis of plant steroids.

REFERENCES

Arnqvist, L., Dutta, P. C., Jonsson, L., *et al.* (2003) *Plant Physiol.*, **131**, 1792–1799.

Ata, A. and Andersh, B. J. (2008) *Alkaloids*, **66**, 191–213.

Atta-ur-Rahman, Zaheer-ul-Haq, Khalid, A., *et al.* (2002) *Helv. Chim. Acta*, **85**, 678–688.

Bajguz, A. and Tretyn, A. (2003) *Phytochemistry*, **62**, 1027–1046.

Báthori, M., Kalász, H., Janicsák, G., *et al.* (2003) *J. Liq. Chromatogr. Relat. Technol.*, **26**, 2629–2649.

Báthori, M. and Pongrácz, Z. (2005) *Curr. Med. Chem.*, **12**, 153–172.

Benjumea, D., Martín-Herrera, D., Abdala, S., *et al.* (2009) *J. Ethnopharmacol.*, **123**, 351–355.

Berges, R. R., Kassen, A. and Senge, T. (2000) *BJU Int.*, **85**, 842–846.

Bhawani, S. A., Sulaiman, O., Hashim, R., *et al.* (2010) *Trop. J. Pharm. Res.*, **9**, 301–313.

Carange, J., Longpré, F., Daoust, B., *et al.* (2011) *J. Toxicol.*, **2011**, 1–13 Article ID 392859, DOI: 10.1155/2011/392859.

Chaurasiya, N. D., Uniyal, G. C., Lal, P., *et al.* (2008) *Phytochem. Anal.*, **19**, 148–154.

Chen, L.-X., He, H. and Qiu, F. (2011) *Nat. Prod. Rep.*, **28**, 705–740.

Clouse, S. D. and Sasse, J. M. (1998) *Annu. Rev. Plant. Physiol. Plant. Mol. Biol.*, **49**, 427–51.

Deng, J., Liao, Z. and Chen, D. (2005) *Phytochemistry*, **66**, 1040–1051.

Devkar, S. T., Badhe, Y. S., Jagtap, S. D., *et al.* (2012) *J. Planar Chromatogr.*, **25**, 290–294.

Devkota, K. P., Lenta, B. N., Wansi, J. D. *et al.* (2008) *J. Nat. Prod.*, **71**, 1482–1484.

Devkota, K. P., Wansi, J. D., Lenta, B. N., *et al.* (2010) *Planta Med.*, **76**, 1022–1025.

Dinan, L., Harmatha, J. and Lafont, R. (2011) *Chromatogr. Sci.*, **102**, 679–708.

Dinan, L. (2001) *Phytochemistry*, **57**, 325–339.

Festucci-Buselli, R. A., Contim, L. A. S., Barbosa, L. C. A., *et al.* (2008) *Botany*, **86**, 978–987.

Furmanowa, M., Gajdzis-Kuls, D., Ruszkowska, J., *et al.* (2001) *Planta Med.*, **67**, 146–149.

Gamoh, K., Abe, H., Shimada, K., *et al.* (1996) *Rapid Commun. Mass Spectrom.*, **10**, 903–906.

Ginzberg, I., Tokuhisa, J. G. and Veilleux, R. E. (2009) *Potato Res.*, **52**, 1–15.

Grove, M. D., Spencer, G. F., Rohwedder, W. K., *et al.* (1979) *Nature*, **281**, 216–217.

Gupta, D., Bhardwaj, R., Nagar, P. K., *et al.* (2004) *Plant Growth Regul.*, **43**, 97–100.

Igarashi, F., Hikiba, J., Ogihara, M. H., *et al.* (2011) *Anal. Biochem.*, **419**, 123–132.

Isidro, J., Knox, R., Singh, A., *et al.* (2012) *Planta*, **236**, 273–281.

Jadhav, A. N., Rumalla, C. S., Avula, B., *et al.* (2007) *Chromatographia*, **66**, 797–800.

Janeczko, A. and Skoczowski, A. (2005) *Folia Histochemica Et Cytobiologica*, **43**, 71–79.

Janeczko, A. and Swaczynová, J. (2010) *Biol. Plant*, **54**, 477–482.

Janeczko, A. (2012) *Steroids*, **77**, 169–173.

Jayaprakasam, B. and Nair, M. G. (2003) *Tetrahedron*, **59**, 841–849.

Jones, P. J. H., MacDougall, D. E., Ntanios, F., *et al.* (1997) *Can. J. Physiol. Pharmacol.*, **75**, 217–227.

Jones, P. J. H., Ntanios, F. Y., Raeini-Sarjaz, M., *et al.* (1999) *Am. J. Clin. Nutr.*, **69**, 1144–1150.

Kim, T.-W., Hwang, J.-Y., Kim, Y.-S., *et al.* (2005) *Plant Cell*, **17**, 2397–2412.

Kreis, W. and Müller-Uri, F. (2010) *Annual Plant Rev.*, **40**, 304–363.

Kumar, N., Singh, B., Bhandari, P., *et al.* (2007) *Chem. Pharm. Bull.*, **55**, 912–914.

Kupchan, S. M., Doskotch, R. W., Bollinger, P., *et al.* (1965) *J. Am. Chem. Soc.*, **87**, 5805–5806.

Lafont, R., Harmatha, J., Marion-Poll, F., *et al.*, *The Ecdysone Handbook*, 3rdon-line, (updated in 2012) edn, 2002) http://ecdybase.org.

Lavie, D., Glotter, E. and Shvo, Y. (1965) *J. Chem. Soc.*, **7517–7531**.

Lee, S. C., Joo, S.-H. and Kim, S.-K. (2011) *J. Plant Biol.*, **54**, 10–14.

Leyon, P. V. and Kuttan, G. (2004) *Phytother. Res.*, **18**, 118–122.

Li, H.-J., Jiang, Y. and Li, P. (2006) *Nat. Prod. Rep.*, **23**, 735–752.

Mirjalili, M. H., Moyano, E., Bonfill, M., *et al.* (2009) *Molecules*, **14**, 2373–2393.

Misico, R. I., Nicotra, V. E., Oberti, J. C., *et al*. (2011) *Prog. Chem. Org. Nat. Prod.*, **94**, 127–229.

Moghadasian, M. H. and Frohlich, J. J. (1999) *Am. J. Med.*, **107**, 588–594.

Moghadasian, M. H. (2000) *Life Sci.*, **67**, 605–615.

Musharraf, S. G., Goher, M., Ali, A., *et al*. (2012) *Steroids*, **77**, 138–148.

Ostlund, Jr., R. E. Racette, S. B., Okeke, A., *et al*. (2002) *Am. J. Clin. Nutr.*, **75**, 1000–1004.

Ostlund, Jr., R. E. (2002) *Annu. Rev. Nutr.*, **22**, 533–549.

Paul, A. T., Vir, S. and Bhutani, K. K. (2008) *J. Chromatogr. A*, **1208**, 141–146.

Pauli, G. F., Friesen, J. B., Gödecke, T., *et al*. (2010) *J. Nat. Prod.*, **73**, 338–345.

Phi, T. D., Pham, V. C., Mai, H. D. T., *et al*. (2011) *J. Nat. Prod.*, **74**, 1236–1240.

Piironen, V., Toivo, J., Puupponen-Pimiä, R., *et al*. (2003) *J. Sci. Food Agric.*, **83**, 330–337.

Samadi, A. K., Tong, X., Mukerji, R., *et al*. (2010) *J. Nat. Prod.*, **73**, 1476–1481.

Sasse, J. M. (1997) *Physiol. Plant.*, **100**, 696–701.

Segura, R., Javierre, C., Lizarraga, M. A., *et al*. (2006) *Br. J. Nutr.*, **96**(Suppl. 2), S36–S44.

Shanker, K., Gupta, S., Srivastava, P., *et al*. (2011) *J. Pharm. Biomed. Anal.*, **54**, 497–502.

Shanker, K., Singh, S. C., Pant, S., *et al*. (2008) *Chromatographia*, **67**, 269–274.

Shi, Q., Yan, S., Liang, M., *et al*. (2007) *J. Pharm. Biomed. Anal.*, **43**, 994–999.

Shinde, P. B., Aragade, P. D., Agrawal, M. R., *et al*. (2011) *Indian J. Pharm. Sci.*, **73**, 240–243.

Simerský, R., Novák, O., Morris, D. A., *et al*. (2009) *J. Plant Growth Regul.*, **28**, 125–136.

Simons, R. G. and Grinwich, D. L. (1989) *Can. J. Bot.*, **67**, 288–296.

Singh, I. and Shono, M. (2005) *Plant Growth Regul.*, **47**, 111–119.

Sondhi, N., Bhardwaj, R., Kaur, S., *et al*. (2008) *Plant Growth Regul.*, **54**, 217–224.

Sun, L., Liu, J., Liu, P., *et al*. (2011) *Process Biochem.*, **46**, 482–488.

Svatoš, A., Antonchick, A. and Schneider, B. (2004) *Rapid Commun. Mass Spectrom.*, **18**, 816–821.

Swaczynová, J., Novák, O., Hauserová, E., *et al*. (2007) *J. Plant Growth Regul.*, **26**, 1–14.

Swartz, M.E. (2005) Ultra Performance Liquid Chromatography (UPLC): An Introduction. *Separation Science Redefined*, http://www.chromatographyonline.com/lcgc/data/articlestandard/lcgc/242005/164646/article.pdf (accessed 30 July 2012).

Tanaka, H., Putalun, W., Tsuzaki, C., *et al*. (1997) *FEBS Lett.*, **404**, 279–282.

Torres, M. C. M., das, F., Pinto, C. L., Braz-Filho, R., *et al*. (2011) *J. Nat. Prod.*, **74**, 2168–2173.

Valkonen, J. T. P., Keskitalo, M., Vasara, T., *et al*. (1996) *Crit. Rev. Plant Sci.*, **15**, 1–20.

Villaseñor, I. M., Angelada, J., Canlas, A. P., *et al*. (2002) *Phytother. Res.*, **16**, 417–421.

Wang, K., Sasaki, T., Li, W., *et al*. (2011) *Chem. Biodivers.*, **8**, 2277–2284.

Woyengo, T. A., Ramprasath, V. R. and Jones, P. J. H. (2009) *Eur. J. Clin. Nutr.*, **63**, 813–820.

Xu, Y.-M., Gao, S., Bunting, D. P., *et al*. (2011) *Phytochemistry*, **72**, 518–522.

Yan, Y.-X., Sun, Y., Chen, J.-C., *et al*. (2011) *Planta Med.*, **77**, 1725–1729.

Yang, Z.-D., Duan, D.-Z., Xue, W.-W., *et al*. (2012) *Life Sci.*, **90**, 929–933.

Zhai, H.-Y., Zhao, C., Zhang, N., *et al*. (2012) *J. Nat. Prod.*, **75**, 1305–1311.

Zhao, X., Wang, W., Wang, J., *et al*. (2006) *J. Sep. Sci.*, **29**, 2444–2451.

Zhou, J.-L., Xin, G.-Z., Shi, Z.-Q., *et al*. (2010) *J. Chromatogr. A*, **1217**, 7109–7122.

Zibareva, L., Volodin, V., Saatov, Z., *et al*. (2003) *Phytochemistry*, **64**, 499–517.

Zibareva, L., Yeriomina, V. I., Munkhjargal, N., *et al*. (2009) *Arch. Insect Biochem. Physiol.*, **72**, 234–248.

Chemical Analysis of Bryophytes

Yoshinori Asakawa[1], Agnieszka Ludwiczuk[2] and Masao Toyota[1]

[1]*Faculty of Pharmaceutical Sciences, Tokushima Bunri University, Yamashiro-cho, Tokushima, Japan* and [2]*Department of Pharmacognosy with Medicinal Plant Unit, Medical University of Lublin, Lublin, Poland*

1 INTRODUCTION

The bryophytes are found everywhere in the world except in the sea. They grow on trees and soil, in lakes, rivers, Antarctic islands, and even on the head of a lizard. The bryophytes are placed taxonomically between algae and pteridophytes; there are about 24,000 species in the world. They are further divided into three phyla: Bryophyta (mosses: 14,000 species) (Figure 1), Marchantiophyta (liverworts: 6000 species) (Figure 2), and Anthocerotophyta (hornworts: 300 species) (Figure 3).

They are considered to be the oldest terrestrial green plants, although no strong scientific evidence for this has appeared in the literature. This hypothesis was mainly based on the resemblance of the present-day liverworts to the first land-plant fossils, the spores of which date back almost 500 million years. Among the bryophytes, almost all the liverworts possess beautiful cellular oil bodies (Figure 4), which are peculiar, membrane-bound cell organelles that consist of ethereal terpenoids and aromatic oils suspended in carbohydrate- or protein-rich matrix, whereas mosses and hornworts do not. These oil bodies are a very important biological marker for the classification in Marchantiophyta (Asakawa, 1982, 1995, 2001, 2004; Asakawa and Ludwiczuk, 2008a, b; Asakawa *et al.*, 1979; Asakawa, Ludwiczuk, and Nagashima, 2013; Ludwiczuk and Asakawa, 2008,

2010; Ludwiczuk *et al.*, 2008, 2011; Sukkharak *et al.* 2011).

Bryophytes, especially mosses, have been considered to be very important green plants in many Japanese temples and shrines in Kyoto and Nara or parks in Kenrokuen, Kanazawa, to keep their gardens beautiful and calm. Phytochemistry of bryophytes has been neglected for a long time because they are morphologically very small and difficult to collect in large amounts as pure samples. Their identification is also very difficult, even under the microscope. They are considered to be nutritionally useless to human diets. In fact, no references concerning their use as food for humans have been seen. However, a number of bryophytes, in particular, mosses, have been widely used as medicinal plants in China to cure burns, bruises, external wounds, snake bites, pulmonary tuberculosis, neurasthenia, fractures, convulsions, scald, uropathy, pneumonia, neurasthenia, and so on (Garnier, Bezaniger-Beauquesne, and Debraux, 1969; Suire, 1975; Ding 1982; Wu, 1982; Ando and Matsuo, 1984; Asakawa, 1999). Many species of liverworts show characteristic fragrant odors and an intense hot and bitter taste. Some *Fissidens* Hedw. and *Rhodobryum* (Schimp.) Limpr. species belonging to Bryophyta produce sweet tasting secondary metabolites. Generally, bryophytes are not damaged by bacteria and fungi, insect larvae and adults, snails, slugs, and other mammals. Furthermore, some liverworts cause intense allergenic

Figure 1 A moss species.
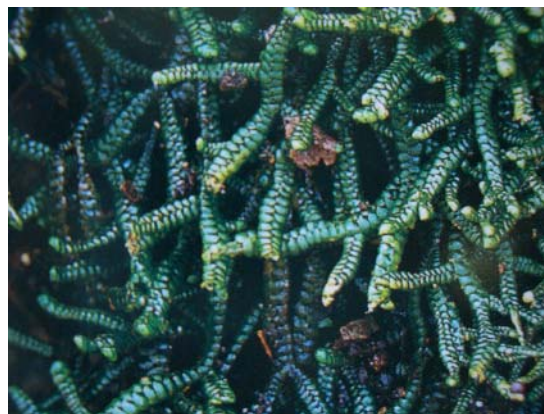
Figure 2 The liverwort, *Porella vernicosa*.

contact dermatitis and allelopathy. The authors have been interested in these biologically active substances found in bryophytes and have studied more than 1000 species of bryophytes collected in North and South America, Australia, China, Europe, French Polynesia, Greece, India, Indonesia, Japan, Madagascar, Malaysia, Nepal, New Zealand, Pakistan, Russia, Taiwan, and Turkey with respect to their chemistry, pharmacology, and application as sources of cosmetics and medicinal or agricultural drugs. The biological activities of liverworts are due to terpenoids, aromatic compounds, and acetogenins, which constitute oil bodies in each species (Asakawa, 1982, 1990a, 1990b, 1993, 1995, 2001, 2004, 2007, 2008a, b; Asakawa and Ludwiczuk, 2009; Asakawa *et al.*, 2008, 2009; Asakawa, Ludwiczuk, and Nagashima, 2013; Harinantenaina and Asakawa, 2007; Xie and Lou, 2009). In this chapter, bio- and chemical diversities, chemical analysis of secondary metabolites in bryophytes, and some biological activity, including characteristic odor and taste as well as chemosystematics of several liverworts, and chemical phylogeny of liverworts are discussed.

2 BIODIVERSITY OF BRYOPHYTES

As shown by the black numbers in Figure 5, there are 54 endemic genera in southern hemispheric countries such as New Zealand and Argentina (Inoue, 1988). In southeast Asia and Japan, relatively, a large number of endemic genera have been recorded; however, South Africa, Madagascar, and both North America

Figure 3 A hornwort.

and Europe are very poor regions of endemic genera. The red numbers in Figure 5 show the numbers of species in bryophytes (Mohamed *et al.*, 2008).

Marchantiophyta (liverworts) includes 2 subclasses, Jungermanniidae and Marchantiidae, 6 orders, 49 families, 130 genera, and 6000 species. Still many new species have been recorded in the literature. The richness of endemic genera of bryophytes in the southern hemisphere suggests that the bryophytes might have originated from the past Antarctic islands since 350,000,000–400,000,000 years ago and developed to the northern hemisphere with a long-range evolutionary process. In the southern hemisphere, New Zealand is the most charming country to see many different species of Marchantiophyta and Bryophyta, which are totally different from those found in northern Asia, including Japan. In

Figure 4 Oil bodies of the liverwort *Frullania vethii*.

Japan, Yaku Island, which has been protected by the UNESCO (United Nations Educational, Scientific and Cultural Organization), is the most interesting place to watch many species of the bryophytes.

In the tropical regions, such as southeast Asia, Borneo, Sumatra, and Papua New Guinea, and Colombia, Ecuador, and Venezuela, there are rain forests where many liverwort species have been found. But many different species, such as the Lejeuneaceae, are intermingled with each other and it is very time consuming to purify all of them. In Columbia and Ecuador, Marchantiophyta species grow in high mountains, over 2000 m where people live and not in the lower level of their lands.

3 CHEMICAL DIVERSITY OF BRYOPHYTES

Liverworts produce a great number of secondary metabolites, lipophilic terpenoids, acetogenins, and aromatic compounds. Typical terpenoids (**1–29**), aromatic compounds (**30–34**), and acetogenins (**35–46**) isolated from liverworts are shown in Figures 6–10. The most characteristic chemical phenomenon of liverworts is that most of the sesqui- and diterpenoids are enantiomers of those found in higher plants, although there are a few exceptions such as germacrane- and guaiane-type sesquiterpenoids. It is very noteworthy that different species of the same genera, such as *Frullania tamarisci* (L.) Dum. and *Frullania dilatata* (L.) Dum. (Frullaniaceae),
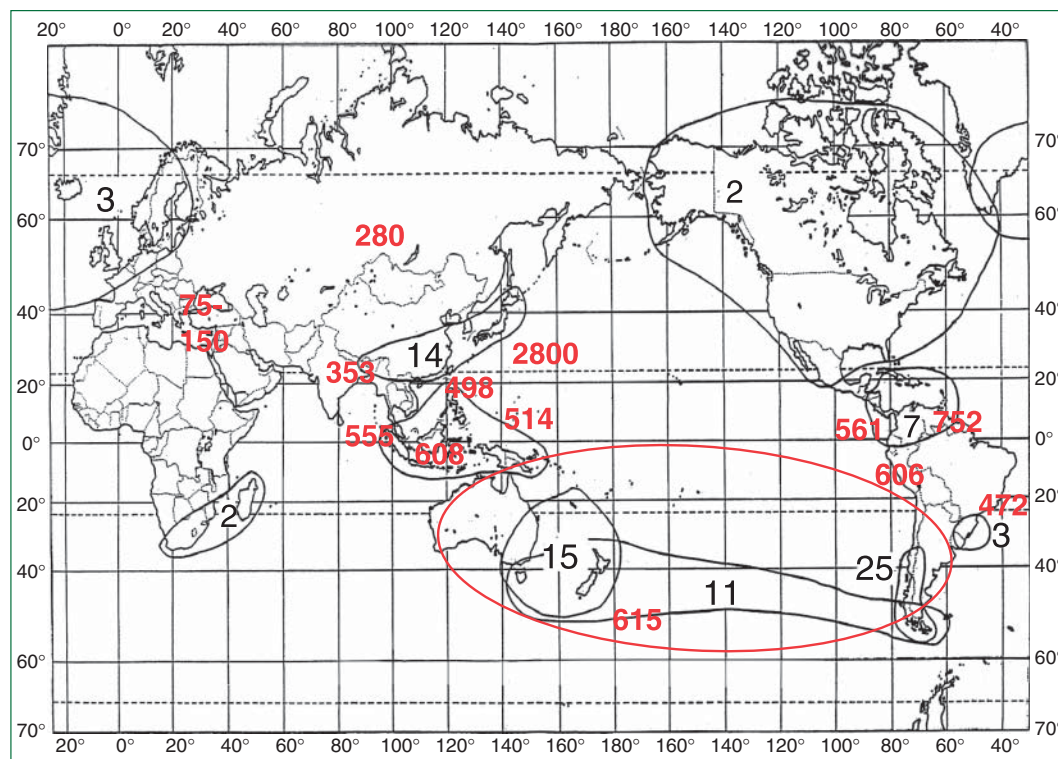
Figure 5 The distribution of bryophytes: Black numbers show endemic genus. The red numbers does species.

1 Norpinguisone   **2** Pinguisanin   **3** Biacutifolone A   **4** Plagiochilide

5 4(15)-Aromadendren-   **6** Plagiochilal A   **7** Furanoplagiochilal   **8** Dumortenol   **9** Dumortenol-6,7-epoxide
12,5-olide
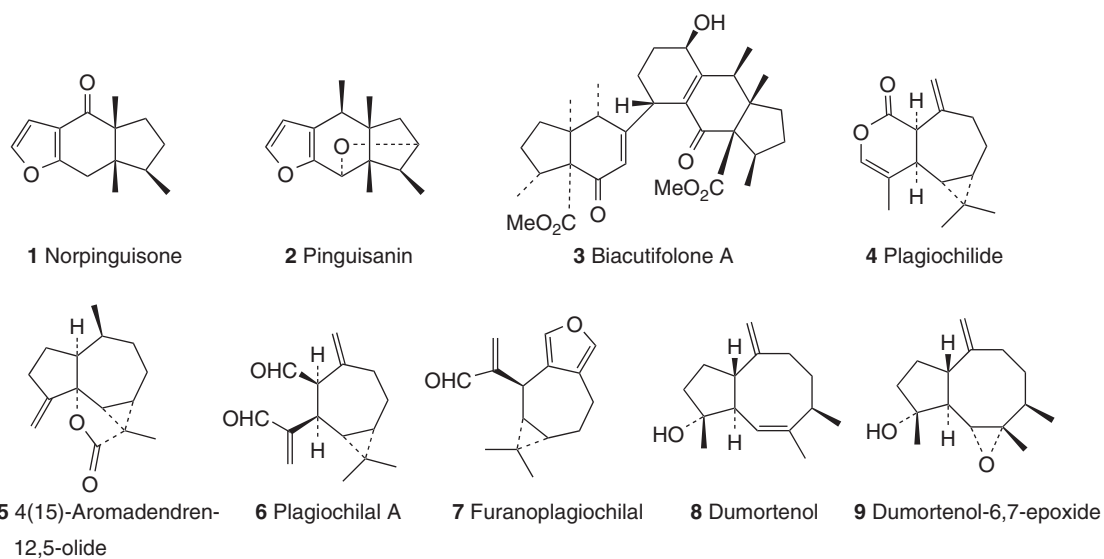
Figure 6   Typical sesquiterpenoids, pinguisane (**1–3**), 2,3-secoaromadendrane (**4–7**), aromadendrane (**5**), and dumortane (**8**,**9**) types.

produce different sesquiterpene lactone enatiomers (**67**, **68**). Some liverworts, such as *Lepidozia* (Dum.) Dum. species (Lepidoziaceae), biosynthesize both the enantiomers. Flavonoids are ubiquitous components in bryophytes and have been isolated from or detected in both Marchantiophyta and Bryophyta.

However, the presence of nitrogen-, sulfur-, or both nitrogen- and sulfur-containing compounds in bryophytes was very rare; recently, several nitrogen-containing compounds (**47–50**) have been isolated from the Mediterranean liverwort, *Corsinia coriandrina* (Spreng.) Lindb. belonging to Corsiniaceae (Marchantiales) (von Reuβ and König, 2005), two prenyl indole derivatives (**51**, **52**) from *Riccardia* A. Gray species belonging to Dilaniaceae (Asakawa, 1995), skatole (**53**) from *Asterella* P. Beauv. or *Mannia* Opiz belonging to Aytoniaceae (Asakawa *et al.*, 1995) and Tahitian *Cyathodium foetidissimum* Schiffn. belonging to Cyathodiaceae (Ludwiczuk *et al.*, 2009), and benzyl- (**54**) and β-phenethyl β-methylthioacrylates (**55**) from Isotachidaceae (Asakawa, 1995) (Figure 11).

Typical acetogenins are sex pheromones, dictyotene (**35**), ectocarpene (**36**), multifidene (**37**), and dictyopterene (**38**) from the liverworts *Fossombronia* Raddi and *Chandonanthus* Mitt. species, that have been isolated from brown algae (Ludwiczuk *et al.* 2008).

Highly evolved liverworts belonging to Marchantiaceae produce phytosterols such as campesterol, stigmasterol, and sitosterol. Almost all liverworts elaborate α-tocopherol and squalene. The characteristic components of Bryophyta are highly unsaturated fatty acids and alkanones, such as 5,8,11,14,17-eicosapentaenoic acid, 7,10,13,16,19-docosapentaenoic acid, and 10,13,16-nonadecatrien-7-yn-2-one, hopane-type triterpenoids, flavonoids, and benzonaphthoxanthenones, and some nitrogen-containing substances. The neolignan is one of the most important chemical markers of Anthocerotophyta (Asakawa, 1995). The presence of hydrophobic terpenoids is very rare in Marchantiophyta. A few bitter kaurene glycosides have been found in *Jungermannia* L. emend. Dum. species. However, a numbers of flavonoid glycosides have been detected in both Marchantiophyta and Bryophyta (Asakawa, 1982, 1995; Asakawa, Ludwiczuk, and Nagashima, 2013).

4   CHEMICAL ANALYSIS OF BRYOPHYTES

4.1   Collection and Purification of Bryophytes

Bryophytes, especially thalloid liverworts such as *Marchantia* species and hornworts, live in humid soil.
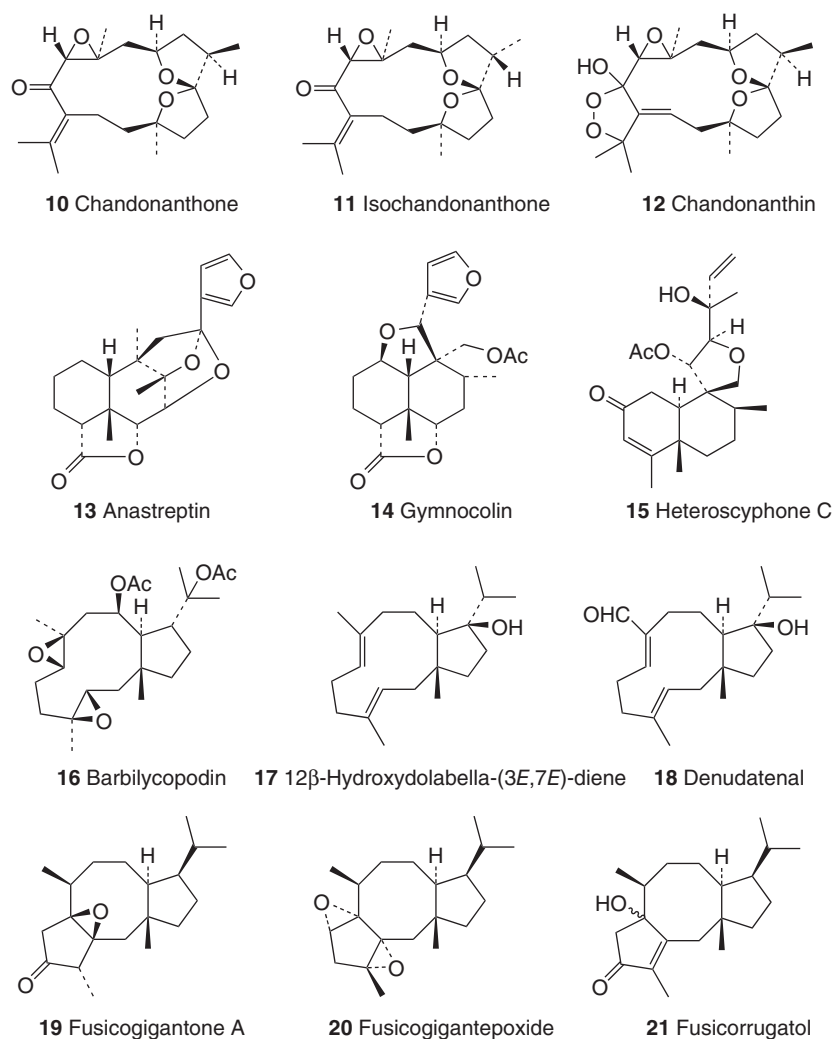
Figure 7 Typical diterpenoids, cembrane (**10–12**), clerodane (**13–15**), dolabellane (**19–20**), and fusicoccane (**19–21**) types.

Stem-leafy liverworts, *Porella* L., *Radula* Dum., and *Plagiochila* (Dum.) Dum. species, grow on wet rocks and many epiphytic liverworts, *Frullania* Raddi, *Ptychanthus* Nees, and *Lejeunea* Lib. species, on trunks of trees (such as oak and birch). *Riccardia* A. Gray, *Chiloscyphus* Corda, and *Gymnocolea* (Dum.) Dum. species have been found in lakes and ponds or rivers. On contrary, mosses, such as *Polytrichum* Hedw., *Mnium* Hedw., *Fissidens* Hedw., and *Pogonatum* P. Beauv., grow on dried or wet rocks or soil and *Sphagnum* L. species in ponds or on wet soil. *Funaria* Hedw. species grow in rivers. When one collects certain bryophyte species, many other species and/or dead leaves, twigs and/or roots of higher plants, and even small animals often intermingle into the species that one wishes to study chemically.

Thus, complete elimination of all such foreign materials should be very carefully carried out. Such time-consuming work is one of the reasons why bryophyte chemistry has been neglected for a century. Purification of bryophyte samples is normally carried out by hands or a sharp pincet; however, even purified samples still include very similar species when one treats very miniature species such as those of Lejeuneaceae family. Plastic trays or boxes, including plasticizer and phthalate esters (di-2-ethylhexyl
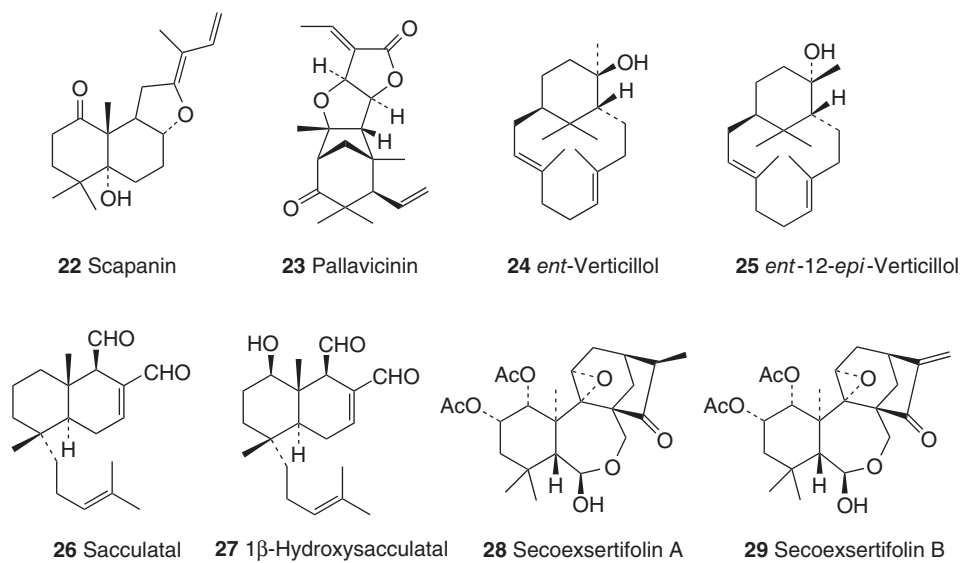
Figure 8 Typical diterpenoids, labdane (**22**, **23**), verticillane (**24**, **25**), sacculatane (**26**, **27**), and 6,7-secokaurane (**28**,**29**) types.

22 Scapanin   **23** Pallavicinin   **24** *ent*-Verticillol   **25** *ent*-12-*epi*-Verticillol

26 Sacculatal   **27** 1β-Hydroxysacculatal   **28** Secoexsertifolin A   **29** Secoexsertifolin B

30 R^1 = R^2 = H, R^3 = Me Riccardin A
31 R^1 = R^2 = R^3 = H   Riccardin C
31a R^1 = Me, R^2 = R^3 = H Riccardin F

32 Marchantin A

33 Pusilatin A

34 Pusilatin D
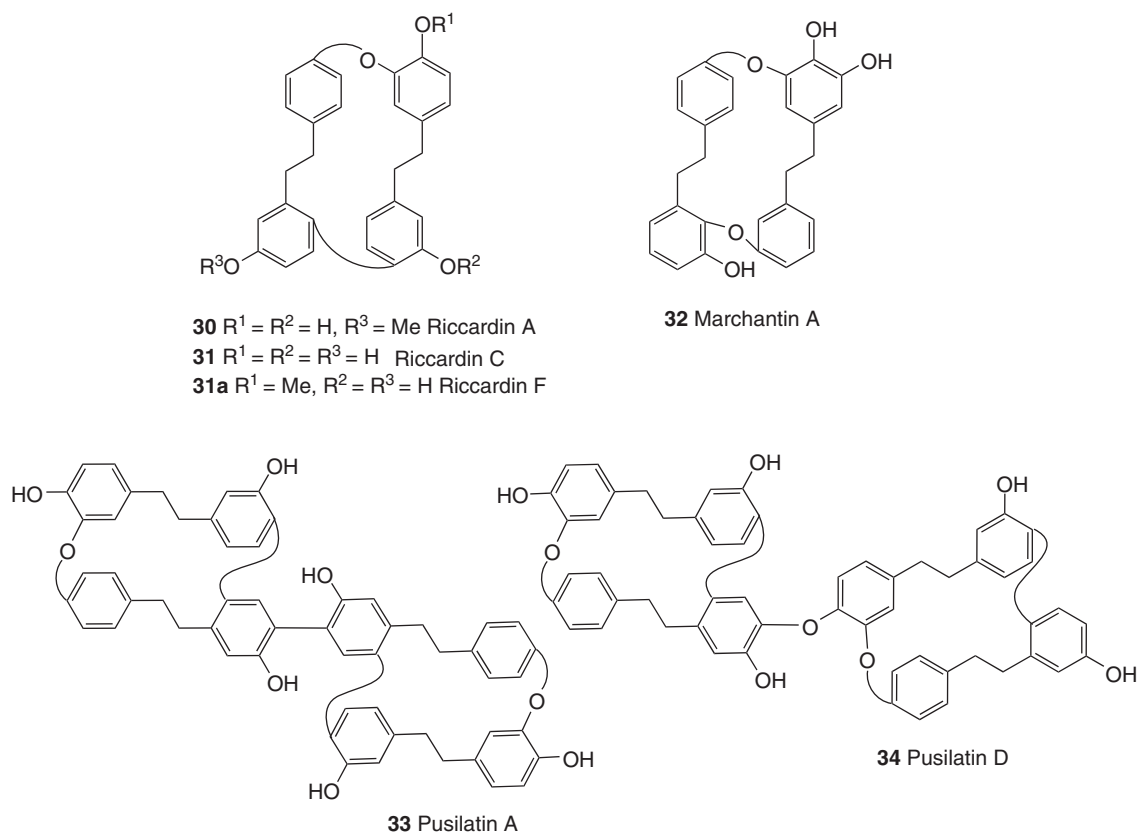
Figure 9 Typical aromatic compounds, bis-bibenzyls (**30–32**) and bis-bibenzyl dimers (**33**, **34**).
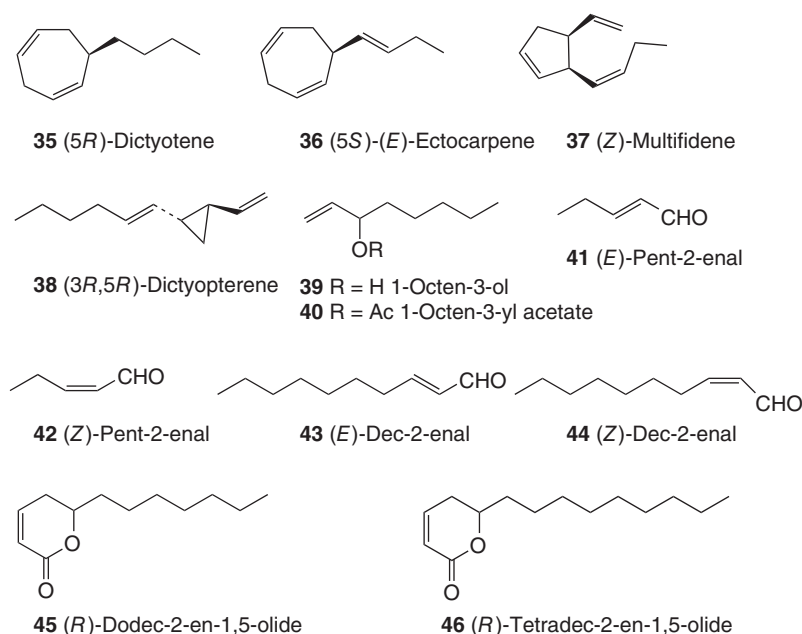
Figure 10 Typical acetogenins (**35–46**).

phthalate), are not recommended to carry liverwort samples. Once the samples are collected, a herbarium should be made (family, genus, species, date of collection and place, condition of collection site, collector name(s), and the identifier of the samples), and each herbarium should be stoked alphabetically on the stoke panel. The remaining samples should be gently dried under a shaded and cooler place and not be exposed to sunlight and stoked at room temperature or not <0°C. The authors recommend to use fresh materials directly to extract; otherwise, many volatile components evaporate or many unstable compounds oxidize during a long stocking and air exposure.

4.2 Extraction of Secondary Metabolites of Bryophytes

4.2.1 Solvent Extraction

The collected and purified fresh samples, such as Metzgeriales species, which are very thin thalloid liverworts, are directly extracted with an organic solvent, *n*-hexane, or ether, recommendable. The cellular oil bodies are easily extracted for a few minutes using ultrasonic apparatus. A large sample is

extracted with the above-mentioned solvents in a large flask by routine method. Methanol is often used for the extraction to obtain hydrophilic compounds. Soxhlet apparatus using ethanol is not recommended to extract the samples because this method damages many unstable terpenoids, including dialdehydes and hemiacetals, and many highly unsaturated fatty acids and aromatic compounds during a long extraction period. Ethyl acetate is hydrolyzed by the acidic components of bryophytes to give acetic acid and ethanol, which react with certain compounds in liverwort extracts during a long-time extraction period, and many different artifacts are created. Almost all of the mosses and some liverworts, for example, *Marchantia* L., *Conocephalum* Hill, and *Wiesnerella* Schiffn. belonging to Marchantiales, are very hard. In order to obtain large amounts of their crude extract, these samples are mechanically ground; however, the machine is heated sometimes until more than 100°C. The authors recommend a machine equipped with water-cooler circulation. The powders obtained are extracted with the solvent mentioned earlier under dark room. Evaporation of the solvent should also be very carefully carried out. Too much vacuum and high temperature are dangerous as many mono- and sesquiterpene hydrocarbons and volatile aromatic compounds are lost.

47 (*Z*)-Coriandrin   **48** (*E*)-Coriandrin   **49** (*Z*)-*O*-Methyltridentatol B

50 (*E*)-*O*-Methyltridentatol A

53 Skatole

51 7-(3-Methyl-2-butenyl)-indole   **52** 6-(3-Methyl-2-butenyl)-indole

54 Isotachin A   **55** Isotachin B   **56** Chandolide
(=Benzyl *trans*-b-methylthioacrylate)   (=β-Phenylethyl *trans*-b-methylthioacrylate)

57a GBB A   **57b** GBB B
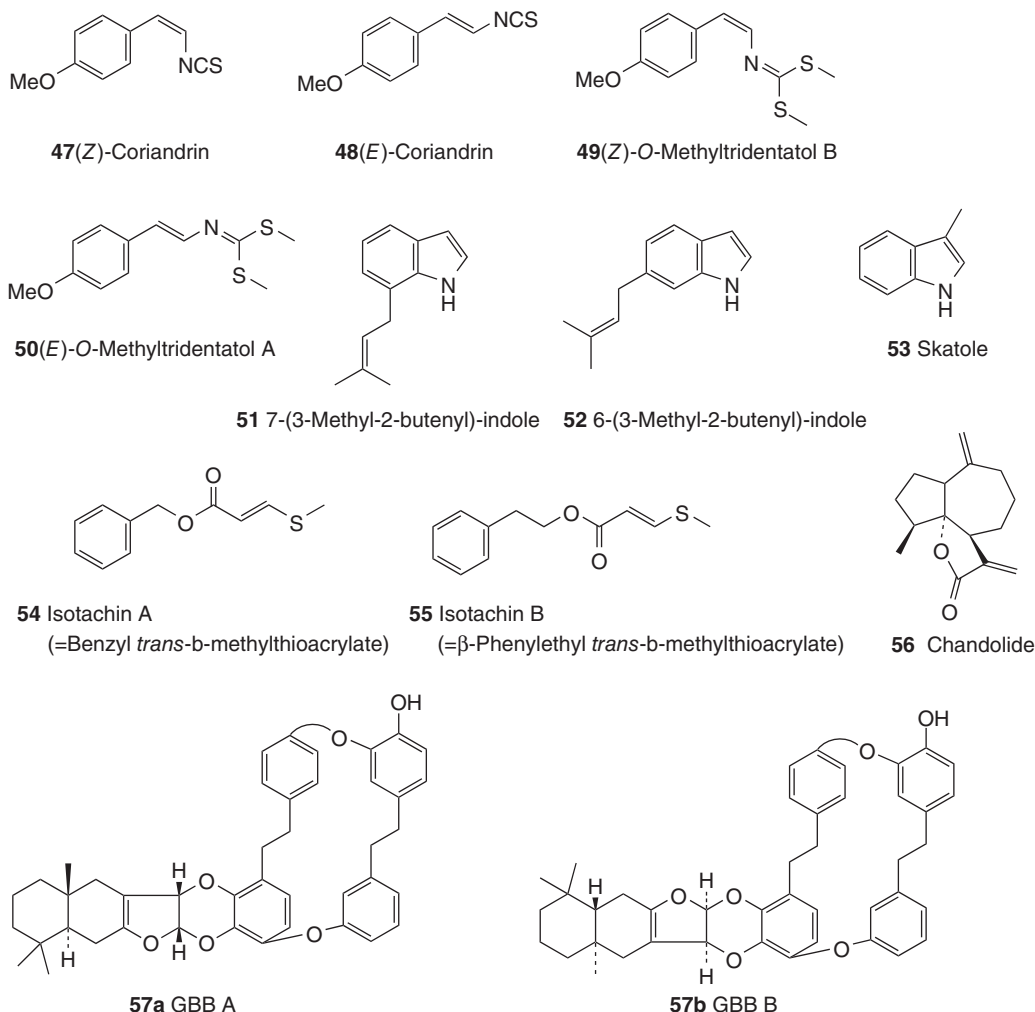
Figure 11 Typical aromatic compounds, nitrogen-sulfur-containing aromatics (**47–50**), nitrogen-containing aromatics (**51–53**), sulfur-containing aromatics (**54**, **55**), zierane type (**56**) and sesquiterpene-bis-bibenzyls (**57a**, **57b**).

4.2.2 Hydrodistillation and Steam Distillation of Bryophytes

In order to obtain volatile components, this method is sometimes used for the extraction of liverworts. However, high temperature and exposure to air and hot water during a long-time extraction decomposes many unstable terpenoids and aromatic acetogenins to form a number of artifacts. But this is the simplest methodology to obtain a small amount of volatile components for gas chromatography (GC) or gas chromatography/mass spectrometry (GC/MS) analysis (see later).

Steam distillation is not usually carried out to obtain volatile components from bryophytes because a large amount of essential oils is not obtained. The disadvantage of this method is the same as mentioned earlier.

A drastic condition of high pressure steam destroys many unstable compounds. Hence, comparison of chemical components obtained by both hydrodistillation and steam distillation and those from solvent extracts is absolutely necessary to recognize whether they are natural or nonnatural products. At present, over several hundred new compounds have been isolated from bryophytes and their structures elucidated

(Asakawa, 1982, 1995, 2004, Asakawa, Ludwiczuk, and Nagashima, 2013). Most of the compounds found in liverworts are composed of lipophilic mono-, sesqui-, and diterpenoids and aromatic compounds, such as typical bibenzyls and bis-bibenzyls, which have been isolated from Marchantiaceae and Aytoniaceae in Marchantiales, Lejeuneaceae, Lepidoziaceae, and Plagiochilaceae in Jungermanniales, and Blasiaceae, Pelliaceae, and Riccardiaceae in Metzgeriales.

4.3 TLC Analysis

The thin-layer chromatography (TLC) analysis is the simplest, fastest, and cheapest method to confirm the presence or absence of a number of components and their properties on the TLC plate. As oil bodies of liverworts are composed of lipophilic terpenoids, aromatic compounds, and acetogenins, TLC analysis is routinely used for the crude extract. *n*-Hexane and ethyl acetate (4 : 1 and 1 : 1) are usually used as the development solvent and the presence of spots is visualized by UV (ultraviolet) lights (254 and 356 nm) or spraying of Godin reagent or 30% sulfuric acid and heating at 110°C.

Figure 12 shows the TLC results of several ether extracts of the New Zealand liverworts. Very beautiful and clear spots are seen in the plates. A major compound is easily purified on the TLC plate using UV light or iodide vapor in the glass vessel.

4.4 Essential Oils

GC/MS of essential oils obtained from three liverworts, *Pallavicinia subciliata* (Austin) Steph., *Plagiochila sciophila* Nees ex Lindb., and *Porella japonica* (Sande Lac.) Mitt., are shown in Figures 13–15.

Each essential oil contains sesquiterpene hydrocarbons and oxygenated sesquiterpenoids. The essential oils and volatile components obtained from hydrodistillation of many liverworts show the presence of various monoterpenoids and complex sesqui- and diterpenoids as well as simple aromatic compounds on GC/MS. These GC/MS fingerprints are a very important tool for the identification of each liverwort.

4.5 GC and GC/MS Analysis

The most powerful methods to identify the chemical constituents of the crude extract of the liverworts and essential oils are GC and GC/MS.

GC columns, 25-m fused silica capillaries with polysiloxane CPSil-5 for analytical column and polysiloxane CPSil-19 (Chrompack) or 25-m fused silica capillaries with octakis(2,6-di-*O*-methyl-3-*O*-pentyl)-γ-cyclodextrin, heptakis(2,6-di-*O*-methyl-3-*O*-pentyl)-β-cyclodextrin, or heptakis(6-*O*-*tert*-butyldimethylsilyl-2,3-di-*O*-methyl-β-cyclodextrin in OV-1701 (50%, w/w), are

yM - 5 yM - 14 yM - 4 yM - 16   yM - 15 yM - 11 yM - 2 yM - 20   yM - 12 yM - 17 yM - 13 yM - 21   yM - 6 yM - 7 yM - 10 yM - 19   yM - 22   yM - 1   yM - 3   yM - 18 yM - 8 yM - 9 yM - 3
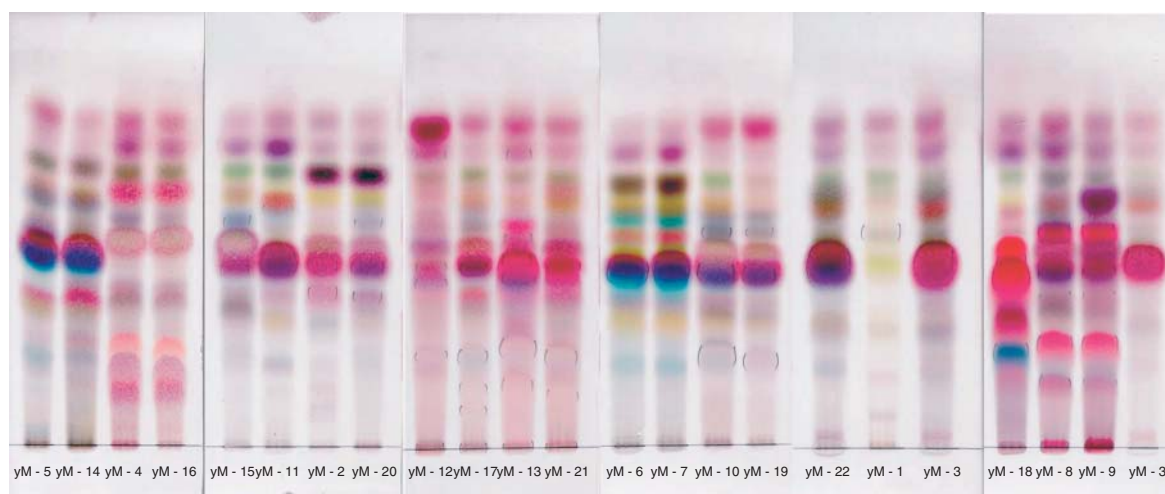
Figure 12 TLC of the ether extracts of several New Zealand liverworts.
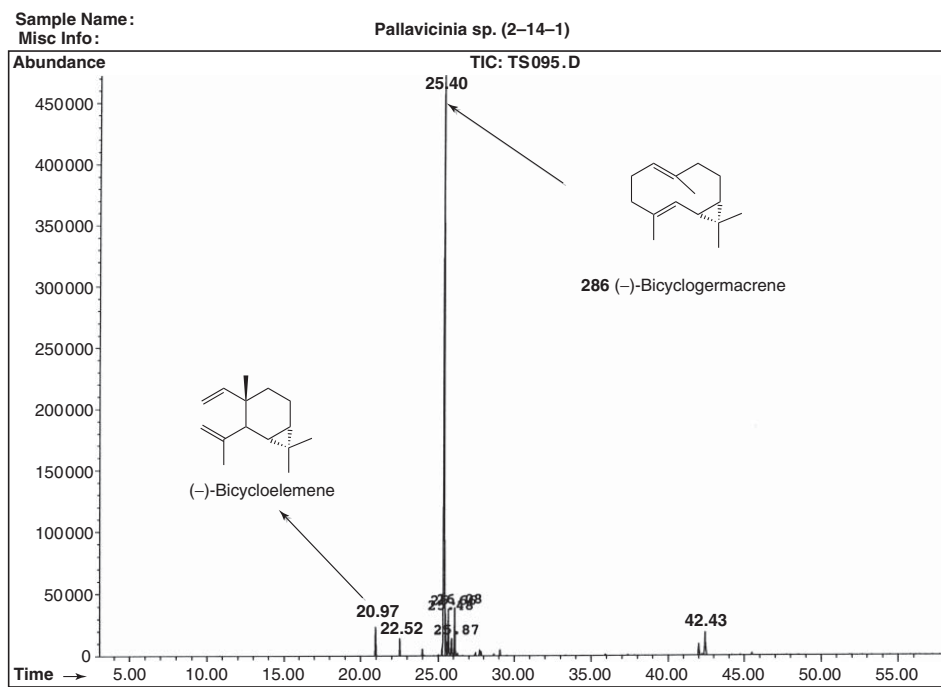
Figure 13   GC/MS of the essential oil from the liverwort *Pallavicinia subciliata*.
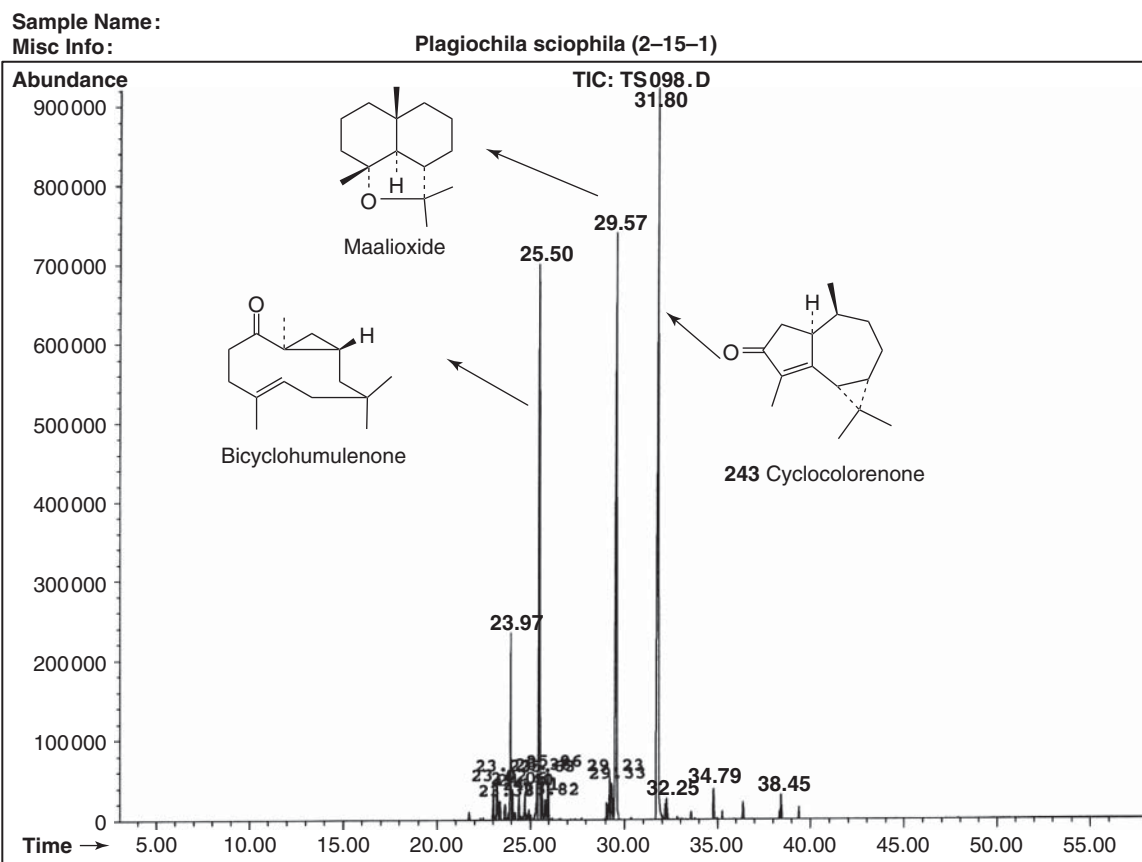
Figure 14 GC/MS of the essential oil from the liverwort *Plagiochila sciophila*.

used for chiral analysis of essential oils, with split injection; slit ratio: about 1 : 30; flame ionization detector (FID); carrier gas: 0.5-bar H_2; and injection and detector temperatures: 200 and 250°C, respectively. For preparative GC, a stainless steel column (1.85 m × 4.3 mm) with 10% polydimethyl-siloxane SE-30 on Chromosorb W-HP or with 2.5% octakis(2,6-di-*O*-methyl-3-*O*-pentyl)-γ-cyclodextrin in OV-1701 (50%, w/w) on Chromosorb W-HP has been used, with FID; He gas as carrier at a flow rate of 240 mL min⁻¹; and injection and detector temperatures of 200 and 250°C, respectively (Paul, Konig, and Wu, 2001).

GC–MS using an Agilent Technologies 6890N gas chromatograph coupled with a mass-selective detector (Agilent Technologies 5973), on an HP-5MS capillary (AgilentTechnologies 5973) and an HP-5MS capillary column (30 m × 0.25 μm, 0.25-mm film thickness), was used at the authors' laboratory. Oven temperature was 50°C with a 3 min initial hold, temperature programmed at 5°C min⁻¹, and 15°C min⁻¹ at 250°C. Injection temperature was 280°C. Carrier gas was He at 1 mL min⁻¹. The detector operated in an electron impact mode (70 eV with 3 scans per second and mass range *m*/*z* 40–500) at 30°C. The retention indices were calculated relative to C_8–C_{27} *n*-alkanes.

Compounds were identified using computer-supported spectral library (MassFinder 2.1 and 4.0, NIST 2008), mass spectra of reference compounds, and MS data from references (Joulain and Konig 1998) and authors' library databases, and then the identities were confirmed by comparison of their retention indices with those of reference compounds and published data.
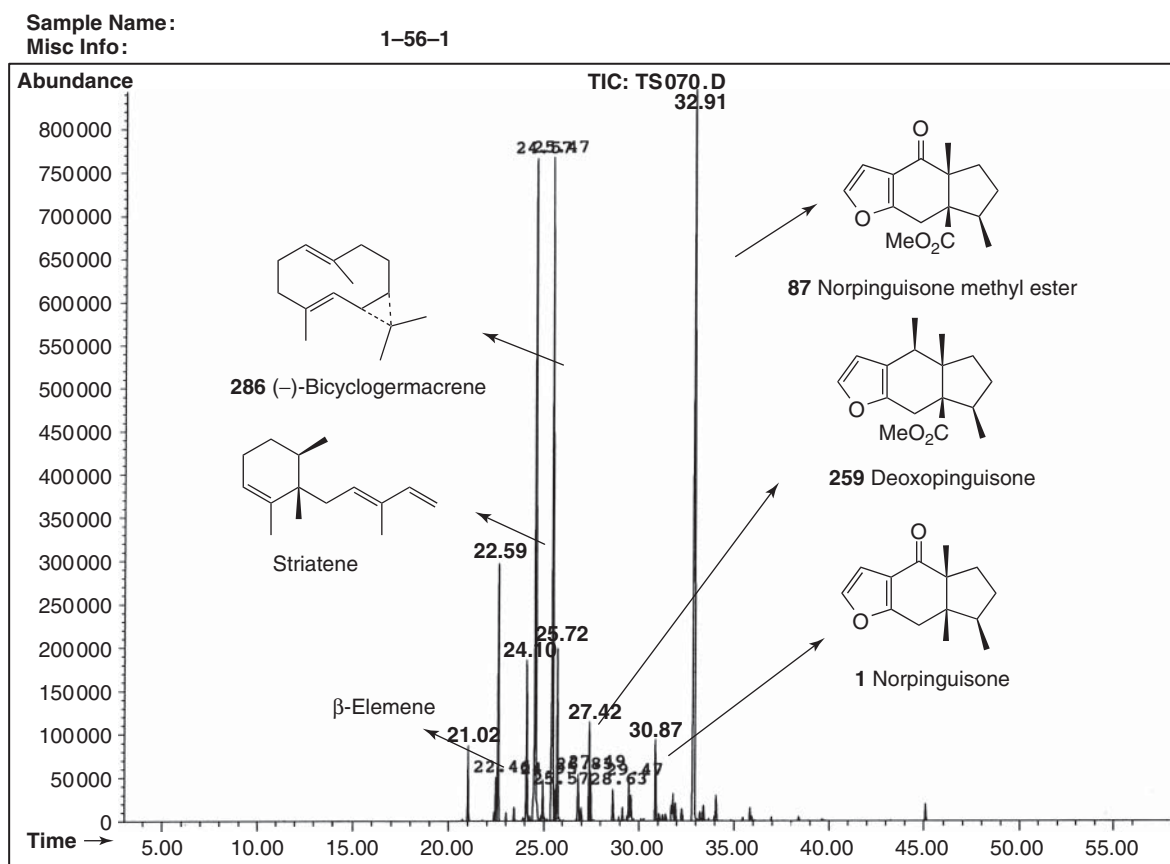
Figure 15   GC/MS of the essential oil from the liverwort *Porella japonica*.

4.6   Purification of the Crude Extracts by Column Chromatography, Preparative TLC, and HPLC

The remaining crude extracts from each liverwort sample were chromatographed on silica gel 60 (70–230 mesh) using *n*-hexane–ethyl acetate gradient and/or on Sephadex LH-20 (Pharmacia Fine Chemicals) using CH_2Cl_2–MeOH (1 : 1) as the solvent. Each fraction was further purified by preparative TLC using *n*-hexane–ethyl acetate as development solvent and preparative HPLC (high performance liquid chromatography) using either Cosmosil 5SL-II (10 × 250 mm), Cosmosil 5C18-AR (10 × 250 mm), or Nucleosil 50-5 using *n*-hexane–ethyl acetate (4 : 1). Reversed-phase silica gel (Cosmosil 75C_{18}) or CHEMCOROB 5-ODS-H column was used for the isolation of

lipophilic compounds using MeCN as development solvent (Nagashima *et al*., 2004a, b). Daiso SP-120-5-ODS-BP (20 mm i.d. × 250 mm) column and 0.1% HCOOH–10% MeCN/H_2O solvent system were used for the isolation of hydrophobic components (Tazaki *et al*., 1999).

4.7   Structural Elucidation

The stereochemical structures, including their absolute configuration of the isolated compounds, are elucidated by spectroscopic methods using IR (infrared), UV, ^{1}H NMR (600, 500, or 400 MHz), ^{13}C NMR (150, 125, or 100 MHz), and two dimension (2D) NMR spectroscopies, high resolution electron impact mass spectrometry (HREIMS), and/or X-ray crystallographic and CD (circular dichroism) spectral

analyses. 2D NMR experiments can be used to correlate both carbon and proton. Heteronuclear multiple quantum correlation (HSQC) experiment correlates the chemical shift of proton with that of the directly bonded carbon. On the bottom axis is a proton spectrum and on the other is a carbon spectrum. Cross peaks show the shift of the corresponding proton and carbon. This technique utilizes the one-bond coupling between carbon and proton. A more powerful method is heteronuclear multiple bond correlation (HMBC), which uses multiple-bond couplings over two or three bonds. Cross peaks are between protons and carbons that are two or three bonds away and often up to four or five bonds apart. Direct one-bond cross peaks resemble correlated spectroscopy (COSY), the proton–proton COSY experiment, in which it gives connectivity information over several bonds. Using these techniques, two examples for assignments of each proton and carbon of a new sesquiterpenoid, chandolide (**56**) (Komala *et al.*, 2009, 2010), have been demonstrated as shown in Figures 16–19.

4.8 Biosynthesis of Bis-Bibenzyls

Only a few biosynthetic papers on terpenoids and aromatic compounds from liverworts have been reported (Asakawa, Ludwiczuk, and Nagashima, 2013). Since the initial reports of two representative macrocyclic bis-bibenzyls, riccardin A (**30**) and marchantin A (**32**), isolated from Japanese *Riccardia multifida* (L.) A. Gray and *Marchantia polymorpha* L. (Figure 9) (Asakawa, 1982), more than 50 new bis-bibenzyls have been isolated from these liverworts and other species belonging to Jungermanniales, Marchantiales, Metzgeriales, and Monocleales and their structures elucidated (Asakawa, 1995). The occurrence, conformation, biosynthesis, biological activity, and synthesis of bis-bibenzyls have been reviewed by Asakawa's group (Asakawa, 1995; Asakawa *et al.*, 2000b, 2009) and Keseru and Nogradi (1995a, 1995b). Such macrocyclic bis-bibenzyls are very rare plant metabolites possessing structures that occur exclusively in Marchantiophyta. Such characteristic bis-
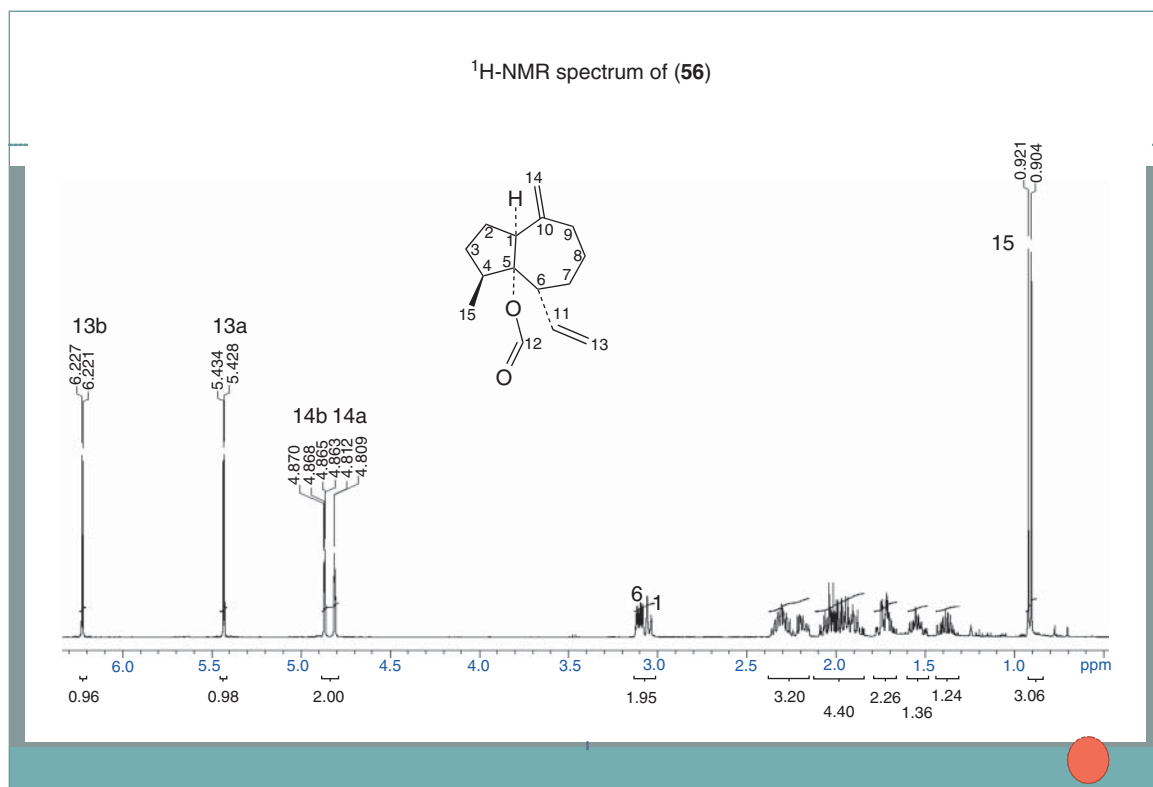
Figure 16 ¹H NMR spectrum of chandolide (**56**).
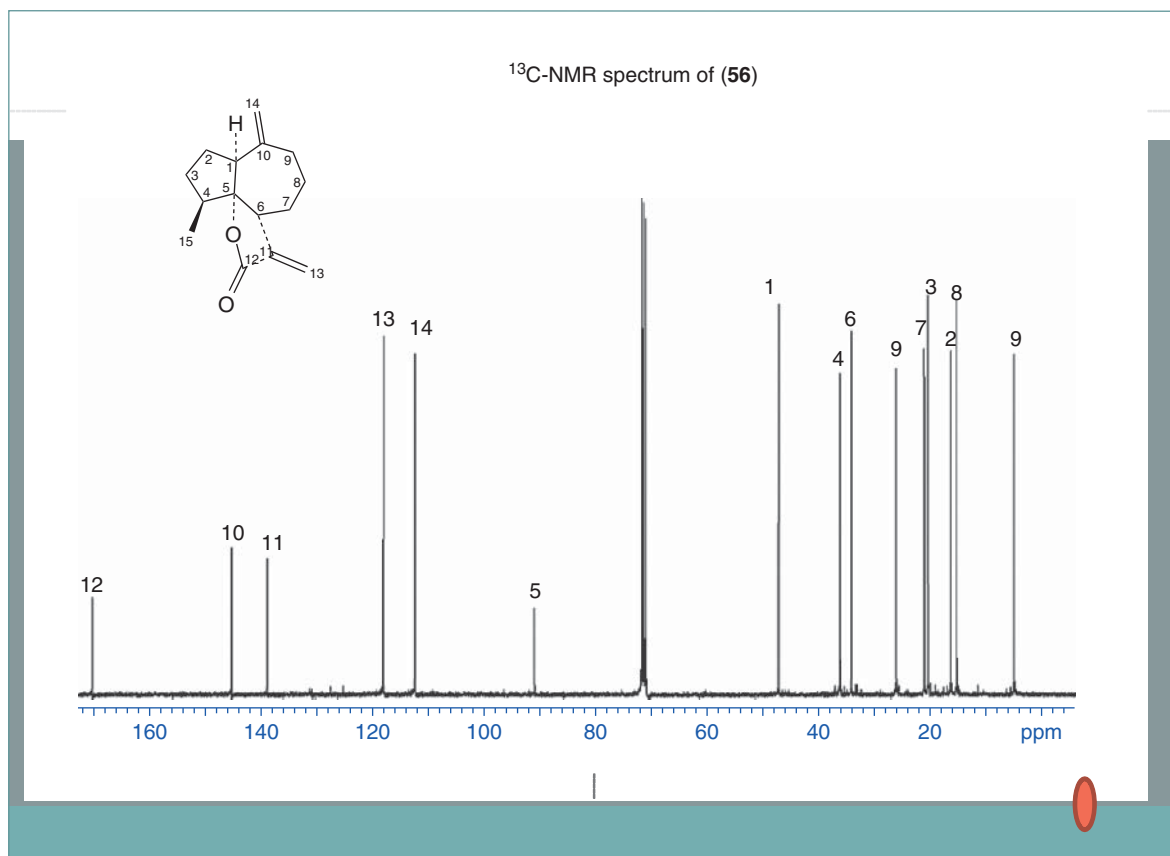
Figure 17 ^{13}C NMR spectrum of chandolide (**56**).

bibenzyls are not only very significant chemical markers of several Marchantiophyta families but also important for considering the phylogeny of the bryophytes and the evolutionary processes of the lower terrestrial spore-forming plants. The naturally occurring bis-bibenzyls are categorized into three structure types, which are made up of macrocyclic rings linked via two biphenyl ether C–O bonds, one biphenyl ether C–O, one biaryl C–C bond, and two biphenyl bonds.

Asakawa and Matsuda (1982) proposed that cyclic bis-bibenzyls, such as riccardin C (**31**) and marchantin A (**32**), might be biosynthesized from bibenzyls that correspond chemically to dihydrostilbenes. This assumption was proved by feeding experiments of radioactive and ^{13}C-labeled precursors, such as L-[U-^{14}C]phenylalanine, [U-^{14}C] dihydro-p-coumaric acid, [2-^{13}C]acetate, and L-[[13]COOH]phenylalanine, as shown in Figure 20

(Friederich *et al.*, 1999a, b). The aqueous precursor solutions were applied to 0.5-cm^2 samples of aseptic thallus tissue of *M. polymorpha* L. and incubated for 24 h. The A and B rings of the marchantin molecule are derived from the benzene ring of L-phenylalanine via *trans*-cinnamic acid and *p*-coumaric acid. Application of the ^{13}C-labeled precursor with subsequent ^{13}C NMR spectroscopy established that dihydrocoumaric acid is an intermediate in marchantin biosynthesis. Enzymatically hydrogenated dihydrocoumaric acid from coumaric acid condenses with three molecules of malonyl-Co A to form prelunularic acid (**58**). The latter is aromatized to yield lunularic acid (**59**) and possibly lunularin (**60**), which is followed by condensation of lunularin or lunularic acid to afford marchantin A (**32**). The mechanism of this final coupling step is still unknown. Two bis-bibenzyls coupled (**57a**, **57b**) with a sesquiterpene ether have been isolated from the New Zealand
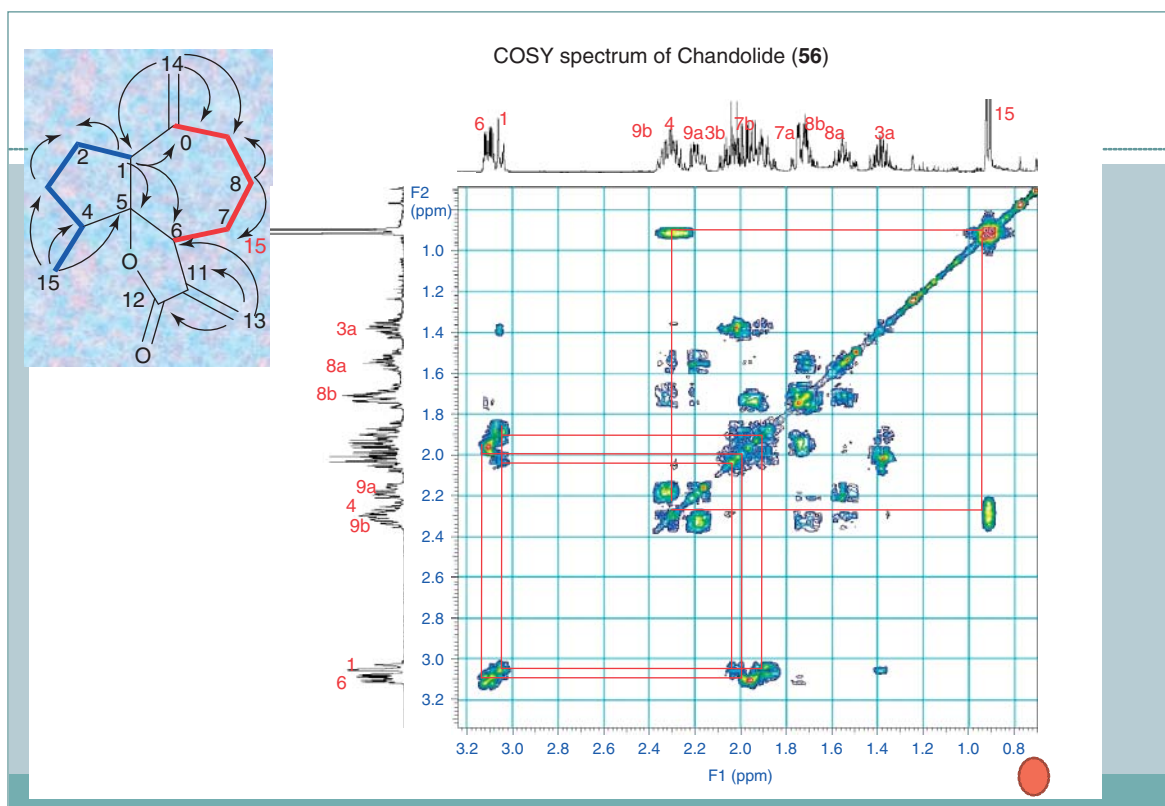
Figure 18 COSY NMR spectrum of chandolide (**56**).

liverwort, *Schistochila glaucescens* (Hook) A. Evans. Their biosynthetic pathway is quite interesting. As cyclic bis-bibenzyls found in liverworts possess unusual structures and various interesting biological activities, several organic chemists have focused on their total synthesis (Gottsegen *et al*., 1990; Harrowven, Woodcock, and Howes, 2005; Harrowven and Kostiuk, 2012; Hioki *et al*., 2009).

5 SCENTS AND BITTER SUBSTANCES FROM BRYOPHTES

5.1 Odorous Components from Liverworts

When bryophytes, especially liverworts, are crushed, they emit intense mushroom-like, sweet woody, or seaweed scents. The presence of 1-octen-3-ol (**39**) and its acetate (**40**) is responsible for the mushroom-like scent of a number of liverworts.

Generally, 1-octen-3-yl acetate is more abundant than the free alcohol. A Malaysian *Asterella* P. Beauv. species emits an incredibly unpleasant odor. It produces skatole (**53**), responsible for this smell, representing 20% of the total extract, and 80% of 3,4-dimethoxystyrene (Asakawa *et al*., 1995). The stink bug smell of the New Zealand *Chiloscyphus pallidus* (Mitt.) Engel & Schust. is attributable to (*E*)-(**41**), (*Z*)-pent-2-enal (**42**), (*E*)-dec-2-enal (**43**), and (*Z*)-dec-2-enal (**44**) (Toyota and Asakawa, 1994). *Cheilolejeunea imbricata* (Nees) S. Hatt. produces enantiomerically pure (*R*)-dodec-2-en-1,5-olide (**45**) and (*R*)-tetradec-2-en-1,5-olide (**46**), with a strong milky odor (Figure 10). A very tiny liverwort, *Leptolejeunea elliptica* (Lehm. &. Lindb.) Schiffn., grows on the leaves of *Camellia* L. or some ferns. It emits a powerful, sweet mold-like odor, which is due to the presence of a large amount of 4-ethylphenol (**61**), 4-ethylanisole (**62**), and 4-ethyl-1-acetoxybenzene (**63**) (Toyota, Koyama,
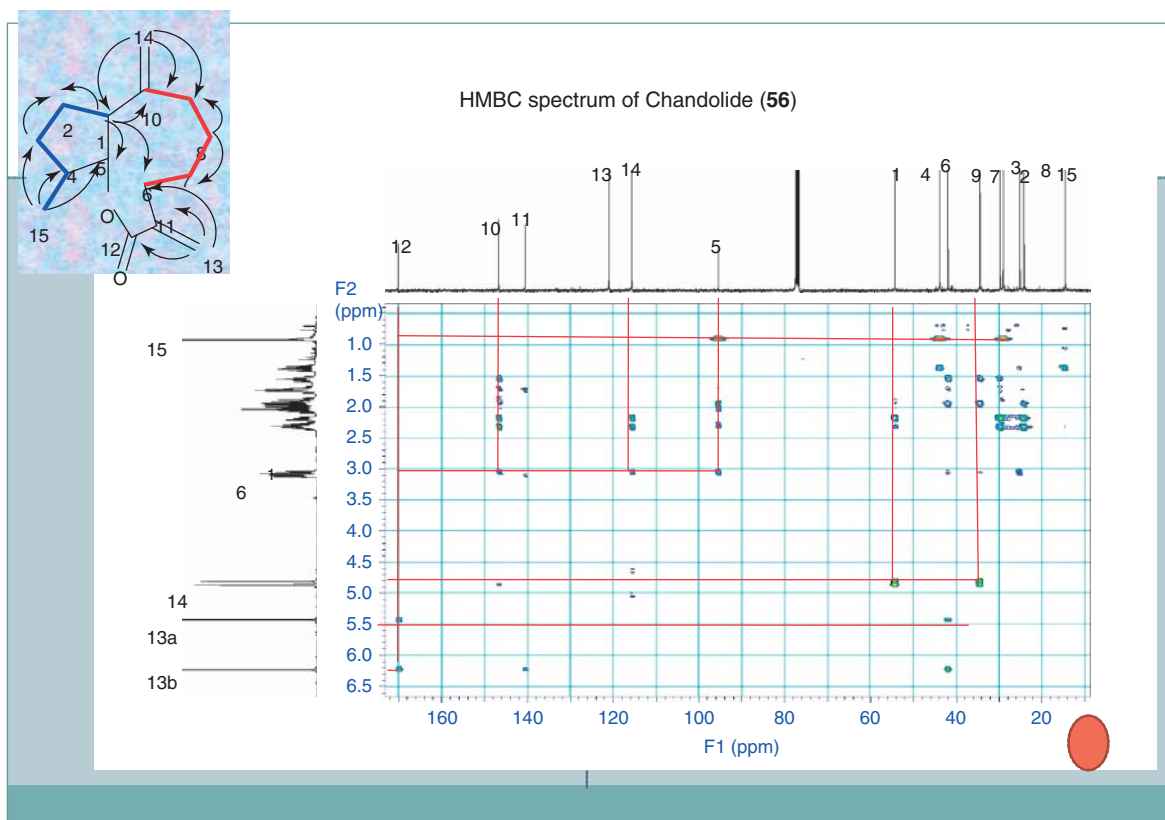
Figure 19 HMBC spectrum of chandolide (**56**).

and Asakawa, 1997a). Isoafricanol (**64**), isolated from the sporophytes of *Pellia epiphylla* (L.) Corda, is responsible for its typical odor (Cullmann and Becker, 1998) (Figure 21).

5.2 Pungent and Bitter Components

Some genera of liverworts contain intense pungent or bitter compounds, which exhibit interesting biological activities as described in the subsequent sections. *Porella arboris-vitae* (With) Grolle, *P. canariensis* (F. Weber) Underw., *P. fauriei* (Steph.) S. Hatt., *P. gracillima* Mitt., *P. obtusata* var. *macroloba* (Steph.) Hatt. et Zhang, *P. roellii* Steph., and *P. vernicosa* Lindb., which are classified as *P. vernicosa* complex, contain very hot tasting compounds. The pungency of *P. vernicosa* complex is due to (−)-polygodial (**65**).

Pellia endiviifolia (Dicks.) Dum. elaborates very hot taste, which is due to two diterpene dialdehydes, sacculatal (**26**) and 1β-hydroxysacculatal (**27**) (Hashimoto *et al*., 1995). The hot taste of *Pallavicinia lyellii* (Hook.) Carruth. (Asakawa, 1995) and *Riccardia lobata* Schiffn. var. *yakushimensis* Hatt. (Asakawa, Harrison, and Toyota, 1985) is due to sacculatal (**26**). *Hymenophyton flabellatum* (Labill.) Dum. ex Trevis. collected in New Zealand produces a pungent-tasting compound, which is due to the presence of 1-(2,4,6-trimethoxyphenyl)-but-(2*E*)-en-1-one (**66**) (Asakawa *et al*., 2001; Toyota *et al*., 2009).

Most of the species belonging to Lophoziaceae produce surprisingly intense bitter substances. *Gymnocolea inflata* (Hunds.) Dum. is persistently bitter and induces vomiting when one chews a few leaves for several seconds. An earlier review already mentioned that this is due to gymnocolin A (**14**) (Asakawa, 1995). It contains additional unknown minor bitter
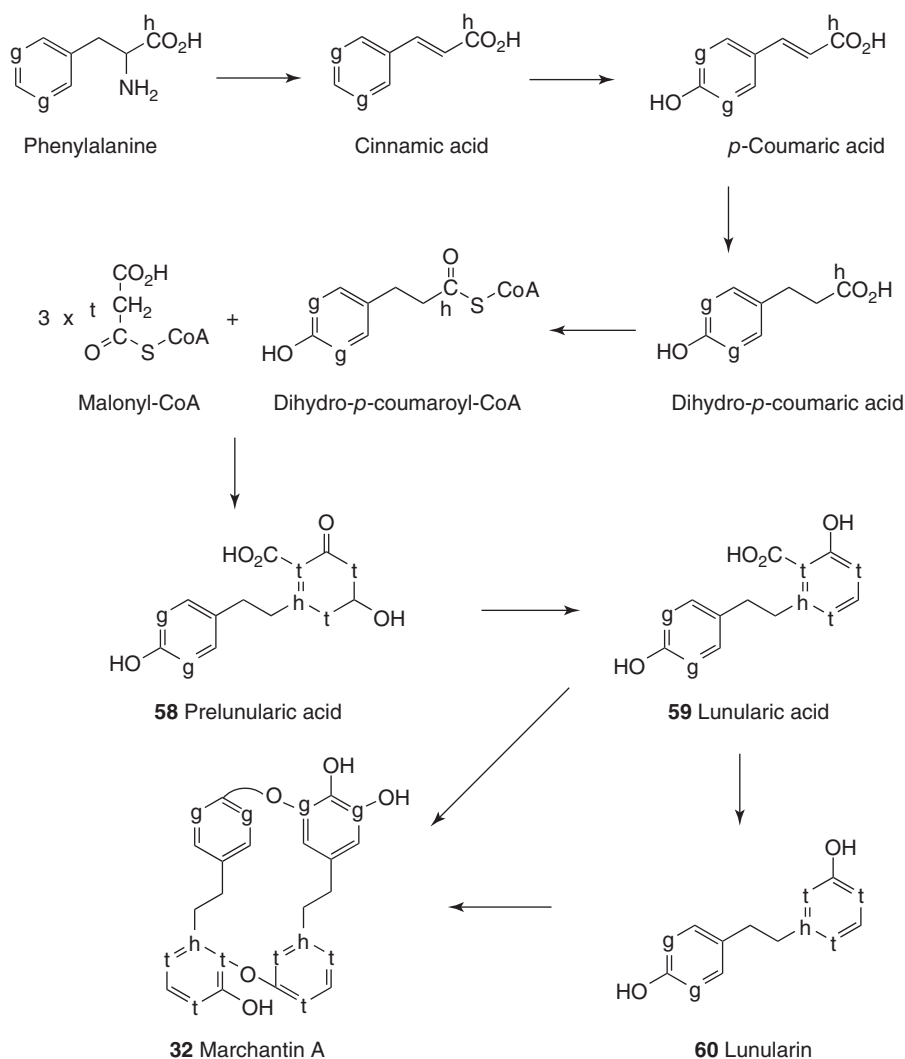
Figure 20 Biosynthetic roots of marchantin A (**32**).

diterpenoids whose structures remain to be clarified. *Jungermannia infusca* (Mitt.) Steph. has an intense bitter taste. This is due to the presence of the infuscasides, such as infuscaside A (**118**) and infuscaside B (**119**), which were the first reported isolations of glycosides from liverworts. The bitterness of *Anastrepta orcadensis* (Hook) Schiffn., *Barbilophozia lycopodioides* (Wallr.) Loeske, and *Scapania undulata* (L.) Dum. is attributable to highly oxygenated diterpenoids anastreptin A (**13**), barbilycopodin (**16**), and scapanin A (**22**), respectively (Asakawa, 1995).

6 BIOLOGICALLY ACTIVE COMPOUNDS FROM BRYOPHYTES

In China, several mosses and liverworts have been used as medicinal plants (Ding, 1982). *Bryum argenteum* Hedw. shows antidotal, antipyretic, and antirhinitic activities and is used for bacteriosis. *Ditrichum pallidum* (Hedw.) Hampe has been used for convulsions, particularly in infants. *Funaria hygrometrica* Hedw. has been used for hemostasis, pulmonary tuberculosis, vomitus cruentus, bruises, and athlete's foot dermatophytosis.

61 R = H 4-Ethylphenol
62 R = Me 4-Ethylanisole
63 R = Ac 4-Ethyl-1-acetoxybenzene

64 Isoafricanol

65 Polygodial

66 1-(2,4,6-Trimethoxyphenyl)but-(2*E*)-en-1-one

67 (–)-Frullanolide

68 (+)-Frullanolide

69 Plagiochiline A
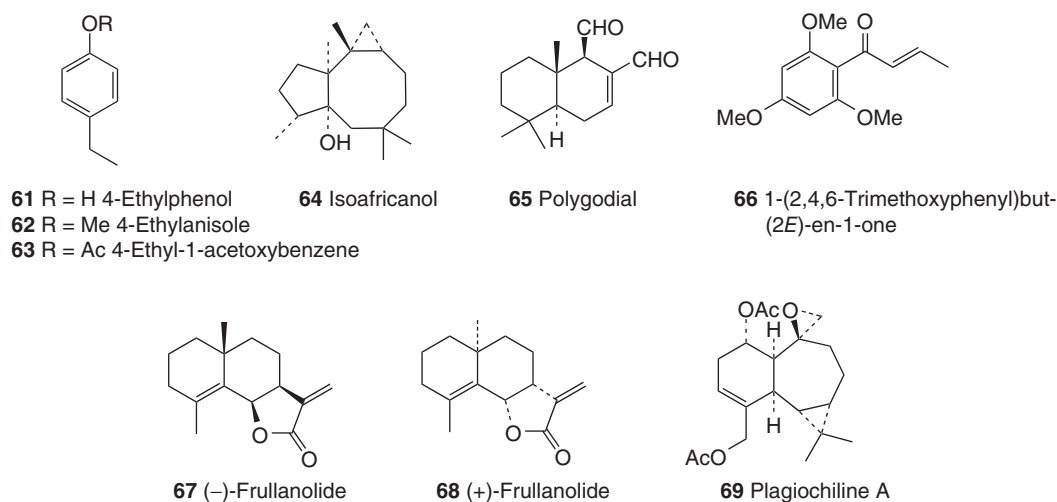
Figure 21 Odorous (**61–64**), pungent (65, 66), and allergy-inducing sesquiterpene lactones (**67**, **68**).

70 Marchantin C

71 R = H Isoplagiochin A
72 R = OH Isoplagiochin B

73 R = H Lepidozenolide
74 R = OOH 5α-Hydroperoxylepidozenolide

75 Marchantin B

76 R = H Marchantin D
77 R = Me Marchantin E

78 Perrottetin F
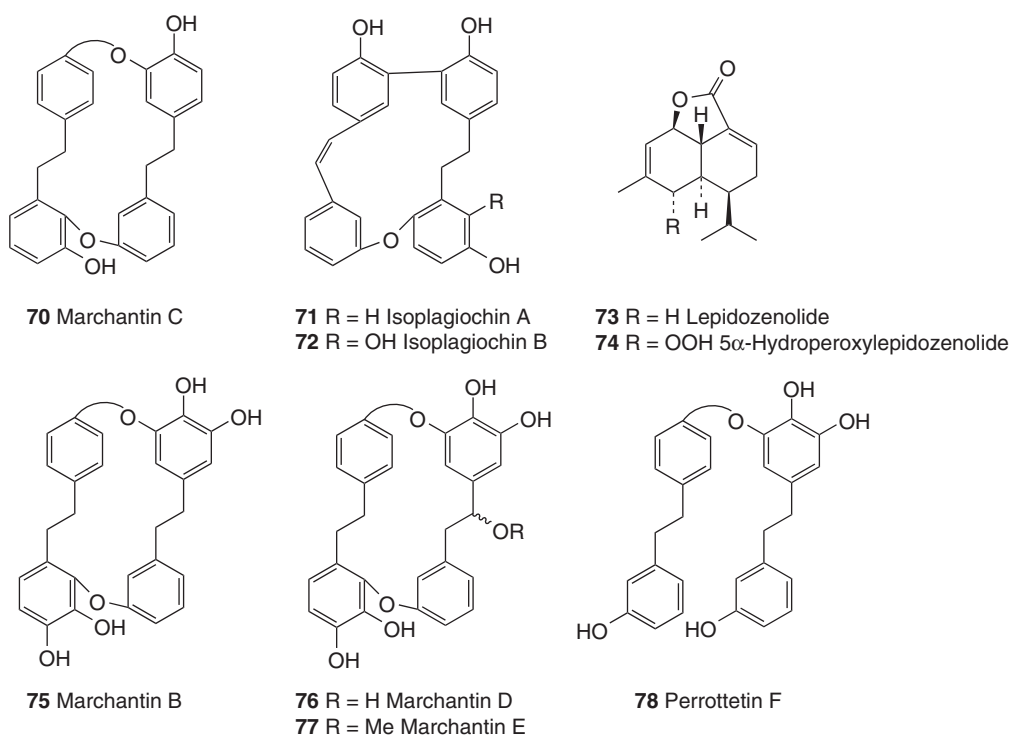
Figure 22 Biologically active compounds (**70–78**) from liverworts.

Polytrichum commune Hedw. shows antipyretic and antidotal activities and is used for hemostasis, cuts, bleeding from gingivae, hematemesis, and pulmonary tuberculosis. *Rhodobryum giganteum* (Schwaegr.) Par. also shows antipyretic, diuretic, and antihypertensive activities and is used for sedation, neurasthenia, psychosis, cuts, cardiopathy, and expansion of heart blood vessels (Asakawa, 2008a). In case of liverworts, four species, *M. polymorpha* L., *Conocephalum conicum* (L.) Dum., *F. tamarisci* (L.)

Dum., and *Reboulia hemisphaerica* (L.) Raddi, have been used as medicinal plants. *C. conicum* (L.) Dum. indicates antimicrobial, antifungal, antipyretic, and antidotal activities and is used to cure cuts, burns, scalds, fractures, swollen tissue, poisonous snake bites, and gallstones.

F. tamarisci (L.) Dum. has antiseptic activity and causes very strong contact dermatitis due to the presence of a high amount of eudesmane-type sesquiterpene lactone, (−)-frullanolide (**67**). On the other hand, *F. dilatata* (L.) Dum. elaborates its enatiomer, (+)-frullanolide (**68**), which also causes powerful allergenic contact dermatitis (Asakawa *et al*., 1976; Asakawa, 1982, 1995) (Figure 21).

One of the most popular thalloid liverworts, *M. polymorpha* L., indicates a similar activity as those mentioned for *C. conicum* (L.) Dum., antipyretic, antihepatic, antidotal, and diuretic activities, and is used to cure cuts, fractures, poisonous snake bites, burns, scalds, and open wounds. *R. hemisphaerica* (L.) Raddi has been used for blotches, hemostasis, external wounds, and bruises. However, the active compounds found in bryophytes have not been studied until recent decade.

Many liverworts contain different biologically active compounds that show allergenic contact dermatitis as mentioned earlier, tublin polymerization inhibitory, vasorelaxation, sex pheromone, plant growth regulatory, nitric oxide production inhibitory, brine shrimp lethality, calcium inhibitory, antithrombin, cathepsins B and L inhibitory, vasopressin antagonist activity, neurotrophic, α-glucosidase inhibitory, antimicrobial, antifungal, antiviral, insect antifeedant, antioxidant, antiplatelet, cytotoxicity and apoptotic, insecticidal, liver X-receptor alpha (LXRα) agonist receptor, muscle-relaxing, farnesoid X-receptor (FXR) activation (Asakawa, 1982, 1995, 2008a, 2008b, Asakawa, Ludwiczuk, and Nagashima, 2013), and antitripanosomal activities (Otoguro *et al*., 2011, 2012). The hot tasting compounds, such as sacculatals (**26**, **27**), polygodial (**65**), and plagiochiline A (**69**), show different biological activities, such as piscicidal, antimicrobial, antifungal, and insect antifeedant, and cytotoxicity, against human cancer cell lines (Toyota, Tanimura, and Asakawa, 1998; Yoshida *et al*. 1996). One of the most interesting biologically active compounds is macrocyclic bis-bibenzyls, such as marchantin, riccardin, and isoplagiochin series. Especially, marchantin A (**32**)-type compounds show anti-influenza, antimicrobial, antifungal, cytotoxic,

anticancer, muscle-relaxing, and cardiotonic activities (Asakawa, 1990a, 1990b; Huang *et al*., 2010; Iwai *et al*., 2011; Shi *et al*., 2009). Riccardin C (**31**) is also very important pharmacologically as it shows antiobesity activity (Tamehiro *et al*., 2005).

In Table 1, biological activity, bioactive compounds, and the names of the bioactive liverworts and mosses are shown. The chemical structures possessing biological activity are demonstrated in Figures 22–34.

7 CHEMOSYSTEMATICS

The identification of bryophytes is very difficult because they are morphologically very small. Thus, a microscope is necessary for the classification of bryophytes. The shape, size, and color of oil bodies in liverworts are also used for the classification of each species. A more important and precise method to classify bryophyte species, especially liverworts, is to use the chemical components. Several examples of chemosystematics of liverworts are presented here (Asakawa, 1982, 1995, 2004, Ludwiczuk *et al*., 2011, Asakawa, Ludwiczuk, and Nagashima, 2013).

7.1 Jungermanniales

7.1.1 Frullaniaceae and Porellaceae

The *Porella* L. and *Frullania* Raddi species are closely related morphologically. However, their chemical constituents are completely different; the former produce pungent substances, but the latter do not produce any pungent components. The *Porella* L. species are further divided into several chemical types (Table 2).

Type II comprises the pungent liverworts, *P. arboris-vitae* (With.) Grolle, *P. canariensis* (F. Weber) Underw., *P. fauriei* (Steph.) S. Hatt., *P. gracillima* Mitt., *P. obtusata* var. *macroloba* (Steph.) S. Hatt., *P. roellii* Steph., and *P. vernicosa* Lindb. (*P. vernicosa* complex). All of these liverworts contain drimane-type sesquiterpenoids, including the hot-tasting dialdehyde, polygodial (**65**) (Asakawa *et al*., 2000a). Liverworts belonging to type II accumulate characteristic sacculatane-type diterpenoids,

Table 1   Bioactive compounds isolated from liverworts and mosses (Asakawa 1982, 1995; Asakawa, Ludwiczuk, and Nagashima, 2013)

Biological activity	Bioactive compounds	Liverworts/ Mosses (M)
Piscicidal	Saccualtatals (**26, 27**)	*Pellia endiviifolia*
	Polygodial (**65**)	*Porella vernicosa*
	Plagiochiline A (**69**)	*Plagiochila tosana*
Cardiotonic activity	Marchantin A (**32**)	*Marchantia polymorpha*
Muscle-relaxant activity	Marchantin A (**32**)	*Marchantia polymorpha*
Calcium-inhibitory activity	Riccardin A (**30**)	*Riccardia multifida*
	Marchantin A (**32**)	*Marchantia polymorpha*
Farnesoid X-receptor (FXR) activation	Marchantin A (**32**)	*Marchantia polymorpha*
	Marchantin E (**77**)	*Marchantia polymorpha*
α-Glucosidase inhibitory activity	Marchantin C (**70**)	*Marchantia polymorpha*
Nitric oxide production inhibition	Norpinguisone (**1**)	*Porella densifolia*
	Marchantin A (**32**)	*Marchantia polymorpha*
	Riccardin A (**30**)	*Marchantia polymorpha*
	Riccardin C (**31**)	*Marchantia polymorpha*
	Riccardin F (**31a**)	*Marchantia polymorpha*
	Marchantin A (**32**)	*Marchantia polymorpha*
	Marchantin B (**75**)	*Marchantia polymorpha*
	Marchantin D (**76**)	*Marchantia polymorpha*
	Marchantin E (**77**)	*Marchantia polymorpha*
	Perrottetin F (**78**)	*Rodula perrottetii*
	α-Herbertenol (**79**)	*Mastigophora diclados*
	β-Herbertenol (**80**)	*Mastigophora diclados*
	Herbertene-1,2-diol (**81**)	*Mastigophora diclados*
	Mastigophorene C (**82**)	*Mastigophora diclados*
	Mastigophorene D (**83**)	*Mastigophora diclados*
	1,2-Cuparenediol (**84**)	*Bazzania decrescens*
	2-Hydroxy-4-methoxycuparene (**85**)	*Bazzania decrescens*
	Myltayl-4(12)-enyl-2-caffeate (**86**)	*Bazzania nitida*
	Norpinguisone (**87**) methyl ester	*Porella densifolia*
	ent-16-kauren-5-one (**88**)	*Porella densifolia*
	Thysaspathone (**89**)	*Thysananthus spathulistipus*
	3β,4β-15,16-Diepoxy-13(16),14-clerodane (**90**)	*Thysananthus spathulistipus*
	Radulanin C (**91**)	*Radula appressa*
	Radulanin L (**92**)	*Radula appressa*
	6-Hydroxy-4-(2-phenylethyl)benzofuran (**93**)	*Radula appressa*
	2-Gernayl-3,4-dihydroxybibenzyl (**94**)	*Radula appressa*
	(2*S*)-2-Methyl-2-(4-methy-3-pentenyl)-7-hydroxy-5-(2-phenylethyl)chromene (**95**)	*Radula appressa*
	O-Cannabicyclol (**96**)	*Radula appressa*
Sex pheromones	Dictyotene (**35**)	*Fossombronia angulosa*
	(*E*)-Ectocarpene (**36**)	*Fossombronia angulosa*
	Multifidene(**37**)	*Chandonanthus hirtellus*
	Dictyopterene (**38**)	*Chandonanthus hirtellus*
		Fossombronia angulosa
Plant growth inhibitory	Lunularic acid (**59**)	*Lunularia cruciata*
	Polygodial (**65**)	*Porella vernicosa*

Table 1 (*continued*)

Biological activity	Bioactive compounds	Liverworts/ Mosses (M)
Allergenic contact dermatitis	(−)-Frullanolide (**67**)	*Frullania tamarisci*
	(+)-Frullanolide (**68**)	*Frullania dilatata*
Tubulin polymerization inhibition	Marchantin C (**70**)	*Marchantia polymorpha*
	Isoplagiochin A (**71**)	*Plagiochila fruticosa*
	Isoplagoichin B (**72**)	*Plagiochila fruticosa*
Vasorelaxation	Lepidozenolide (**73**)	*Lepidozia fauriana*
	(-)-5β-hydroperoxylepidozenolide (**74**)	*Lepidozia fauriana*
Brine shrimp lethality activity	β-Phenylethyl benzoate (**97**)	*Balantiopsis cancellata*
	(*R*)-2-Hydroxy-2-phenylethyl benzoate (**98**)	*Balantiopsis cancellata*
	β-Phenylethyl (*Z*)-cinnamate (**99**)	*Balantiopsis cancellata*
	β-Phenylethyl (*E*)-cinnamate (**100**)	*Balantiopsis cancellata*
	Benzyl (*E*)-cinnamate (**101**)	*Riccardia polyclada*
	2,6-Dichloro-3-hydroxy-4′-methoxybibenzyl (**102**)	*Riccardia polyclada*
	2,6,3′-Trichloro-3-hydroxy-4′-methoxybibenzyl (**103**)	*Riccardia polyclada*
	2,4,6,3′-Tetrachloro-3-hydroxy-4′-methoxybibenzyl (**104**)	*Riccardia polyclada*
	2,4,6,3′-Tetrachloro-3,4′-dihydroxybibenzyl (**105**)	*Riccardia polyclada*
Antithrombin activity	Perrottetin E (**106**)	*Jungermannia* species
Antiplatelet activity	Lepidozenolide (**73**)	*Lepidozia fauriana*
		Lepidozia vitrea
	(−)-5β-Hydroperoxylepidozenolide (**74**)	*Lepidozia fauriana*
		Lepidozia vitrea
	Marchantiaquinone (**107**)	*Plagiochila elegans*
	Plagiochiline C (**108**)	*Reboulia hemisphaerica*
	Isoplagiochilide (**109**)	*Plagiochila elagans*
Cathepsin L and B inhibitory activity	Isomarchantin C (**110**)	*Bryopteris filicina*
	Infuscaic acid (**111**)	*Jungermannia infusca*
	11,13-Dehydoporelladiolide (**112**)	*Porella japonica*
Vasopressin (VP) antagonist activity	2-(3-Methyl-2-butenyl)-3,5-dihydroxybibenzyl (**113**)	*Radula perrottetii*
Neurotrophic activity	Plagiochilide (**4**)	*Plagiochila fruticosa*
	Mastigophorene D (**83**)	*Mastigophora diclados*
	Mastigophorene A (**114**)	*Mastigophora diclados*
	Mastigophorene B (**115**)	*Mastigophora diclados*
	Plagiochilal B (**116**)	*Plagiochila fruticosa*
	Plagiochin A (**117**)	*Plagiochila fruticosa*
	Infuscaside A (**118**)	*Jungermannia infusca*
	Infuscaside B (**119**)	*Jungermannia infusca*
Antibacterial Activity	2-Hydroxy-4,6-dimethoxyacetophenone (**120**)	*Plagiochila fasciculata*
(Gram positive and/or negative)	2-Hydroxy-3,4,6-trimethoxyacetophenone (**121**)	*Plagiochila fasciculata*
	2,4,6-Trichloro-3-hydroxybibenzyl (**123**)	*Riccardia marginata*
Antiviral activity	Marchantin A (**32**)	*Marchantia polymorpha*
((Influenza)	Marchantin B (**75**)	*Marchantia polymorpha*
	Perrottetin F (**78**)	*Marchantia polymorpha*
	Marchantin E (**77**)	*Marchantia polymorpha*
	Plagiochin A (**117**)	*Marchantia polymorpha*

(*Continued overleaf*)

Table 1 (*continued*)

Biological activity	Bioactive compounds	Liverworts/ Mosses (M)
Antiviral activity (HIV-1)	Marchantin A (**32**)	*Marchantia polymorpha*
	Marchantin B (**75**)	*Marchantia polymorpha*
	Marchantin D (**76**)	*Marchantia polymorpha*
	Perrottetin F (**78**)	*Radula perrottetii*
	Paleatin B (**122**)	*Marchantia paleacea* var. diptera
Antifungal activity (*Candida*)	Sacculatal (**26**)	*Pellia endiviifolia*
	Riccardin C (**31**)	*Plagiochasma intermedium*
	Riccardin F (**31a**)	*Plagiochasma intermedium*
	Marchantin A (**32**)	*Marchantia polygmorpha*
	Isotachin B (**55**)	*Balantiopsis cancellata*
	Lunularin (**60**)	*Dumortiera hirsuta*
	Polygodial (**65**)	*Porella vernicosa* complex
	Marchantin C (**70**)	*Schistochila glaucescens*
	Lepidozenolide (**73**)	*Lepidozia fauriana*
	Marchantin B (**75**)	*Marchantia polymorpha*
	Marchantin E (**77**)	*Marchantia polymorpha*
	α-Herbertenol (**79**)	*Mastigophora diclados*
	β-Herbertenol (**80**)	*Mastigophora diclados*
	Herbertene-1,2-diol (**81**)	*Mastigophora diclados*
	Mastigophorene C (**82**)	*Mastigophora diclados*
	(-)-Mastigophorene D (**83**)	*Mastigophora diclados*
	2-Phenylethyl benzoate (**97**)	*Balantiopsis cancellata*
	(*R*)-2-Hydroxy-2-phenylethyl benzoate (**98**)	*Balantiopsis cancellata*
	2-Phenylethyl (*Z*)-cinnamate (**99**)	*Balantiopsis cancellata*
	its (*E*)-cinnamate (**100**)	*Balantiopsis cancellata*
	2,6-Dichloro-3-hydroxy-4′-methoxybibenzyl (**102**)	*Riccardia polyclada*
	2,6,3′-Trichloro-3-hydroxy-4′-methoxybibenzyl (**103**)	*Riccardia polyclada*
	2,4,6,3′-Tetrachloro-3-hydroxybibenzyl (**104**)	*Riccardia polyclada*
	Perrottetin E (**106**)	*Asterella angusta*
	Mastigophorene A (**114**)	*Mastigophora diclados*
	2-Hydroxy-4,6-dimethoxy-acetophenone (**119**)	*Plagiochila fasciculata*
	2-Hydroxy-3,4,6-trimethoxy-acetophenone (**120**)	*Plagiochila fasciculata*
	2,4,6-Trichloro-3-hydroxy-bibenzyl(**123**)	*Riccardia marginata*
	2,4-Dichloro-3-hydroxy-bibenzyl (**124**)	*Riccardia marginata*
	2-Chloro-3-hydroxy-bibenzyl (**125**)	*Riccardia marginata*
	Mastigophoric methyl ester (**126**)	*Mastigophora diclados*
	α-Formyl herbertenol (**127**)	*Mastigophora diclados*
	1,2-Dihydroxyherberten-12-al (**128**)	*Mastigophora diclados*
	Viridiflorol (**129**)	*Bazzania trilobata*
	Gymnomitrol (**130**)	*Bazzania trilobata*
	5-Hydroxycalamenene (**131**)	*Bazzania trilobata*
	7-Hydroxycalamenene (**132**)	*Bazzania trilobata*
	Drimenol (**133**)	*Bazzania trilobata*
	Drimenal (**134**)	*Bazzania trilobata*

Table 1 (*continued*)

Biological activity	Bioactive compounds	Liverworts/ Mosses (M)
	Dehydrocostus lactone (**135**)	*Targionia lorbeeriana*
	Acetyltrifloculoside lactone (**136**)	*Targionia lorbeeriana*
	11α*H*-dihydrodehydrocostuslactone (**137**)	*Targionia lorbeeriana*
	Riccardiphenol C (**138**)	*Riccardia crassa*
	Glaucescenolide (**139**)	*Schistochila glaucescens*
	ent-1β-Hydroxykauran-12-one (**140**)	*Paraschistochila pinnatifolia*
	Riccardin D (= plagiochin E) (**141**)	*Marchantia polymorpha*
	13,13′-*O*-isopropylidene riccardin D (**142**)	*Marchantia polymorpha*
	Neomarchantin A (**143**)	*Schistochila glaucescens*
	Riccardin H (**144**)	*Marchantia polymorpha*
	Riccardin B (**145**)	*Asterella angusta*
	Marchantin H (**146**)	*Plagiochasma intermedium*
	Marchantin M (**147**)	*Plagiochasma intermedium*
	Marchantin P (**148**)	*Plagiochasma intermedium*
	Asterellin A (**149**)	*Asterella angusta*
	Asterellin B (**150**)	*Asterella angusta*
	11-*O*-Demethylmarchantin I (**151**)	*Asterella angusta*
	Dihydroptychantol A (**152**)	*Asterella angusta*
	Isoriccardin C (**153**)	*Asterella angusta*
	Pakyonol (**154**)	*Asterella angusta*
	6′,8′-Dichloroisoplagiochin C (= bazzanin B) (**155**)	*Bazzania trilobata*
	Isoplagiochin D (**156**)	*Bazzania trilobata*
	6′-Chloroisoplagiochuin D (= bazzanin S) (**157**)	*Bazzania trilobata*
	Tomentellin (**158**)	*Trichocolea tomentella*
	Demethoxytomentellin (**159**)	*Trichocolea tomentella*
	Artanorin (**160**)	*Homalia trichomanoides*
	3*a*-Methoxyserrat-14-en-21β-ol (**161**)	*Homalia trichomanoides*
	Neomarchantins B (**162**)	*Lepidozia fauriana*
	ent-Trachyloban-17-al (**163**)	*Jungermannia exsertifolia* subsp. *cordifolia*
	ent-3β-Acetoxy-19-hydroxy-trachylobane (**164**)	*Jungermannia exsertifolia* subsp. *cordifolia*
	ent-Trachylobane-3-one (**165**)	*Jungermannia exsertifolia* subsp. *cordifolia*
	ent-3β-Hydroxytrachylobane (**166**)	*Jungermannia exsertifolia* subsp. *cordifolia*
	ent-3β-Acetoxytrachylobane (**167**)	*Jungermannia exsertifolia* subsp. *cordifolia*

Table 1   (*continued*)

Biological activity	Bioactive compounds	Liverworts/ Mosses (M)
Insect antifeedant	Isotachin B (=β-phenethyl *trans*-β-methylthioacrylate) (**55**)	*Balantiopsis cancellata*
	1-(2,4,6-Trimethoxyphenyl)-but-2(*E*)-en-1-one (**66**)	*Hymenophyton flavellatum*
	Plagiochiline A (**69**)	*Plagiochila fruticosa*
	2,6,3′-Trichloro-3-hydroxy-4′-methoxybibenzyl (**103**)	*Riccardia polyclada*
	2,4,6,3′-Tetrachloro-3,4′-dihydroxybibenzyl (**105**)	*Riccardia polyclada*
	Clavigerin A (**168**)	*Lepodolaena clavigera*
	Clavigerin B (**169**)	*Lepodolaena clavigera*
	Clavigerin C (**170**)	*Lepodolaena clavigera*
	Methoxyclavigerin C (**171**)	*Lepodolaena clavigera*
Antioxidant activity	Marchantin A (**32**)	*Marchantia polymorpha*
	(−)-Herbertane-1,2-diol (**81**)	*Mastigophora diclados*
	(−)-Mastigophorene C (**82**)	*Mastigophora diclados*
	(−)-Mastigophorene D (**83**)	*Mastigophora diclados*
	3,5-Dihydroxy-2-(3-methyl-2-butenyl)-bibenzyl (**113**)	*Plagiochila fruticosa*
	Marchantin H (**147**)	*Marchantia polymorpha*
	Perrottetin D (**172**)	*Radula perrottetii*
	Plagiochin D (**173**)	*Plagiochila fruticosa*
	2,3,4,5,6-Pentamethoxyacetophenone (**174**)	*Adelanthus decipiens*
	Subulatin (**175**)	*Jungermannia sublata*
	α-Tocopherol (**176**)	*Porella, Pellia, Radula etc.*
Antitrypanosomal activity	Marchantin A (**32**)	*Marchantia polymorpha*
	2(*R*)-2-Isopropenyl-6,7-dihydroxy-4-(2-phenylethyl) dihydrobenzofuran (**93**)	*Radula perrottetii*
	Plagiochin A (**116**)	*Plagiochila fruticosa*
	α-Eudesmol (**177**)	*Porella perrottetinana*
Cytotoxic and Apoptotic activity	Sacculatal (**26**)	*Pellia endiviifolia*
	Perrottetin F (**31a**)	*Radula perrottetii*
	Marchantin A (**32**)	*Marchantia polymorpha*
	Chandolide(**56**)	*Chandonanthus hirtellus*
	Bis-bibenzyl dimer GBB A (**57a**)	*Schistochila glaucescens*
	Bis-bibenzyl dimer GBB B (**57b**)	*Schistochila glaucescens*
	Lunularin (**60**)	*Dumortiera hirsuta*
	Polygodial (**65**)	*Porella vernicosa*
	Plagiochiline A (**69**)	*Plagiochila ovalifolia*
	Marchantin C (**70**)	*Schistochila glaucescens*
	Lepidozenolide (**73**)	*Lepidozia fauriana*
	Marchantin B (**75**)	*Schistochila glaucescens*
	Marchantin D (**76**)	*Schistochila glaucescens*
	α-Herbertenol (**79**)	*Mastigophora diclados*
	(−)-Herbertene-1,2-diol (**81**)	*Mastigophora diclados*

Table 1 (*continued*)

Biological activity	Bioactive compounds	Liverworts/ Mosses (M)
	Mastigophorene C (**82**)	*Mastigophora diclados*
	Mastigophorene D (**83**)	*Mastigophora diclados*
	ent-16-Kauren-15-one (**88**)	*Jungermannia truncata*
	2-Hydroxy-4,6-dimethoxyacetophenone (**120**)	*Plagiochila fasciculata*
	2-Hydroxy-3,4,6-trimethoxyacetophenone (**121**)	*Plagiochila fasciculata*
	Paleatin B (**122**)	*Marchantia paleacea* var. *diptera*
	Riccardiphenol C (**138**)	*Riccardia crassa*
	Glaucescenolide (**139**)	*Schistochila glaucescens*
	ent-1*b*-Hydroxykauran-12-one (**140**)	*Paraschistochila pinnataifolia*
	Neomarchantin A (**143**)	*Schistochila glaucescens*
	Tomentellin (**158**)	*Trichocolea lanata*
		Trichocolea tomentella
	Demethoxytomentellin (**159**)	*Trichocolea tomentella*
	Atranorin (**160**)	*Frullania species*
	Neomarchantin B (**162**)	*Schistochila glaucescens*
	ent-Trachyloban-17-al (**163**)	*Jungermannia exsertifolia* ssp. c*ordifolia*
	ent-3β-Acetoxy-19-hydroxytrachylobane (**164**)	*Jungermannia exsertifolia* ssp. c*ordifolia*
	ent-Trachylobane-3-one (**165**)	*Jungermannia exsertifolia* ssp. c*ordifolia*
	ent-3β-Hydroxytrachylobane (**166**)	*Jungermannia exsertifolia* ssp. c*ordifolia*
	ent-3β-Acetoxytrachylobane (**167**)	*Jungermannia exsertifolia* ssp. c*ordifolia*
	2α,5β-Dihydroxybornane-2-cinnamate (**178**)	*Conocephalum conicum*
	Plagiochiline A-15-yl octanoate (**179**)	*Plagiochila ovalifolia*
	14-Hydroxyplagiochiline A-15-yl (2*E*,4*E*)-dodecadienoate (**180**)	*Plagiochila ovalifolia*
	Plagiochiline A-15-yl decanoate (**181**)	*Plagiochila ovalifolia*
	13,18,20-tri-*epi*-Chandonanthone (**182**)	*Chandonanthus hirtellus*
	Anadensin (**183**)	*Chandonanthus hirtellus*
	6α-Methoxyfusicoauritone (**184**)	*Chandonanthus hirtellus*
	13-Hydroxychiloscyphone (**185**)	*Chilosyphus rivularis*
	(−)-*ent*-Arbusculin B (**186**)	*Hepatostolonophora paucistipula*
	(−)-*ent*-Costunolide (**187**)	*Hepatostolonophora paucistipula*
	Costunolide (**188**)	*Frullania nisquallensis*
	Naviculyl caffeate (**189**)	*Bazzania novae-zelandiae*
	(−)-Diplophyllolide (**190**)	*Mastigophora diclados*
	1α-Hydroxy-*ent*-sandaracopimara-8(14),15-diene (**191**)	*Trichocolea mollissima*
	Rabdoumbrosanin (**192**)	*Lepidolaena taylorii*
	8,14-Epoxyrabdoumbrosanin (**193**)	*Lepidolaena taylorii*
	Atisane-2 derivative (**194**)	*Lepidolaena clavigera*
	α-Zeorin (**195**)	*Reboulia hemisphaerica*

(*Continued overleaf*)

Table 1   (*continued*)

Biological activity	Bioactive compounds	Liverworts/ Mosses (M)
	(+)-3α-(4′-Methoxybenzyl)-5,7-dimethoxyphthalide(**196**)	*Frullania* species
	(−)-3α-(3′-Methoxy-4′,5′-methylene-dioxybenzyl)-5,7-dimethoxyphthalide (**197**)	*Frullania* species
	3-Methoxy-3′,4′-methylene-dioxybibenzyl (**198**)	*Frullania* species
	2,3,5-Trimethoxy-9,10-dihydrophenanthrene (**199**)	*Frullania* species
	Lichexanthone (**200**)	*Frullania* species
	Tulipinolide (**201**)	*Frullania* species
	4α,5β-epoxy-8-*epi*-Inunolide(**202**)	*Porella perrottetiana*
	7-keto-8-Carbomethoxypinguisenol (**203**)	*Porella perrottetiana*
	Acutifolone A (**204**)	*Porella perrottetiana*
	Perrottetianal A (**205**)	*Porella perrottetiana*
	Pusilatin B (**206**)	*Blasia pusilla*
	Pusilatin C (**207**)	*Blasia pusilla*
	Methyl-4-[(2*E*)-3,7-dimethyl-2,6-octadieyl]oxy]3-hydroxy-benzoate (**208**)	*Tricocolea hatcheri*
	ent-3α-Hydroxy-16-kauren-15-one (**209**)	*Jungermannia truncata*
	(16*R*)-*ent*-11α-Hydroxy-16-kauran-15-one (**210**)	*Jungermannia truncata*
	ent-11α-Hydroxy-16-kauren-15-one (**211**)	*Jungermannia infusca*
	ent-Kaurene-11α,15α-diol (**212**)	*Jungermannia truncata*
	ent-6b-Hydroxy-16-kauren-15-one (**213**)	*Unidentified Jungermannia*
	16,17-*Di*hydrorostronol F (**214**)	*Jungermannia truncata*
	ent-14 α,15α-diHydroxy-16-kaurene (**215**)	*Jungermannia truncata*
	ent-20-Acetoxy-11α-hydroxy-16-kauren-15-one (**216**)	*Jungermannia truncata*
	ent-11α-Acetoxy-7β,14α-dihydroxy-16-kauren-15-one (**217**)	*Jungermannia truncata*
	(16*R*)-*ent*-Acetoxy-3α-hydroxy-16-kauran-15-one (**218**)	*Jungermannia truncata*
	ent-1β-Hydroxy-9(11),16-kauradien-15-one (**219**)	Unidentified *Jungermannia*
	ent-9(11),16-Kauradien-12,15-dione (**220**)	Unidentified *Jungermannia*
	Jungermannenone A (**221**)	Unidentified *Jungermannia*
	Jungermannenone B(**222**)	Unidentified *Jungermannia*
	Jungermannenone C(**223**)	Unidentified *Jungermannia*
	Jungermannenone D (**224**)	Unidentified *Jungermannia*
	ent-12β-Hydroxykaurane (**225**)	*Paraschistochila pinnatifolia*
	Ursolic acid (**226**)	*Pilidium pulcherrimum (M)*
	Acetoxyursolic acid (**227**)	*Pilidium pulcherrimum (M)*

Table 1 (*continued*)

Biological activity	Bioactive compounds	Liverworts/ Mosses (M)
	2*a*,3*b*-Dihydroxyurs-12-en-28-oic acid (**228**)	*Ptilidium pulcherrimum* (M)
	Pallidisetin A (**229**)	*Ptilidium pulcherrimum* (M)
	Pallidisetin B (**230**)	*Ptilidium pulcherrimum* (M)
	1-*O*-Methylohioensin B (**231**)	*Ptilidium pulcherrimum* (M)
	1-*O*-Methyldihydroohioensin B (**232**)	*Ptilidium pulcherrimum* (M)
	1,14-di-*O*-Methyldihydroohioensin B (**233**)	*Ptilidium pulcherrimum* (M)
Insecticidal activity	Atisane 2 (**194**)	*Lepidolaena clavigera*
	Hodgsonox (**234**)	*Lepidolaena hodgsoniae*
	Metacalypogin (**235**)	*Metacarypogeia alternifolia*
Nematode larval motility	Polygodial (**65**)	*Porella vernicosa* complex
	3-Methoxy-4′-hydroxy- bibenzyl (**236**)	*Plagiochila stephensoniana*
	3,4-Dimethoxybibenzyl (**237**)	*Plagiochila stephensoniana* *Frullania falciloba*

[a] *Source*: (Asakawa 1982, 1995; Asakawa, Ludwiczuk, and Nagashima, 2013).

which have not yet been found in any other plant group. The most characteristic compound for this chemotype is perrottetianal A (**205**). There are a few *Porella* L. species that, besides sacculatanes, produce a remarkably large amount of pinguisane-type sesquiterpenoids, for example, **1–3**, **87**. These are *Porella platyphylla* (L.) Pfeiff., *Porella grandiloba*. Lindb., *Porella elegantula* (Mont.) E. A. Hodgs., *P. japonica* (Sande Lac.) Mitt., and *Porella acutifolia* ssp. *tosana* (Steph.) S. Hatt. The former three species are included in the sacculatane–pinguisane type (III). However, the latter two liverworts have been separated from this group because besides the compounds mentioned, they biosynthesize very characteristic guaiane- (**112**) and germacrane-type sesquiterpene lactones (**202**) (type IV) (Asakawa, Ludwiczuk, and Nagashima, 2013; Toyota, Ueda, and Asakawa, 1991). *Porella recurva* (Taylor) Kuhnem., *P. cordaeana* (Huebener) Moore, *P. navicularis* (Lehm. & Lindb.) Lindb., as well as *P. densifolia* ssp. *appendiculata* (Steph.) S. Hatt., and *P. densifolia* var. *fallax* (C. Massal.) S. Hatt. (type V) produce pinguisanes as the main components. Pinguisane-type sesquiterpenoids are widespread in the genus *Porella*

L., but the species belonging to this group do not seem to produce many other sesquiterpenoids besides pinguisanes. *Porella* L. species that biosynthesize africane-type sesquiterpenoids (**238**, **239**) make up type VI. This class of compounds is very rare in nature, and these are characteristic for *Porella swartziana* (Weber) Trevis, *P. subobtusa* (Steph.) S. Hatt., and *P. caespitans* (Steph.) S. Hatt., var. *setigera* (Steph.) Matt. (Ludwiczuk and Asakawa, 2008; Asakawa, Ludwiczuk, and Nagashima, 2013). The taxonomy of the genus *Porella* L., based on morphology, has been regarded as notoriously difficult (Schuster, 1980). Recent DNA-based studies have brought new insights to the phylogeny and taxonomy of these plants (Hentschel *et al*., 2007). Comparison of the molecular classification with recognized *Porella* chemotypes shows some striking correlations (Table 3) (Ludwiczuk *et al*., 2011).

It appears that the distribution of chemotypes II, IV, and V over the molecular tree is scattered, but chemotypes I, III, and VI have more limited distributions and are phylogenetically informative. Thus, chemotypes III and VI occur only in species of lineage B, whereas the pungent chemotype I is
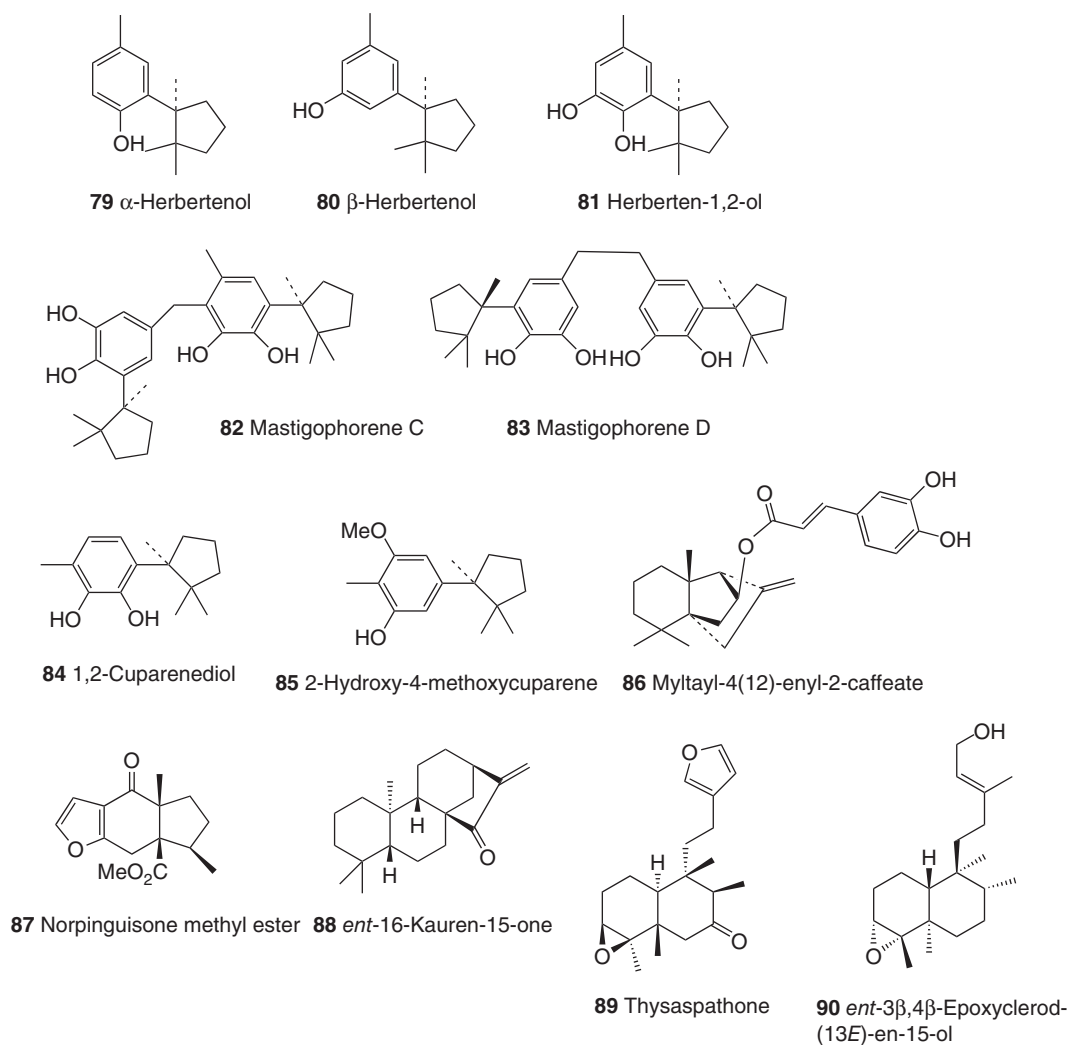
Figure 23 Biologically active compounds (**79–90**) from liverworts.

exclusive to members of lineage A. In fact, all the members of chemotype I (i.e., the *P. vernicosa* complex: *P. arboris-vitae* (With.) Grolle, *P. canariensis* (F. Weber) Underw., *P. obtusata* (Taylor) Trevis, *P. gracillima* Mitt., *P. roellii* Steph., *P. vernicosa* Lindb.), with the exception of *P. fauriei* (Steph.) S. Hatt., belong to the same molecular clade (A1). The chemical data suggest that the *P. vernicosa* complex deserves recognition as a separate section of *Porella*, *Porella* section *Vernicosae* Ludwiczuk, Gradstein, Nagashima & Asakawa, sect. nov. (Ludwiczuk *et al*., 2011).

With an estimated 300–375 species, *Frullania* Raddi is the largest genus of the order Porellales,

family Frullaniaceae, and forms a major clade of leafy liverworts (Hentschel *et al*., 2009). This genus is a rich source of sesquiter-penoids, especially sesquiterpene lactones, diterpenoids, and bibenzyl derivatives, such as 3-hydroxy-3′,4′-methylenedioxybibenzyl (**240**) and brittonin A (**241**) (Asakawa 2004; Asakawa, Ludwiczuk, and Nagashima, 2013). *Frullania* Raddi species are divided chemically into 7 major groups, namely, type I: sesquiterpene lactone-bibenzyl, type II: sesquitepene lactone, type III: bibenzyl, type IV: labdane, type V: bazzanane, type VI: pacifigorgiane-type, and type VII: cyclocolorenone (**242**), as shown in Table 4.

91 Radulanin C **92** Radulanin L **93** 6-Hydroxy-4-(2-phenylethyl)benzofuran

95 (2*S*)-2-Methyl-2-(4-methyl-3-pentenyl)-7-hydroxy-5-(2-phenylethyl)chromene

96 *o*-Canabicyclol

94 2-Geranyl-3,5-dihydroxybibenzyl

97 β-Phenethyl benzoate

98 (*R*)-2-Hydroxy-2-phenylethyl benzoate

99 β-Phenethyl (*Z*)-cinnamate

100 Benzyl-(*E*)-cinnamate

101 β-Phenethyl (*E*)-cinnamate
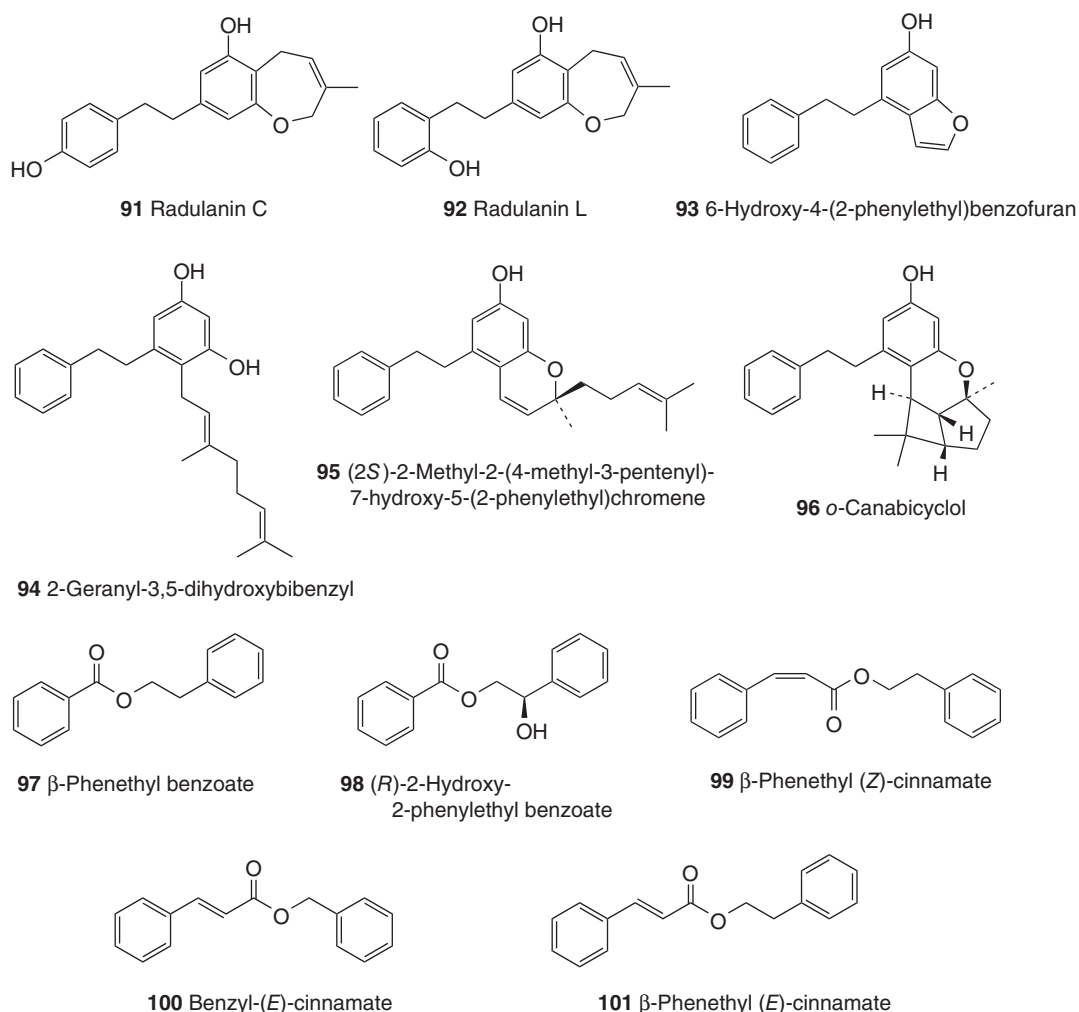
Figure 24 Biologically active compounds (**91–101**) from liverworts.

The most important chemical markers for these species are eudesmanolides (**67**, **68**) and eremophilanolide (**243**). Such compounds with an α-methylene-γ-lactone cause the miserable allergenic contact dermatitis (Asakawa *et al*., 2000a, 2000b). Besides the above-mentioned lactones, *Frullania* Raddi species also produce eremanolides, germacranolides, and guaianolides. There are some *Frullania* Raddi species that produce totally different secondary metabolites depending on the collection place. One of the examples is the Japanese *F. tamarisci* subsp. *obscura* (Verd.) S. Hatt. There are two district chemotypes of this species. The first, type T, contains tamariscol (**244**),

which shows remarkable aroma reminiscent of woody note and powdery green note of moss, and 5α,7β(*H*)-eudesma-4α,6α-diol (**245**) as the major component. The second chemotype, type O, lacks these two components and produces mainly a eudesmane-type sesquiterpene lactone, 4-*epi*-arbusculin A (**246**). Representatives of type T have been found in high mountains at altitudes of 1500–3000 m, whereas type O occurs more frequently at lower altitudes (Asakawa *et al*., 1991).

Despite the fact that liverworts belonging to Porellaceae and Frullaniaceae families are quite similar morphologically, they are very different from the

Table 2 Chemotypes of *Porella* species.

Types	*Porella* species	Sesquiterpenoids									Diterpenoids		
		DRI[a]	PIN	ARO	GER	GUA	AFR	MON	SAN	ELE	SAC	LAB	KAU
I	*Porella arboris-vitae*	++++	+	+	—	—	—	—	—	—	—	—	—
	Porella canariensis	+	+	+	—	—	—	—	—	—	—	—	—
	Porella fauriei	++	—	—	—	—	—	—	—	—	—	—	—
	Porella gracillima	+++	++++	++	—	—	—	—	—	—	—	—	—
	Porella obtusata ssp. *macroloba*	+++	++++	++	—	—	—	—	—	—	—	—	—
	Porella roellii	+++	—	—	—	—	—	—	—	—	—	—	—
	Porella vernicosa	+++	++++	+	—	—	—	—	—	—	—	—	—
II	*Porella camphylophylla*	—	—	—	—	—	—	—	—	—	+	—	—
	Porella perrottetiana	—	—	—	—	—	—	—	—	—	++	++	—
	Porella stephaniana	—	—	—	—	—	—	—	—	+	+	—	—
III	*Porella elegantula*	—	++	—	—	—	—	—	—	—	+	—	—
	Porella grandiloba	—	++++	—	—	—	+	—	—	+	+	—	—
	Porella navicularis	—	++++	—	—	—	—	+	+	+	+	+	—
	Porella platyphylla	—	++++	—	—	—	—	—	—	—	++++		
IV	*Porella acutifolia* ssp. *tosana*	—	++++	+	++++	++	—	—	—	+	+	—	—
	Porella japonica	—	++	++	++	++++	—	—	—	+	+	—	—
V	*Porella cordaeana*	++	++++	—	—	—	—	+	—	+	—	—	—
	Porella densifolia	—	+++	+	—	—	—	++	—	—	—	—	++++
	Porella recurva	—	++	—	—	—	—	—	—	—	—	—	—
VI	*Porella caespitans* ssp. *setigera*	—	—	+	—	—	++	—	+	+	—	—	—
	Porella subobtusa	—	—	—	—	—	+++	++	++	—	—	—	—
	Porella swartziana	—	—	+	++	++	++++	—	—	—	—	—	—

[a] DRI, drimanes; PIN, pinguisanes; ARO, aromadendranes; GER, germacranes; GUA, guaianes; AFR, africanes; MON, monocyclofarnesanes; SAN, santalanes; ELE, elemanes; LAB, labdanes; SAC, sacculatanes; KAU, kauranes.
Source: Ludwiczuk *et al.*, (2011)

chemical point of view, except for the presence of germacranolides. Such sesquiterpene lactones can be found in both the families and these are precursors of eudesmanolides, eremophilanolides, and guaianolides in both genera (Figure 35).

7.1.2 Plagiochilaceae

There are at least 1600 species of *Plagiochila* (Dum.) Dum., such as *Porella* L. species, which are divided mainly into two chemical groups, pungent and nonpungent species. The pungency of *Plagiochila* (Dum.) Dum. species is because of the presence of 2,3-secoaromadendrane-type sesquiterpene hemiacetals, such as plagiochiline A (**69**), which were degradated enzymatically to give two pungent components, furanoplagiochilal (**7**) and plagiochilal B

(**116**) (Hashimoto, Tanaka, and Asakawa, 1994). 2,3-Secoaromadendrane-type sesquiterpenoids are rare naturally occurring compounds and they are the most important chemical markers of *Plagiochila* (Dum.) Dum. genus. These compounds, however, are absent in the species belonging to the nonpungent chemotype. For both groups (pungent and nonpungent), several further chemical groups can be distinguished as shown in Table 5 (Asakawa, Ludwiczuk, and Nagashima, 2013). Among the *Plagiochila* (Dum.) Dum. species containing compounds of the plagiochiline series, there are liverworts that are characteristic to produce diterpenoids, for example, fusicoccanes (**19–21**), dolabellane (**17**), and cyathanes (**247**, **248**), as well as bibenzyls and bis-bibenzyls. *Plagiochila trabeculata* Steph., which belongs to type III, is quite isolated from the other *Plagiochila*

Table 3 Correlation between molecular classification and chemistry of *Porella* species.

Molecular classification based on maximum likelihood analysis (statistical support in brackets) (Hentschel *et al.*, 2007)	Chemotypes (Ludwiczuk *et al.*, 2011)
Clade A1 (69%)	
P. arboris-vitae	I
P. vernicosa	I
P. gracillima	I
P. obtusata	I
P. canariensis	I
P. roellii	I
Clade A2 (98%)	
P. densifolia	V
P. stephaniana	II
Clade A3 (<50%)	
P. japonica	IV
Clade A4 (<50%)	
P. fauriei	I
Clade B1 (97%)	
P. platyphylla	III
P. cordaeana	V
P. navicularis	III
Clade B2 (89%)	
P. acutifolia	IV
P. camphylophylla	II
P. perrottetiana	II
Clade B3 (99%)	
P. caespitans	VI
P. subobtusa	VI
Clade B4 (100%)	
P. swartziana	VI
Clade B5 (81%)	
P. grandiloba	III

Source: Ludwiczuk *et al.* (2011).

(Dum.) Dum. species, because this liverwort produces gymnomitrane-type sesquiterpenoids (**249**, **250**) as the main components (Asakawa, Ludwiczuk, and Nagashima, 2013). Nine chemotypes have been found in the nonpungent *Plagiochila* (Dum.) Dum. species. Most characteristic components of these liverworts are aromatic compounds, especially bibenzyls (**236**, **237**) and bis-bibenzyls (**115**, **177**, **173**), which constitute types IV, VI, VII, and VIII. Type XII of the *Plagiochila* (Dum.) Dum. species is characterized by the occurrence of everninic acid methyl ester (**251**), together with a wide range of other aromatic compounds, especially 9,10-dihydrophenanthrenes

(**252**, **253**). Types IX–XI produce mainly sesquiterpenoids, whereas type V produces diterpenoids. *Plagiochila rutilans* Lindb. is different from other species, as it contains various monoterpenoids that are responsible for a peppermint-like odor caused by the presence of several menthane monoterpenoids, including, most notably, pulegone (**254**), menthone (**255**), and isomenthone (**256**). All these monoterpenoids have been found in Bolivian and Brazilian collections, whereas a Costa Rican specimen produced 3,7-dimethyl-2,6-octadien-1,6-olide (**257**) as the principal monoterpenoid in place of pulegone (**254**) (Rycroft and Cole, 2001).

Until present, more than 90 *Plagiochila* (Dum.) Dum. species have been investigated chemically. This is only a small percentage of all species belonging to this genus, and because of this, the chemical classification of this genus is not complete.

7.1.3 Lejeuneaceae, Ptilidiaceae, and Trichocoleaceae

Species belonging to the former two families are very small morphologically. One of the important chemical markers of the three families are pinguisane sesquiterpenoids (**258**, **259**), which have been isolated from *Trichocoleopsis sacculata* (Mitt.) Okam. belonging to Trichocoleaceae, many *Porella* species belonging to Porellaceae as mentioned earlier, and also from *Ptilidium ciliare* (L.) Hampe and *Ptilidium pulcherrimum* (G. Web.) Hampe. All these families belong to Jungermanniales. Thus, Lejeuneaceae, Ptilidiaceae, Trichocoleaceae, and Porellaceae might have originated from the same ancestor. The presence of pinguisanes (**258**, **259**) has been found in *Aneura pinguis* (L.) Dum. (Aneuraceae) belonging to the order Metzgeriales. Thus, both the orders might have originated from the same ancestor (Asakawa *et al.*, 1995; Asakawa, Ludwiczuk, and Nagashima, 2013).

7.1.4 Radulaceae

There are about 150–200 species of *Radula* Dum. in the world. In Asia, 60 species have been known. The *Radula* is divided taxonomically into three groups, *Radula* Dum., *Cladoradula* Spruce, and *Odotoradula* K. Yamada, which is the only genus in Radulaceae family. Their chemical constituents are totally

Table 4 Chemotypes of *Frullania* species.

Type I(SL–BB)[a]	Type II(SL)	Type III(BB)	Type IV(LAB)	Type V(BAZ)	Type VI(PAC)	Type VII(CYCLOCOL)	Other types
F. brasiliensis	*F. apiculata*	*F. amplicrania*	*F. fugax*	*F. falciloba*	*F. fragilifolia*	*F. diversitexta*	*F. clavata*
F. convoluta	*F. asagrayana*	*F. anomala*	*F. hamatiloba*	*F. squarrosula*	*F. tamarisci*	*F. gaudichaudi*	*F. kagoshimensis*
F. davurica	*F. aterrima* var. *aterrima*	*F. bonincola*		unidentified	*F. tamarisci*		*F. motoyana*
F. davurica	*F. bicornistipula*	*F. ericoides*		*Frullania* sp.	ssp. *asagrayana*		*F. solanderiana*
F. dilatata	*F. brotheri*	*F. jackii*					*F. taradakensis*
F. dilatata var. anomala	*F. californica*	*F. patula*					*F. truncata*
F. incumbens	*F. chevalierii*	*F. pedicellata*					
F. muscicola	*F. congesta*	*F. pycnantha*					
F. osumiensis	*F. densiloba*	*F. scandens*					
F. parvistipula	*F. deplanata*	*F. spinifera*					
F. ptychantha	*F. inflata*						
F. serrata	*F. lobulata*						
F. tamarisci	*F. magellanica*						
ssp. *obscura*	*F. media*						
F. tamarisci	*F. monocera*						
ssp. *tamarisci*	*F. nepalensis*						
F. usamiensis	*F. nisqualensis*						
	F probosciphora						
	F. ramuligera						
	F. rostrata						
	F. sphaerocephala						
	F. tamarisci ssp. nisqualensis						
	F. ternatensis						
	F. vethii						
	F. yunnanensis						
	unidentified						
	Frullania sp.						

[a] SL–BB, sesquiterpene lactones-bibenzyls; SL, sesquiterpene lactones; BB, bibenzyls; LAB, labdanes; BAZ, bazzananes; PAC, pacifigorgianes; CYCLOCOL, cyclocolorenone.
Source: Asakawa, Ludwiczuk, and Nagashima (2013).

Table 5 The recognized chemotypes of the *Plagiochila* species.

Plagiochila species with the presence of 2,3-secoaromadendranes[a]

Type I (2,3-secoARO)	Type II (2,3-secoARO + DIT)	Type III (2,3-secoARO + GYM)
P. adiantoides	*P. aerea*	*P. trabeculata*
P. asplenioides	*P. cristata*	
P. atlantica	*P. moritziana*	
P. beskeana	*P. ovalifolia*	
P. carringtonii	*P. peculiaris*	
P. cipaconensis	*P. pulcherrima*	
P. cristatissima	*P. yokogurensis*	
P. cucullata		
P. dilatata		
P. dura		
P. elegans		
P. ericicola		
P. falcata		
P. fruticosa		
P. gayana		
P. goebeliana		
P. guayrapurinensis		
P. guilleminiana		
P. hattoriana		
P. hookeriana		
P. incurvicola		
P. lecheri		
P. magna		
P. micropterix		
P. orbicularis		
P. pittieri		
P. porelloides		
P. satoi		
P. scopulosa		
P. semidecurrens		
P. squamulosa var. sinuosa		
P. tenerrima		

***Plagiochila* species with the absence of 2,3-secoaromadendranes**

Type IV (DIT + BBB)	Type V (DIT)	Type VI (BB)	Type VII (BB + DIT)	Type VIII (BB + BBB)	Type IX (PIN)	Type X (CUP)	Type XI (SL)	Type XII (EVEA + DIPHE)
P. barteri	*P. circinalis*	*P. arbuscula*	*P. spinulosa*	*P. permista var. integerrima*	*P. alternans*	*P. bispinosa*	*P. hondurensis*	*P. bifaria*
P. sciophila	*P. corrugata*	*P. buchtiniana*	*P. stepchensoniana*	Unidentified I	*P. retro-spectans*	*P. duricaulis*		*P. killarniensis*
	P. deltoidea	*P. chacabucensis*			*P. rosariensis*	*P. subdura*		*P. retrorsa*
	P. dusenii	*P. diversifolia*						*P. stricta*
	P. geniculata	*P. exigua*						
	P. panamensis	*P. fasciculata*						
	P. validissima	*P. fuegiensis*						
	Unidentified II	*P. gymnocalycina*						
		P. longispina						
		P. maderensis						
		P. rutilans						
		P. standleyi						
		P. trichostoma						

(Continued overleaf)

Table 5 (*continued*)

Plagiochila species with the presence of 2,3-secoaromadendranes

Type I (2,3-secoARO)[a]		Type II(2,3-secoARO + DIT)	Type III(2,3-secoARO +GYM)
Other *Plagiochila* species			
P. acanthola	*P. elata*	*P. oresitropha*	*P. suborbiculata*
P. amazonica	*P. engelii*	*P. oxyphylla*	*P. tabinensis*
P. bursata	*P. excisa*	*P. pachyloma*	*P. tambillensis*
P. cornicula	*P. jamesoni*	*P. parvidens*	*P. terebrans*
P. dendroides	*P. kroneana*	*P. parvitexta*	*P. verruculosa*
P. dichotoma	*P. neesiana*	*P. pavidens*	

[a] *2,3-sec-ARO: 2,3-Secoaromadendranes; DIT: Diterpenoids; **GYM**: Gymnomitranes; **BBB**: Bis-bibenzyls; BB: Bibenzyls; PIN: Pinguisanes; **CUP**: Cuparanes; **SL**: Sesquiterpene lactones; **EVEA**: Eveminic acid methyl ester; **DIPHE**: 9,10-Dihydrophenanthre

Source: Asakawa, Ludwiczuk, and Nagashima (2013).

102 R^1 = R^2 = H, R^3 = OMe 2,6-Dichloro-3-hydroxy-4′-methoxybibenzyl
103 R^1 = H, R^2 = Cl, R^3 = OMe 2,6,3′-Trichloro-3-hydroxy-4′-methoxybibenzyl
104 R^1 = R^2= Cl, R^3 = OMe 2,4,6,3′-Tetrachloro-3-hydroxy-4′-methoxybibenzyl
105 R^1 = R^2 = Cl, R^3 = OH 2,4,6,3′-Tetrachloro-3,4′-dihydroxybibenzyl

106 Perrottetin E

107 Marchantiaquinone

108 Plagiochiline C

109 *iso*-Plagiochilide

110 Isomarchantin C

111 Infuscaic acid

112 11,13-Dehydroporelladiolide

113 2-(3-Methyl-2-butenyl)-3,5-dihydroxy-bibenzyl

114 Mastigophorene A

115 Mastigophorene B

Figure 25 Biologically active compounds (**102–115**) from liverworts.

different from the other genera of Marchantiophyta. Generally, they contain a large amount of bibenzyls, especially prenylated bibenzyls (**91–96**, **113**, **172**). The most significant chemical indicators for this genus are 2-geranyl-3,5-dihydroxybibenzyl (**94**) and 3,5-dihydroxy-2-(3-methyl-2-butenyl)bibenzyl (**113**). *Radula* Dum. species are also known for the presence of bibenzyls with a dihydro-oxepin skeleton (**91**, **92**), cyclopropane chroman derivatives (**260**), benzofuran (**93**), 2-phenylethylchromene (**261**), and

bibenzyl cannabinoids (**96**, **262**, **263**) (Asakawa, 2004) (Figure 36).

Bibenzyls possessing dihydro-oxepin skeleton are characteristic for *Radula javanica* Gott., *R. appressa* Mitt., *R. buccinifera* (Hook.f. et Taylor) Taylor, *R. complanata* (L.) Dum., *R. constricta* Steph., *R. grandis* Steph., *R. javanica* Gott., and *R. lindenbergiana* Gott. ex Hartm., and *Radula tokiensis* Steph., *R. appressa* Mitt., and *R. buccinifera* (Hook.f. et Taylor) Taylor also elaborate benzofuran together

116 Plagiochilal B

117 Plagiochin A

118 Infuscaside A

119 Infuscaside B

120 2-Hydroxy-4,6-dimethoxy-acetophenone

121 2-Hydroxy-3,4,6-trimethoxy-acetophenone

122 Paleatin B

123 R^1 = R^2 = Cl 2,4,6-Trichloro-3-hydroxybibenzyl
124 R^1 = Cl, R^2 = H 2,4-Dichloro-3-hydroxybibenzyl
125 R^1 = R^2 = H 2-Chloro-3-hydroxybibenzyl
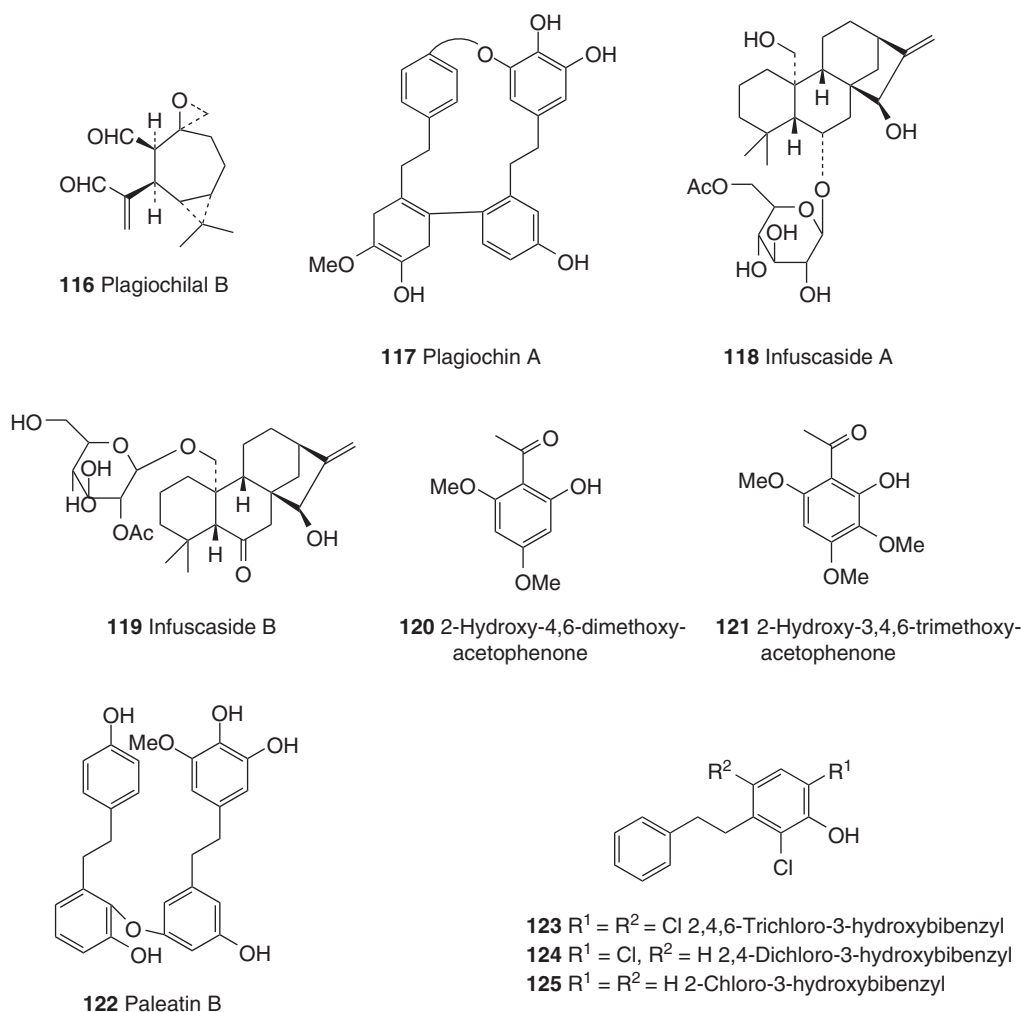
Figure 26   Biologically active compounds (**116–125**) from liverworts.

with dihydro-oxepin compounds. Very interesting compounds occurring in some *Radula* Dum. species are bibenzyl cannabinoids. All cannabinoids isolated from *Radula* Dum. thus far belong to three types: (i) the *o*-cannabichromene (**95**), (ii) *o*-cannabicyclol (**96**), and (iii) tetrahydrocannabinol types (**262**, **263**), with the *o*-cannabichromene type as the most prevalent. Such compounds have been found in *R. appressa* Mitt., *R. buccinifera* (Hook.f. et Taylor) Taylor, *R. campanigera* Mont., *R. javanica* Gott., *R. kojana*

Steph., *R. laxiramea* Steph., *R. lindenbergiana* Gott. ex Hartm., *R. obtusiloba* Steph., *R. perrottetii* Gott, ex Steph., and *R. tokiensis* Steph. (Asakawa, 1982, 1995; Asakawa, Ludwiczuk, and Nagashima, 2013).

7.1.5   Jackiellaceae

The genus *Jackiella* Schiffn. is the only member of the family Jackiellaceae. *Jackiella javanica*
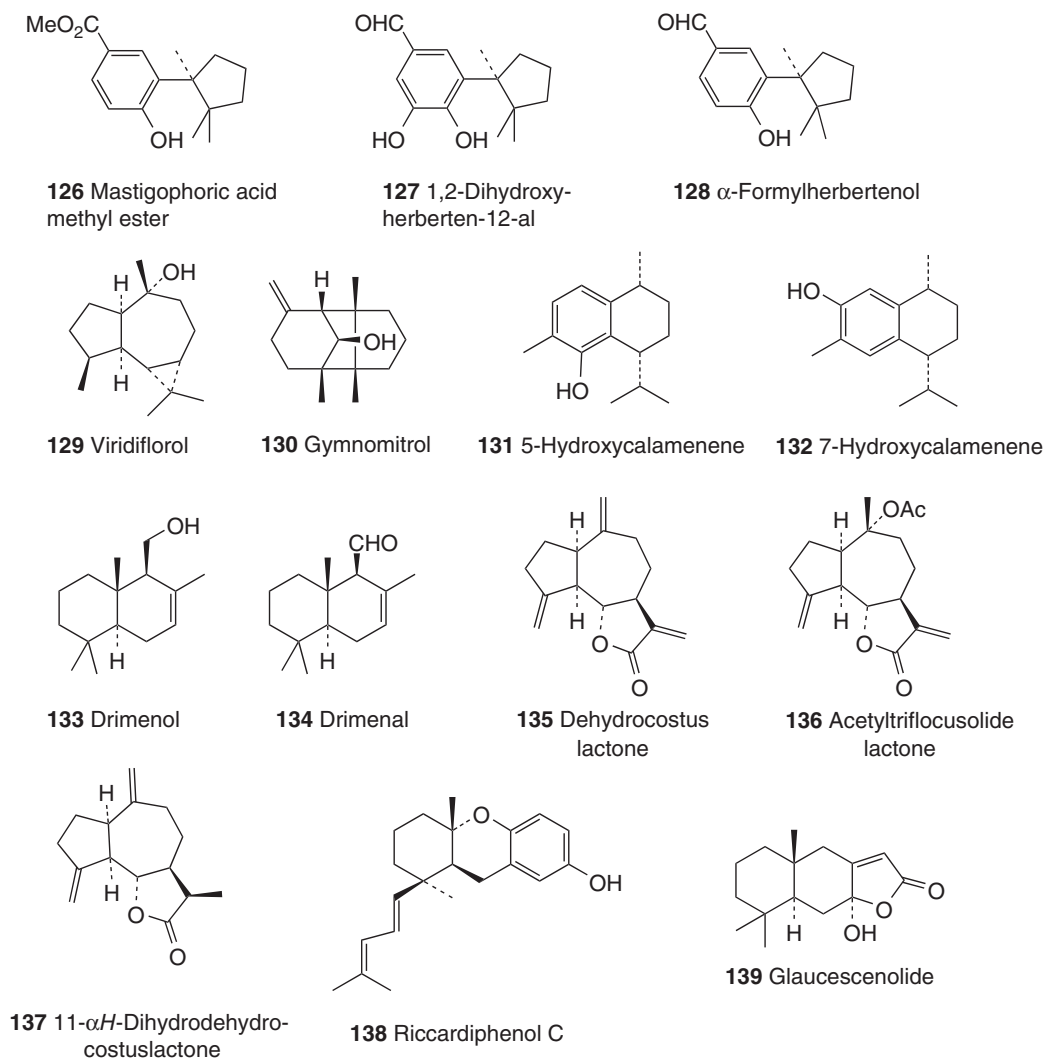
Figure 27 Biologically active compounds (**126–139**) from liverworts.

Schiffn. has been investigated chemically and is characterized by the presence of a wide range of *ent*-verticillane-type diterpenoids such as **23–25** (Nagashima *et al*., 1997, 2005, 2008). Verticillanes are very rare diterpenoids in nature. They are considered as putative biosynthetic precursors of the taxanes (Koepp *et al*., 1995). While these compounds are also related biogenetically to the cembranes (Karlsson *et al*., 1978; Basar, Koch, and König, 2001), until now, no cembranoids have been detected in *J. javanica* Schiffn. Cembranoids (**10–12**, **182**) together with *ent*-verticillol (**23**) and *ent*-12-epi-verticillol (**24**) have been found in Tahitian *Chandonanthus hirtellus* belonging to Scapaniaceae (Ludwiczuk *et al*., 2009; Komala *et al*., 2009, 2010).

140 *ent*-1β-Hydroxy-kauran-12-one

141 Riccardin D

142 13,13′-O-Isopropylidenericcardin D

143 Neomarchantin A

144 Riccardin H

145 Riccardin B

146 Marchantin H

147 Marchantin M

148 Marchantin P

149 R = H Asterelin A
150 R = Me Asterelin B

151 11-*O*-Demethylmarchantin I

152 Dihydroptychantol A
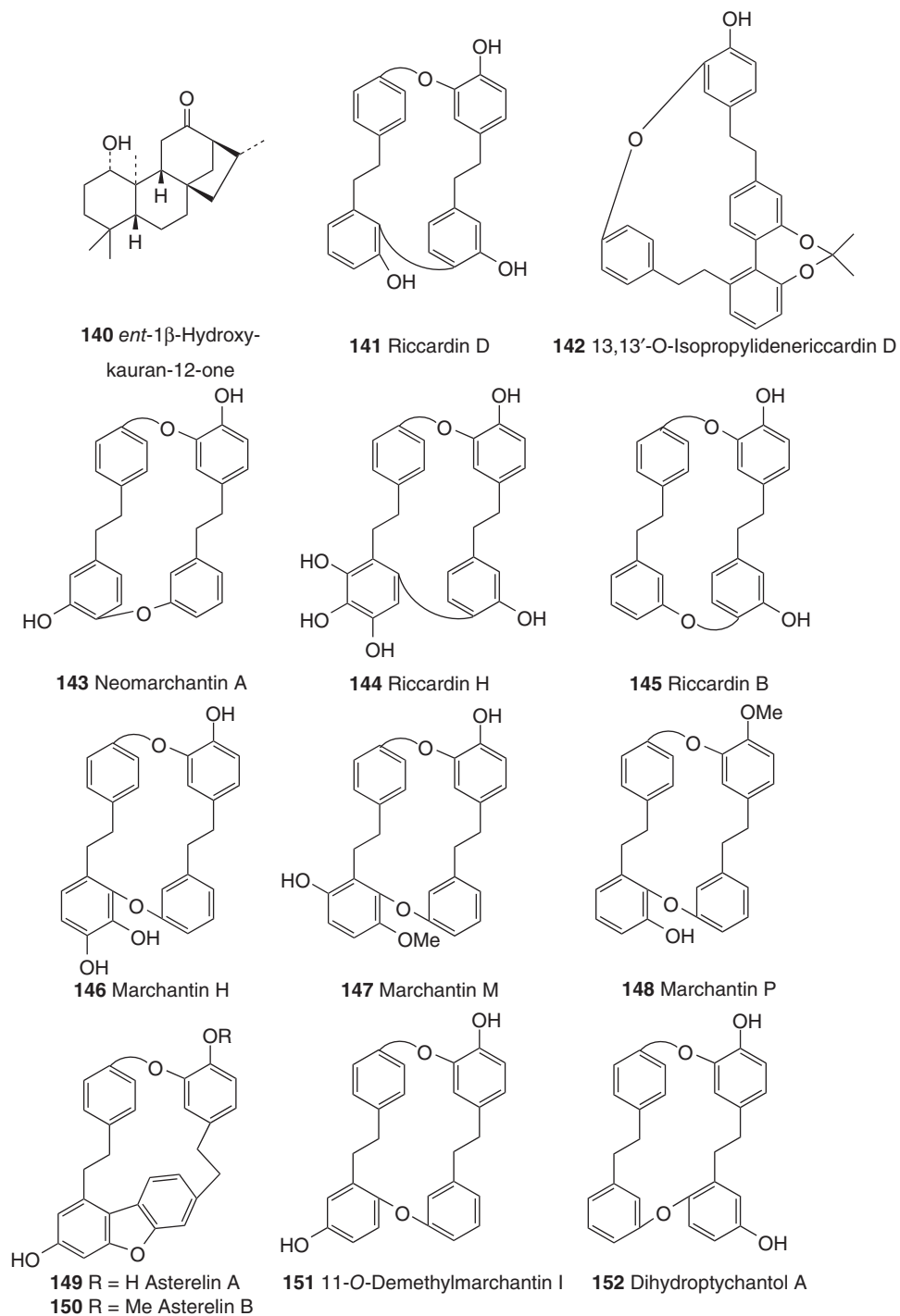
Figure 28   Biologically active compounds (**140–152**) from liverworts.
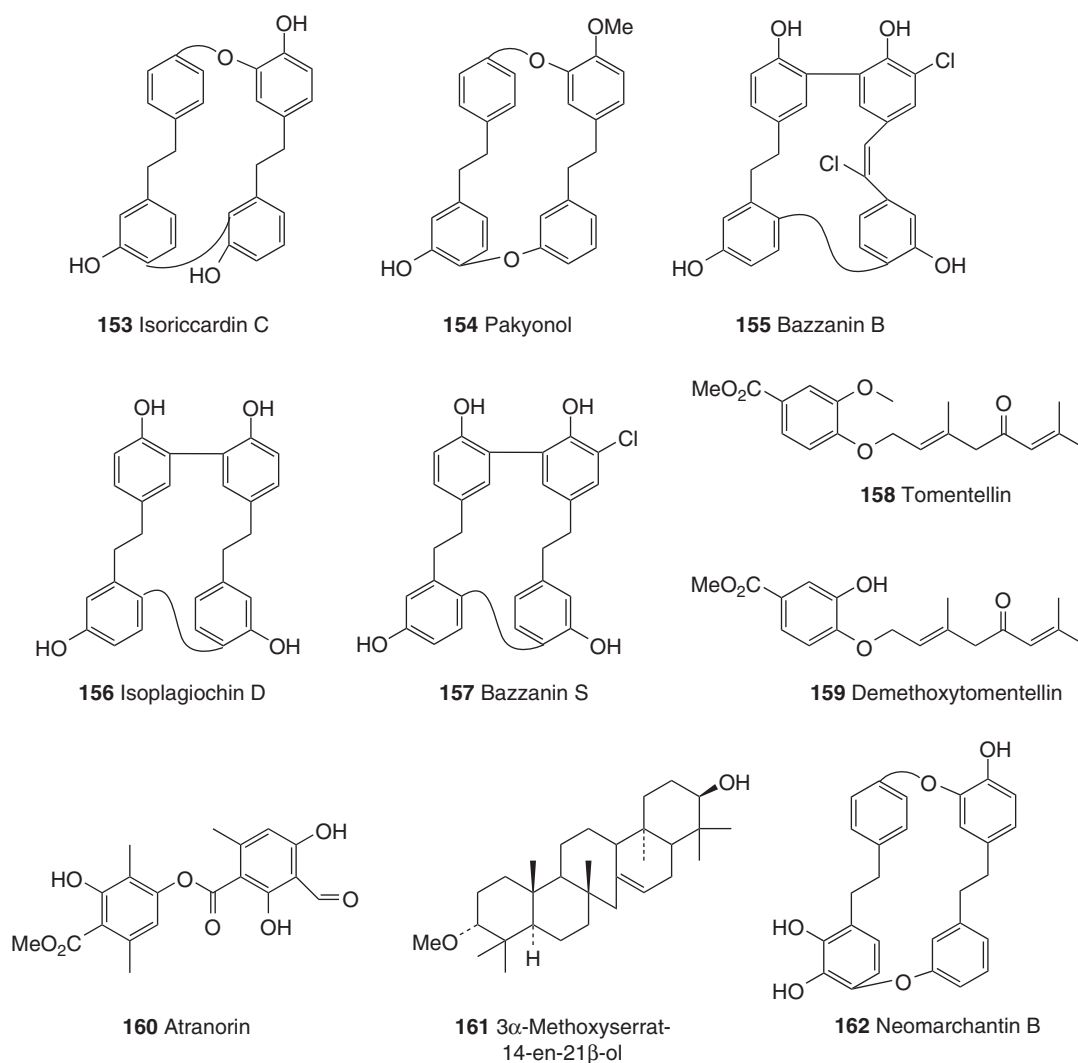
Figure 29 Biologically active compounds (**153–162**) from liverworts.

7.1.6 Calypogeiaceae

Some *Calypogeia* Raddi species, *C. azurea* Stotler & Crotz, *C. granulata* Inoue, and *C. muelleriana* (Schiffn.) Müll. Frib., have characteristic bluish-violet oil bodies, which indicate clearly the presence of azulenes in the live plants. 1,4-Dimethylazulene (**264**) and other azulenoids together with aromadendrane sesquiterpenoids have been found in these species (Asakawa, 1982, 1995; Tazaki *et al*., 1998; Warmers and König, 1998).

Azulenoids are very rare in nature. Thus, azulenoids in *Calypogeia* Raddi species are very characteristic and they seem to be significant chemical markers of this genus. This class of compounds has also been found in some *Barbilophozia* Loeske (Lophoziaceae), *Macrolejeunea* (Spruce) Schiffn. (Lejeuneaceae), and *Plagiochila* (Dum.) Dum. species (Plagiochilaceae) (Asakawa, 1995; Heinrichs *et al*., 2000; Adio and König, 2007). However, Calypogeiaceae and the other four families cannot be

163 R^1 = H, R^2 = R^3 = Me, R^4 = CHO *ent*-Trachyloban-17-al
164 R^1 = OAc, R^2 = CH$_2$OH, R^3 = R^4 = Me *ent*-3β-Acetoxy-19-hydroxytrachylobane
165 R^1 = O, R^2 = R^3 = R^4= Me *ent*-Trachyloban-3-one
166 R^1 = OH, R^2 = R^3 = R^4 = Me *ent*-3β-Hydroxytrachylobane
167 R^1 = OAc, R^2 = R^3 = R^4 = Me *ent*-3β-Acetoxytrachylobane

168 R^1 = OAc, R^2 = Ac Clavigerin A
169 R^1 = H, R^2 = Ac Clavigerin B

170 R = Ac clavigerin C
171 R = Me Methoxy clavigerin C

172 Perrottetin D

173 Plagiochin D

175 Subulatin

174 2,3,4,5,6-Pentamethoxyacetophenone

176 α-Tocopherol

177 α-Eudesmol

Figure 30 Biologically active compounds (**163–177**) from liverworts.

considered to be related morphologically. Aromadendranes, such as 1,2-dehydro-3-oxo-β-gurjunene (**265**), isolated from *C. azurea* Stotler & Crotz, may be related biogenetically to azulenes (Tazaki *et al*., 1998).

Metacalypogeia (S. Hatt.) Inoue species do not produce azulenoids but chroman derivatives. The Japanese *Metacalypogeia cordifolia* (Steph.) Inoue produces chroman compounds **266–270** (Toyota, Omatsu, and Asakawa, 2001) as major components.

178 2α,5β-Dihydroxybornane-2-cinnamate

179 R¹ = octanoyl, R² = H Plagiochiline A-15-yl octanoate
180 R¹ = decanoyl, R² = H Plagiochiline A-15-yl decanoate
181 R¹ = (2E),(4E)-dodecadienoyl, R² = OH
14-Hydroxyplagiochiline A-15-yl (2E),(4E)-dodecadienoate

182 13,18,20-epi-iso-Chandonanthone

183 Anadensin

184 6α-Methoxyfusocoauritone

185 13-Hydroxychiloscyphone

186 (−)-epi-Arbusculin B

187 (−)-ent-Costunolide

188 Ccostunolide

189 ent-Diplophyllolide

190 Naviculyl caffeate

191 1α-Hydroxy-ent sandaracopimara- 8(14),15-diene

192 Rabdoumbrosanin

193 8,14-Epoxy-rabdoumbrosanin

194 Atisane-type 2

Figure 31 Biologically active compounds (**178–194**) from liverworts.

Dihydroisochromene and metacalypogin (**235**) are the major constituents of the Taiwanese *Metacalypogeia alternifolia* (Nees) Grolle and *M. cordifolia* (Steph.) Inoue (Shy *et al.*, 2002). Thus, *Calypogeia* Raddi and *Metacalypogeia* (S. Hatt.) Inoue are very distinct chemically.

195 α-Zeorin

196 3α-(4'-Methoxybenzyl)-
5,7-dimethoxyphthalide

197 (−)-3α-(3'-Methoxy-4',5'-
methylenedioxybenzyl)-
5,7-dimethoxyphthalide

198 3-Methoxy-3',4'-methylenedioxy
bibenzyl

199 2,3,5-Trimethoxy-9,10-dihydro-
phenanthrene

200 Lichexanthone

201 Tulipinolide

202 4α,5β-Epoxy-**8**-*epi*-
innunolide

203 7-Keto-8-carbomethoxy-
pinguisenol

204 Acutifolone A

205 Perrottetianal

206 Pusilatin B

207 Pusilatin C

Figure 32   Biologically active compounds (**195–207**) from liverworts.

7.2   Metzgeriales

7.2.1   Aneuraceae, Pallaviciniaceae, Pelliaceae, Fossombroniaceae (Codoniaceae), and Makinoaceae

The German and Japanese *P. endiviifolia* (Dicks.) Dum. and European *P. epiphylla* (L.) Corda belonging to the Pelliaceae possess a persistent pungent taste that is due to the diterpene dialdehyde, sacculatal (**26**), that has been isolated from *T. sacculata*

(Mitt.) S. Okamura, classified in the order Jungermanniales as mentioned earlier. The same compound has also been isolated from *P. lyellii* (Hook.) Gray classified in Pallaviciniaceae. Thus, the three families mentioned earlier are chemically very closely related (Asakawa *et al.*, 2009). A similar sacculatane (**205**) has also been found in *Fossombronia* Raddi species belonging to Fossombroniaceae (Codoniaceae) and *Makinoa crispata* (Steph.) Miyake belonging to Makinoaceae. The latter species produces drimane sesquiterpenoids (**271**, **272**), which are one of the most important chemical markers of the *Porella*

208 Methyl 4-[[(2*E*)-3,7-dimethyl-
2,6-octadienyl]oxy]-3-hydroxybenzoate

209 *ent*-3α-Hydroxy-
16-kauren-15-one

210 16(*R*)-*ent*-11α-Hydroxy
kauran-15-one

211 *ent*-11α-Hydroxy-
16-kauren-15-one

212 *ent*-16-Kaurene-11α,15α-diol

213 *ent*-6β-hydroxy-
16-kauren-15-one

214 16,17-Dihydrorostronol F

215 *ent*-14α,15α-Dihydroxy-
16-kaurene

216 *ent*-20-Acetoxy-11α-
16-kauren-15-one

217 *ent*-11α-Acetoxy-7β,14α
dihydroxy-16-kauren-15-one

218 ((16*R*)-*ent*-3α-Hydroxy-
kauran-15-one

219 R^1 = OH, R^2 = H, H *ent*-1β-Hydroxy-9(11),
16-kauradien-15-one

220 R^1 = H, R^2 = O   *ent*-9(11),16-Kauradien-12,
15-dione

221 R^1 = OH, R^2 = R^4 = H, R^3 = O Jungermannenone A
222 R^1= R^2 = R^4 = H, R^3 = O Jungermannenone B
223 R^1 = R^2 = H, R^3 = O, R^4 = OH Jungermannenone C
224 R^1 = R^4 = OH, R^2 = H, R^3 = O Jungermannenone D

Figure 33   Biologically active compounds (**208–224**) from liverworts.

225 *ent*-12β-Hydroxy-
kauran

226 R¹ = R² = H Ursolic acid
227 R¹ = Ac, R² = H Acetoxy ursolic acid
228 R¹ = H, R² = OH 2α,3β-Dihydroxyurs-
12-en-28-oic acid

229 Pallidisetin A

230 Pallidisetin B

231 1-*O*-Methylohioensin B

232 R = H 1-*O*-Methyldihydroohioensin B
233 R = Me 1,14-Di-*O*-methyldihydro-
ohioensin B

234 Hodgsonox

235 Metacalypogin

236 R = H 3-Methoxy-4′-hydroxybibenzyl
237 R = Me 3,4′-Dimethoxybibenzyl

Figure 34 Biologically active compounds (**225–237**) from liverworts and mosses.

species belonging to Jungermanniales. *Fossombronia angulosa* (Dicks.) Raddi also produces cyathane-type diterpenoids, (−)2β,9α-dihydroxyverrucosane (**273**) and 5,18-dihydroxy-*epi*-homo-verrucosane (**274**), which are distributed in some Jungermanniales. Thus, the presence of sacculatane-type diterpenoids and cyatane diterpenoids is a further evidence that both Jungermanniales and Metzgeriales originated from the same ancestor, although they are not closely related morphologically to each other.

In the modern classification of Marchantiophyta, Jungermanniales and Metzgeriales are united within the subclass Jungermanniidae (Asakawa *et al*., 2009). The chemical evidence mentioned earlier shows that

several species belonging to both Jungermanniales and Metzgeriales produce very specifically similar terpenoids and aromatic compounds.

7.3   Marchantiales

7.3.1   *Aytoniaceae*

R. hemisphaerica (L.) Raddi belonging to Aytoniaceae and grown in Japan has a few chemo types. One of the significant types produces aristolane sesquiterpenes (**275**, **276**). The second

238 Caespitenone   **239** Swartzianin A   **240** 3-Hydroxy-3′,4′-methylenedioxybibenzyl

241 Brittonin A   **242** Cyclocolorenone   **243** Eremofrullanolide

234 Tamariscol   **235** 5α,7β(H)-Eudesma-4α,6aα-diol   **236** 4-*epi*-Abusculin A   **247** 13-*epi*-Neoverrucosan-5β-ol

248 13-*epi*-Homoverrucosan-5β-ol   **249** (+)-Gymnomitr-3(15)-en-4α-ol   **250** (+)-Gymnomitran-4-one   **251** Eeverninic acid methyl ester

252 R¹ = OH, R² = R³ = OMe
2-Hydroxy-3,5-dimethoxy-9,10-dihydrophenanthrene
253 R¹ = R² = R³ = OMe
2,3,5-Trimethoxy-9,10-dihydrophenanthrene

254 Pulegone   **255** Menthone   **256** Isomenthone

Figure 35   Chemical markers (**238–256**) of some liverworts.

type is the barbatane sesquiterpenoids (**277–279**) and the third one is classified into cyclomyltaylanes such as sesquiterpenoid (**280**). Some specific monoterpenoids are also used as chemical markers. Monoterpenes amounted to 77% of all compounds present in the volatiles of *Asterella venosa* (Lehm.

& Lindenb.) A. Evans. The main component was geranyl acetate (**281**), a compound with a lemon-like odor. In contrast to the above-mentioned species, which have a pleasant smell, *Asterella* P. Beauv. species grown in Malaysia emit an intense, unpleasant odor, which is due to skatole, composing about

257 3,7-Dimethyl-2,6-octadien-1,6-olide

258 Deoxopinguisone

259 Pinguisone

260 Radulanin K

261 2,2-Dimethyl-7,8-dihydroxy-5-(2-phenylethyl)chromene

262 R = H Perrottetinene
263 R = CO$_2$H Perrottetinenic acid

264 1,4-Dimethylazulene

265 1,2-Dehydro-3-oxo-β-gurjunene

266 R^1 = OH, R^2 = H 6-Hydroxy-5-methoxy-2,2-dimethyl-7-(3-methyl-2-butenyl)-chroman
267 R^1 = H, R^2 = OH 8-Hydroxy-5-methoxy-2,2-dimethyl-7-(3-methyl-2-butenyl)-chroman

268 2,2-Dimethyl-7-(3-methyl-2-butenyl)-chroman-5,8-dione

269 R = H 5-Hydroxy-2,2,2′,2′-tetramethyldichroman
270 R = Me (2,2,2′,2′-Tetramethyl dichroman-5-methyl ether

271 Cinnamolide

272 7α-Chloro-6β-hydroxy-confertifoline

273 2β,9α-Dihydroxy-verrucosane

274 5β,18-Dihydroxy-epi-homo-verrucosane

Figure 36   Chemical markers (**257–274**) of some liverworts.

275 *ent*-Aristolone

276 *ent*-8β-Hydroxyaristolene

277 Gymnomitr-3(15)-en-9-one

278 Gymnomitrol

279 Gymnomitr-3(15)-en-5a-ol

280 Cyclomyltayran-5b-ol

281 Geranyl acetate

282 Conocephalenol

283 Isolepidozene

284 *ent*-Bicyclogermacrene

285 Bicyclogermacrenal

286 4β-Hydroxy-germacra-1(10),5-diene

287 Dihydrocostunolide

288 β-Cyclocostunolide

289 Zaluzanin C

290 R = H Zaluzanin D
291 R = OAc 8α-Acetoxy-zaluzanin D

292 Nenyl acetate

293 Dihydrotulipinolide

294 Dumort-4,6-dien-3-one

295 6β,7β-Epoxy-dumort-4-ene

296 Rearranged dumortenol

297 4α-Hydroxy-nordumort-5-en-6-one

298 R = H Dumortin A
299 R = OMe Dumortin B

300 Dumortin C

Figure 37 Chemical markers (**275–300**) of some liverworts.

20% of the total extract (Asakawa *et al.*, 2009) (Figure 37).

7.3.2 Conocephalaceae

There are two species, *C. conicum* (L.) Dum. and *Conocephalum japonicum* (Thunb.) Grolle, belonging to Conocephalaceae (Marchantiales). The former taxon is four times bigger than the latter one. Both the species are different morphologically. There are three chemotypes of *C. conicum* (L.) Dum in Japan. Types 1, 2, and 3 emit (−)-sabinene, (+)-bornyl acetate, and methyl cinnamate as the major components, respectively, and are responsible for the characteristic odor of each type (Toyota *et al.*, 1997b; Asakawa *et al.*, 2009). In Europe, there are several chemotypes of *C. conicum* (L.) Dum. The typical chemotype mainly elaborates very characteristic sesquiterpene alcohol conocephalenol (**282**) that has not been found in any Asian taxa. It is noteworthy that the German *C. conicum* (L.) Dum. and the Japanese *C. japonicum* (Thunb.) Grolle contain the same monoterpene hydrocarbon, β-sabinene and a sesquiterpene hydrocarbon, isolepidozene (**283**).

The Japanese *C. japonicum* (Thunb.) Grolle produces only limonene as a minor monoterpene hydrocarbon. The major component is isolepidozene (**283**), the diastereomer of the widespread sesquiterpene hydrocarbon, bicyclogermacrene (**284**). Bicyclogermacrene (**284**) and two other germacranes, bicyclogermacrenal (**285**) and 4β-hydroxygermacra-1(10),5-diene (**286**), as well as costunolide (**188**), dihydrocostunolide (**287**), and β-cyclocostunolide (**288**), have been detected as minor components. *C. japonicum* (Thunb.) Grolle is chemically different from the Japanese *C. conicum* (L.) Dum. and rather similar to *Wiesnerella denudata* (Mitt.) Steph., which was recently separated from Conocephalaceae family. *C. conicum* (L.) Dum. does not produce germacrane-type sesquiterpene lactones, while *W. denudata* (Mitt.) Steph. elaborates the same lactones as those found in *C. japonicum* (Thunb.) Grolle (Asakawa, 2004; Asakawa *et al.*, 2009).

7.3.3 Wiesnerellaceae

W. denudata (Mitt.) Steph. was moved from Conocephalaceae to Wiesnerellaceae family, which has two genera, *Wiesnerella* Schiffn. and *Dumortiera* Nees. There are three different chemical races of *W. denudata* (Mitt.) Steph., the costunolide–guaianolide type, the costunolide type, and the guaianolide type (Asakawa, Ludwiczuk, and Nagashima, 2013). Guaianolides, for example, zaluzanin C (**289**), zaluzanin D (**290**), and 8α-acetoxyzaluzanin D (**291**), alongside a large amount of neryl acetate (**292**), have been detected with tulipinolide (**201**), dihydrotulipinolide (**293**), and its cyclized β-cyclocostunolide (**290**) in the Japanese *W. denudata* (Mitt.) Steph. The presence of tulipinolide and guaianolides in *W. denudata* (Mitt.) Steph. indicates that this liverwort is more evolved chemically than *C. japonicum* (Thunb.) Grolle as the latter species produces neither tulipinolide (**201**) nor guaianolides (**289–291**).

There are a few chemotypes of *Dumortiera hirsuta* (Sw.) Nees. The Argentinean *D. hirsuta* (Sw.) Nees elaborates dumortane sesquiterpenoids (**8**, **9**, **294**, **295**), a rearranged dumortane (**296**), and a nordumortane (**299**), along with marchantin C (**70**). The Ecuadorian *D. hirsuta* elaborates the peculiar α-pyrone derivatives, dumortins A–C (**298–300**) (Asakawa, 2004; Asakawa, Ludwiczuk, and Nagashima, 2013). The chemical constituents of *W. denudata* (Mitt.) Steph. are totally different from *D. hirsuta* (Sw.) Nees, although both genera are classified as the same family.

7.3.4 Marchantiaceae

There are two genera of Marchantiaceae, *Marchantia* L. and *Preissia* Corda. The former genus has three species in Japan, *M. polymorpha* L., which is the most popular liverwort in the world, *Marchantia paleacea* var. *diptera* (Mont.) Hatt., and *Marchantia emarginata* subsp. *tosana* (Steph.) Bischl. The three species are morphologically different from each other. *M. paleacea* var. *diptera* (Mont.) Hatt. is much harder than *M. polymorpha* L. and *M. emarginata* subsp. *tosana* (Steph.) Bischl., which are smaller than the first two species. The first species grows on the soil near living houses and in green houses and the latter two species have been found on wet soil in the mountains. These three species produce the bis-bibenzyl, marchantin A (**32**), which is one of the most characteristic metabolites of liverworts, along with its many marchantin analogs. The GC–MS pattern of the ether extract of *M. polymorpha* L.,

301 (−)-Cyclopropane-cuparenol

302 α-Cuprenene

303 β-Cuprenene

304 δ-Cuprenene

305 *ent*-Thujopsene

306 *ent*-Thujopsan-7β-ol

307 *ent*-Thujopsenone

308 α-Chamigrene

309 *ent*-9-Oxo-α-chamigrene

310 Long chain butenolide A

311 Long chain butenolide B

312 3,5,4′-Trimethoxybibenzyl

313 3-(4′-Methoxybenzyl)-5,6-dimethoxyphthalide

314 β-Barbatene
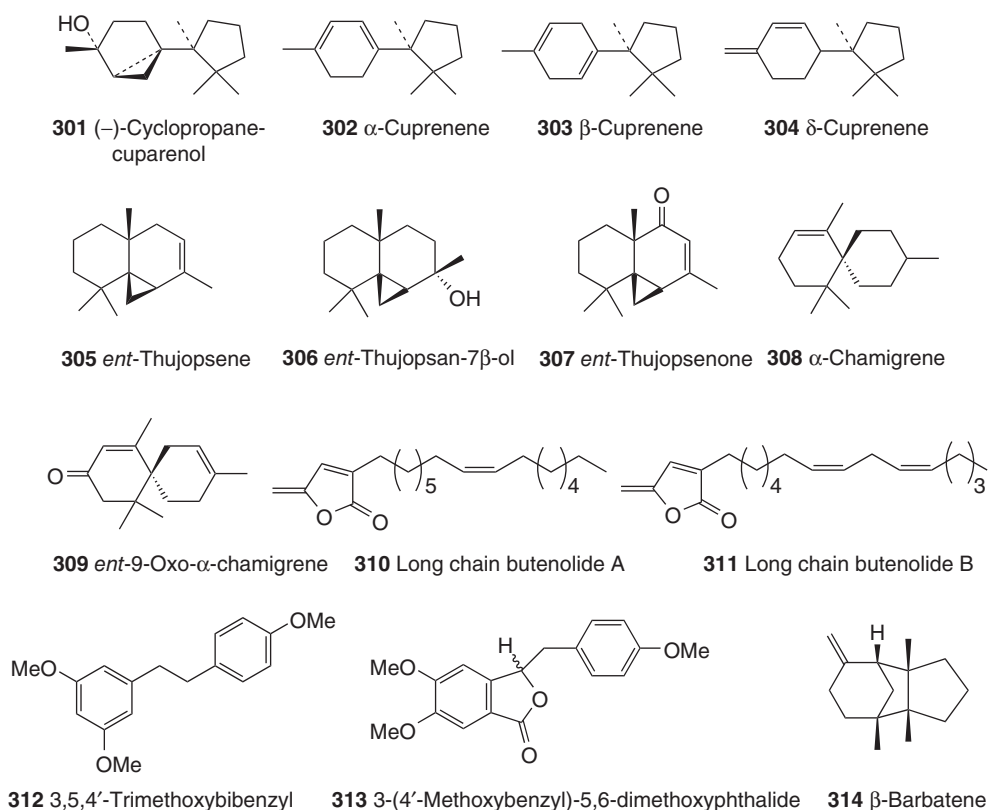
Figure 38 Chemical markers (**301–314**) of some liverworts.

which were collected from Hatacho, Tokushima, and Toyama in Japan, is almost the same as that of the Polish specimen. Thus, the geographical difference in the volatile components has not been found in this species.

(−)-Cyclopropanecuparenol (**301**) is the main component identified in the Japanese *M. polymorpha* L. thalli and female sporophytes. This liverwort also produces other cuparanes, α- (**302**), β- (**303**), and δ-cuparenes (**304**); thujopsene (**305**), thujopsane-2β-ol (**306**), thujopsenone (**307**), α-chamigrene (**308**), and *ent*-9-oxo-α-chamigrene (**309**) have also been identified. These sesquiterpenoids are also significant chemical markers of *M. polymorpha* L. The sesquiterpenoid composition of *M. polymorpha* L. thalli was quite similar to that of female sporophytes. In contrast to thalli, female sporophytes elaborated hexadecanoic acid, its ethyl ester, (*Z,Z*)-9,12-octadecadienoic acid, and oleic acid. *M. paleacea* var. *diptera* (Mont.) Hatt. elaborates long-chain butenolides (**310**, **311**), a

characteristic secondary metabolite of this species that has not yet been isolated from or detected in *M. polymorpha* L. and *M. emarginata* subsp. *tosana* (Steph.) Bischl. (Asakawa *et al.*, 2009) (Figure 38).

7.3.5 Monosoleniaceae

Monosolenium tenerum Griff. grows near living houses and green houses of botanical gardens. The secondary metabolites of this species are very simple. In the German and Chinese *M. tenerum* Griff., only two components 3,5,4′-trimethoxybibenzyl (**312**) and 3-(4′-methoxybenzyl)-5,6-dimethoxyphthalide (**313**) have been identified (Asakawa *et al.*, 2009). The compound was isolated previously from the Australian liverwort *Frullania falciloba* Taylor, ex Lehm., belonging to Jungermanniales. The chemical constitution of *M. tenerum* Griff. is absolutely different from that of all Marchantiale species so far investigated chemically. It is quite interesting to note

Table 6 The distribution of specific sesqui- and diterpenoids and bis(bibenzyls) in liverworts.

Order genus	Drimanes	Pinguisanes	Sacculatanes	Cyathanes	Bis(bibenzyls)
Metzgeriales					
Aneurea	—	+	—	—	—
Riccardia	—	—	+	—	+
Makinoa	+	—	+	—	—
Pellia	—	—	+	—	+
Pallavicinia	—	—	+	—	—
Blasia	—	—	—	—	+
Fossombronia	—	—	+	+	—
Jungermanniales					
Bazzania	+	—	—	—	+
Ptychantus	—	+	—	—	+
Porella	+	+	+	—	—
Trichocoleopsis	—	+	+	—	—
Ptilidium	+	+	—	—	—
Mylia	—	—	—	+	—
Jamesoniella	—	—	—	+	—
Herbertus	—	—	+	—	+
Plagiochila	—	+	—	—	+
Frullania	—	—	—	—	+
Schistochila	—	—	—	—	+
Lepidolaena	—	+	+	—	—

Table 7 The distribution of volatile components of the investigated liverwort species (Ludwiczuk *et al*., 2008).

	Monoterpenoids	Sesquiterpenoids	Diterpenoids	Aromatics	Acetogenins
Metzgeriales					
Pellia endiviifolia[a]	—	+	+	+	+
Pellia endiviifolia[b]	—	+	+	+	+
Pellia epiphylla	—	+	+	—	+
Makinoa crispata	—	+	+	—	+
Symphyogyna brasiliensis	—	+	+	—	—
Noteroclada confluens	—	+	—	—	—
Fossombronia angulosa	—	+	—	—	—
Jungermanniales					
Plagiochila sciophila	—	+	—	+	+
Frullania tamarisci subsp. *obscura*	—	+	—	—	+
Radula perrottetii	—	+	—	—	—
Trocholejeunea sandvicensis	—	+	—	—	—
Marchantiales					
Reboulia hemisphaerica	+	+	—	+	+
Asterella echinella	+	+	—	—	+
Asterella venosa	+	+	—	—	+
Conocephalum conicum[a]	+	+	—	—	+
Conocephalum conicum[b]	+	+	—	—	+
Conocephalum japonicum	+	+	—	—	+[c]
Wiesnerella denudata	+	+	—	—	+
Dumortiera hirsuta	—	+	—	—	+
Marchantia polymorpha	—	+	—	—	—
Marchantia tosana	—	+	—	—	—
Marchantia paleacea var. *diptera*	—	+	—	—	—
Monoselenium tenerum	—	—	—	—	—

[a] Species collected in Japan.
[b] Species collected in Germany.
[c] Only in sporophytes.

that the present species is closely related chemically to *Frullania* Raddi (Jungermannales) chemotype II, which produce bibenzyls as the major product.

In Table 6, the distribution of specific sesqui- and diterpenoids and bis-bibenzyls in liverworts has been summarized.

The volatiles of the liverwort species are characterized by a large number of sesquiterpenoids, as shown in Table 7. Several acyclic, mono-, bi-, and tricyclic sesquiterpene systems are present, including many unique combinations. The investigated liverworts produced mainly sesquiterpene hydrocarbons, with bicyclogermacrene (**284**) and β-barbatene (**314**) being the most widespread in the present and previously analyzed species. Because sesquiterpene hydrocarbons are prevalent throughout liverworts and most liverwort species elaborate the same compounds, it is difficult to use these compounds as chemosystematic markers. However, other sesquiterpenoids, such as eudesmane, eremophilane, germacrane and guaiane sesquiterpene lactones, pinguisane, cuparane, and gymnomitrane sesqui- and sacculatane and cyatane diterpenoids can be used as chemosystematic markers.

8   CHEMICAL PHYLOGENY OF BRYOPHYTES

The most interesting species belonging to the order Metzgeriales is *F. angulosa* (Dicks) Raddi. The species collected in Greece contained three compounds previously reported in brown algae (Ludwiczuk *et al*., 2008). These are dictyotene (**35**), the main compound of the volatile fraction (14.1%), (*Z*)-multifidene (**37**), and dictyopterene (**38**) (Boland, 1995). It is also noteworthy that *C. hirtellus* (Web.) Mitt. collected in Tahiti and Japan produces the same sex pheromones, dictyotene (**35**) and (*E*)-ectocarpene (**36**) (Komala *et al*., 2009, 2010).

When the species belonging to Metzgeriales are dried, they immediately emit a seashore algae-like odor. The genuine component of this odor is dimethyl sulfide, which is the most important odor of almost all algal species.

This may suggest that some liverworts originated from algae (Asakawa, Ludwiczuk, and Nagashima, 2013). The marine algal components in *Fossombronia* Raddi species are extremely interesting as the species is morphologically quite similar to marine algae, for example, *Ulva* L. species belonging to Chlorophyceae, although their chemical constituents are totally different.

9   CONCLUSION

The bryophytes are very tiny plant groups that are distributed everywhere in the world, except sea. They are considered to have evolved from the Southern Hemisphere to the Northern Hemisphere. Over 20,000 species of the bryophytes have been recognized; however, only a few percent of liverworts have chemically been studied and more than 1000 compounds, lipophilic terpenoids, acetogenins, and aromatic compounds, have been isolated from or detected in bryophytes. The hydrophobic components, such as flavonoid glycosides, have been found in bryophytes. On contrary, the distribution of terpenoids and the other aromatic glycosides are very poor in this group. Further chemical analysis of the species grown in high mountains and subtropical and tropical liverworts will promise the discovery of a great number of interesting compounds and their biological activities.

10   RELATED ARTICLES

Sesquiterpenes and Other Terpenoids; Analytical Strategies for Multipurpose studies of a plant volatile fraction; Headspace Sampling and Gas Chromatography: A Successful Combination to Study the Composition of a Plant Volatile Fraction; Metabolomics.

ACKNOWLEDGMENTS

REFERENCES

Adio, A. M., and König, W. A. (2007) *Tetrahedron Asymm.*, **18**, 1693–1770.

Ando, H., and Matsuo, A. (1984) Applied bryology, in *Advances in Bryology*, ed. W. Schultze-Motel, J. Cramer, Vaduz, vol. 2, pp. 133–224.

Asakawa, Y. (1982) Chemical constituents of the Hepaticae, in *Progress in the Chemistry of Organic Natural Products*, eds. W. Herz, H. Grisebach and G. W. Kirby, Springer, Vienna, vol. 42, pp. 1–285.

Asakawa, Y. (1990a) Biologically active substances from bryophytes, in *Bryophyte Development: Physiology and Biochemistry*, eds. R. N. Chopra and S. C. Bhatla, CRC Press, Boca Raton, pp. 259–287.

Asakawa, Y. (1990b) Terpenoids and aromatic compounds with pharmacological activity from bryophytes, in *Bryophytes: Their Chemistry and Chemical Taxonomy*, eds. D. H. Zinsmeister and R. Mues, Oxford University Press, Oxford, pp. 369–410.

Asakawa, Y. (1993) Biologically active terpenoids and aromatic compounds from liverworts and inedible mushroom *Cryptoporus volvatus*, in *Bioactive Natural Products: Detection, Isolation, and Structural Determination*, eds. S. M. Colegate and R. J. Molyneux, CRC Press, Boca Raton, pp. 319–347.

Asakawa, Y. (1995) Chemical constituents of the bryophytes, in *Progress in the Chemistry of Organic Natural Products*, eds. W. Herz, W. B. Kirby, R. E. Moore, W. Steglich and C. Tamm, Springer, Vienna, vol. 65, pp. 1–618.

Asakawa, Y. (1999) Phytochemistry of bryophytes: biologically active terpenoids and aromatic compounds from liverworts, in *Phytochemicals in Human Health Protection, Nutrition, and Plant Defense*, ed. J. Romeo, Kluwer Academic/Plenum Publishers, New York, pp. 319–342.

Asakawa, Y. (2001) *Phytochemistry*, **56**, 297–312.

Asakawa, Y. (2004) *Phytochemistry*, **65**, 623–669.

Asakawa, Y. (2007) *Chenia*, **9**, 73–104.

Asakawa, Y. (2008a) *Curr. Pharm. Design*, **14**, 3067–3088.

Asakawa, Y. (2008b) *Nat. Prod. Commun.*, **3**, 77–92.

Asakawa, Y., and Ludwiczuk, A. (2008a) *Rośliny Lecznicze w Polsce i na Świecie (Medical Plants in Poland and in the World)*, **2**, 33–53.

Asakawa, Y., and Ludwiczuk, A. (2008b) *Rośliny Lecznicze w Polsce i na Świecie (Medical Plants in Poland and in the World)*, **3**(4), 43–54.

Asakawa, Y., and Ludwiczuk, A. (2009) *Malaysian J. Sci.*, **28**, 229–257.

Asakawa, Y., and Matsuda, R. (1982) *Phytochemistry*, **21**, 2143–2144.

Asakawa, Y., Muller, J. C., Ourisson, G., *et al*. (1976) *Bull. Soc. Chim. Fr.*, 1465–1466.

Asakawa, Y., Tokunaga, N., Toyota, M., *et al*. (1979) *J. Hattori. Bot. Lab.*, **45**, 395–407.

Asakawa, Y., Harrison, L. J., and Toyota, M. (1985) *Phytochemistry*, **24**, 261–262.

Asakawa, Y., Sono, M., Wakamatsu, M., *et al*. (1991) *Phytochemistry*, **30**, 2295–2300.

Asakawa, Y., Toyota, M., Tanaka, H., *et al*. (1995) *J. Hattori. Bot. Lab.*, **78**, 183–188.

Asakawa, Y., Toyota, M., Nagashima, F., *et al*. (2000a) *Heterocycles*, **54**, 1057–1093.

Asakawa, Y., Toyota, M., Tori, M., *et al*. (2000b) *Spectroscopy*, **14**, 149–175.

Asakawa, Y., Toyota, M., Oiso, Y., *et al*. (2001) *Chem. Pharm. Bull.*, **49**, 1380–1381.

Asakawa, Y., Toyota, M., Nagashima, F., *et al*. (2008) *Nat. Prod. Commun.*, **3**, 289–300.

Asakawa, Y., Ludwiczuk, A., Nagashima, F., *et al*. (2009) *Heterocycles*, **77**, 99–150.

Asakawa, Y., Ludwiczuk, A., and Nagashima, F. (2013) Chemical constituents of bryophytes: bio- and chemical diversity, biological activity and chemosystematics, in *Progress in the Chemistry of Organic Natural Products*, eds. D. A. Kinghorn, H. Falk and J. Kobayashi, Springer, Vienna, vol. **95**, pp. 1–796.

Basar, S., Koch, A., and König, W. A. (2001) *Flavour. Fragr. J.*, **16**, 315–318.

Boland, W. (1995) *Proc. Natl. Acad. Sci. U. S. A.*, **92**, 37–43.

Cullmann, F., and Becker, H. (1998) *J. Hattori. Bot. Lab.*, **84**, 285–295.

Ding, H. (1982) *Zhong guo Yao Yun Bao zi Zhi Wu*, Kexue Jishu Chuban She, Shanghai, pp. 1–409.

Friederich, S., Maier, U. H., Deus-Meumann, B., *et al*. (1999a) *Phytochemistry*, **50**, 589–598.

Friederich, S., Rueffer, M., Asakawa, Y., *et al*. (1999b) *Phytochemistry*, **52**, 1192–1202.

Garnier, G., Bezaniger-Beauquesne, L., and Debraux, G. (1969) *Ressources médicinales de la flore francaise*, Vigot Frères Éditeurs, Paris, vol. 1, pp. 78–81.

Gottsegen, A., Nógrádi, M., Vermes, K.-P. M., *et al*. (1990) *J. Chem. Soc. Perkin. Trans.*, **1**, 315–320.

Harinantenaina, L., and Asakawa, Y. (2007) *Nat. Prod. Commun.*, **2**, 701–709.

Harrowven, D. C., and Kostiuk, S. L. (2012) *Nat. Prod. Rep.*, **29**, 223–242.

Harrowven, D. C., Woodcock, T., and Howes, P. D. (2005) *Angew. Chem. Int. Ed.*, **44**, 3899–3901.

Hashimoto, T., Tanaka, H., and Asakawa, Y. (1994) *Chem. Pharm. Bull.*, **42**, 1542–1544.

Hashimoto, T., Okumura, Y., Suzuki, K., *et al*. (1995) *Chem. Pharm. Bull.*, **43**, 2030–2032.

Heinrichs, J., Anton, H., Gradstein, S. R., *et al*. (2000) *Plant Syst. Evol.*, **220**, 115–138.

Hentschel, J., Zhu, R.-L., Long, D. G., *et al*. (2007) *Mol. Phylogent. Evol.*, **45**, 693–705.

Hentschel, J., von Konrat, M., Pócs, T., *et al*. (2009) *Mol. Phylogenet. Evol.*, **52**, 142–156.

Hioki, H., Shima, N., Kawaguchi, K., *et al*. (2009) *Bioorg. Med. Chem. Lett.*, **19**, 738–741.

Huang, W.-J., C-L, W. U., Lin, C.-W., *et al*. (2010) *Cancer Lett.*, **291**, 101–119.

Inoue, H. (1988) *Kagakuasahi*, **8**, 116–121.

Iwai, Y., Murakami, K., Gomi, Y., *et al*. (2011) *PLoS One*, **6**, e19825.

Joulain, D., and Konig, W. A. (1998) The atlas of spectral data of sesquiterpene hydrocarbons, E. B.-Verlag, Hamburg, pp. 1–658.

Karlsson, B., Pilotti, A.-M., Söderholm, A.-C., *et al*. (1978) *Tetrahedron*, **34**, 2349–2354.

Keseru, G. M., and Nógrádi, M. (1995a) *Bioorg. Med. Chem.*, **3**, 1511–1517.

Keseru, G. M., and Nógrádi, M. (1995b) *Nat. Prod. Rep.*, **12**, 69–75.

Koepp, A. E., Hezari, M., Zajicek, J., *et al*. (1995) *J. Biol. Chem.*, **270**, 8686–8890.

Komala, I., Ito, T., Nagashima, F., *et al*. (2009) New sesqui- and diterpenoids from the liverwort *Chandonanthus hirtellus*. 53rd Symposium on Chemistry of Terpenes, Essential Oils and Aromatics, Nara, Japan, November 7–9, Symposium Papers, pp. 266–268.

Komala, I., Ito, T., Nagashima, F., *et al*. (2010) *Phytochemistry*, **71**, 1387–1394.

Ludwiczuk, A., and Asakawa, Y. (2008) *Fieldiana Bot.*, **47**, 37–58.

Ludwiczuk, A., and Asakawa, Y. (2010) *Tropic Bryol.*, **31**, 33–42.

Ludwiczuk, A., Nagashima, F., Gradstein, S., *et al*. (2008) *Nat. Prod. Commun.*, **3**, 133–140.

Ludwiczuk, A., Komala, I., Pham, A., *et al*. (2009) *Nat. Prod. Commun.*, **4**, 1387–1392.

Ludwiczuk, A., Gradstein, S. R., Nagashima, F., *et al*. (2011) *Nat. Prod. Commun.*, **6**, 315–321.

Mohamed, H., Baki, B. B., Nasrulhaq-Boyce, A., *et al*. eds. (2008) *Bryology in the New Millenium*, University of Malaya, Kuala Lumpur, pp. 1–513.

Nagashima, F., Tamada, A., Fujii, N., *et al*. (1997) *Phytochemistry*, **46**, 1203–1208.

Nagashima, F., Murakami, M., Takaoka, S., *et al*. (2004a) *Chem. Pharm. Bull.*, **52**, 949–952.

Nagashima, F., Sekiguchi, T., Takaoka, S., *et al*. (2004b) *Chem. Pharm. Bull.*, **52**, 556–560.

Nagashima, F., Kishi, K., Yamada, Y., *et al*. (2005) *Phytochemistry*, **66**, 1662–1670.

Nagashima, F., Wakayama, K., Ioka, Y., *et al*. (2008) *Chem. Pharm. Bull.*, **56**, 1184–1188.

Otoguro, K., Iwatsuki, M., Ishiyama, A., *et al*. (2011) *Phytochemistry*, **72**, 2024–2030.

Otoguro, K., Ishiyama, A., Iwatsuki, M., *et al*. (2012) *J. Nat. Med.*, **66**, 377–382.

Paul, C., Konig, W. A., and Wu, C.-L. (2001) *Phytochemistry*, **58**, 789–798.

von Reuβ, S. H., and König, W. A. (2005) *Eur. J. Org. Chem.*, 1184–1188.

Rycroft, D. S., and Cole, W. J. (2001) *Phytochemistry*, **57**, 479–488.

Schuster, R. M. (1980) *The Hepaticae and Anthocerotae of North America*, Columbia University Press, New York, vol. **1**, p. 758.

Shi, Y.-Q., Zhu, C.-J., Yuan, H.-Q., *et al*. (2009) *Cancer Lett.*, **276**, 160–170.

Shy, H.-S., Wu, C.-L., Paul, C., *et al*. (2002) *J. Chin. Chem. Soc.*, **49**, 593–598.

Suire, C. (1975) *Rev. Bryol. et Lichenol.*, **41**, 105–256.

Sukkharak, P., Ludwiczuk, A., Asakawa, Y., *et al*. (2011) *Cryptogamie Bryol.*, **32**, 199–209.

Tamehiro, N., Sato, Y., Suzuki, T., *et al*. (2005) *Fed. Europ. Biochem. Soc. Lett.*, **579**, 5299–5304.

Tazaki, H., Hayashida, T., Ishikawa, F., *et al.* (1999) *Tetrahedron Lett.*, **40**, 101–104.

Tazaki, H., Kithara, T., Koshino, H., *et al*. (1998) *Phytochemistry*, **48**, 147–149.

Toyota, M., and Asakawa, Y. (1994) *Flav. Fragr. J.*, **9**, 237–240.

Toyota, M., Ueda, A., and Asakawa, Y. (1991) *Phytochemistry*, **30**, 567–573.

Toyota, M., Koyama, H., and Asakawa, Y. (1997a) *Phytochemistry*, **44**, 1261–1264.

Toyota, M., Saito, T., Matsunami, J., *et al*. (1997b) *Phytochemistry*, **44**, 1265–1270.

Toyota, M., Tanimura, K., and Asakawa, Y. (1998) *Planta Med.*, **64**, 462–464.

Toyota, M., Omatsu, I., and Asakawa, Y. (2001) *Chem. Pharm. Bull.*, **49**, 924–926.

Toyota, M., Omatsu, I., Braggins, J., *et al*. (2009) *Chem. Pharm. Bull.*, **57**, 1015–1018.

Warmers, U., and König, W. A. (2000) *Phytochemistry*, **53**, 645–650.

Wu, P. C. (1982) *Bryol. Times*, **13**, 5.

Xie, C. F., and Lou, H. X. (2009) *Chem. Biodivers.*, **6**, 303–312.

Yoshida, T., Hashimoto, T., Takaoka, T., *et al*. (1996) *Tetrahedron*, **52**, 14487–14500.

Cyclotide Analysis

Susan E. Northfield, Aaron G. Poth, Charlotte D'Souza and David J. Craik

The University of Queensland, Brisbane, Queensland, Australia

1 INTRODUCTION

Cyclotides are a large family of plant-derived mini-proteins. They are characterized by a head-to-tail cyclic peptide backbone and a cystine knot arrangement of their three conserved disulfide bonds. They contain up to 37 amino acids (size ranges from 28 to 37 residues) but most are typically around 30 amino acids in size. They first came to notice based on their presence as a bioactive component in an indigenous medicine used to facilitate childbirth in the Congo region of Africa in the 1960s (Gran, 1970). It was not until the mid-1990s, however, that the cyclotide field started in earnest, with the first determination of a cyclotide structure (Saether *et al*., 1995) and later, the recognition that this prototypic example might be part of a large family of topologically unique proteins (Craik *et al*., 1999).

The first cyclotide discovered was kalata B1, so named after a native medicine called *kalata–kalata* from the plant *Oldenlandia affinis*, which was used to accelerate the process of childbirth. In 1970, the peptide had been isolated and partially characterized by a Norwegian physician, Lorents Gran, who discovered that it comprised 29 amino acids but at that stage had not determined any structural information (Gran, 1970). Some 20 years later the three-dimensional structure of kalata B1 was determined (Saether *et al*., 1995) and found to contain the cyclic cystine knot (CCK) motif (Craik *et al*., 1999). The exceptional stability of kalata B1 was apparent early on, given its isolation from the medicinal tea, which was prepared through boiling of the plant material. Typically, peptides are susceptible to degradation at high temperatures and also in the milieu of proteolytic agents encountered in the GI tract. However, the oral activity of kalata B1 suggests that it survives both of these challenges. Later, work by Colgrave and Craik (2004) demonstrated that the CCK motif structurally defined by kalata B1 was responsible for its exceptional stability against thermal, chemical, and enzymatic proteolysis.

Cystine knots are present in a wide range of plants and animals but in no other case is it combined with a circular backbone and no other protein families have yet been discovered that adopt the CCK motif. Cyclotides are of great interest because of their unique structure and resultant exceptional stability. Their unique structure and potential for pharmacological activity mean that cyclotides have attracted attention as scaffolds for drug design applications, the concept being that, for example, the native kalata B1 scaffold can be chemically reengineered with the insertion of biologically active epitopes into the peptide backbone to produce a desired biological activity in a stable peptide framework (Clark, Daly, and Craik, 2006; Gunasekera *et al*., 2008; Garcia and Camarero, 2010). The target sites for such grafting applications are typically one of the six backbone loops between the six conserved cysteine residues that form the core cystine knot motif (Jagadish and Camarero, 2010; Chan *et al*., 2011; Pinto *et al*., 2012a; Poth, Chan, and Craik, 2013).
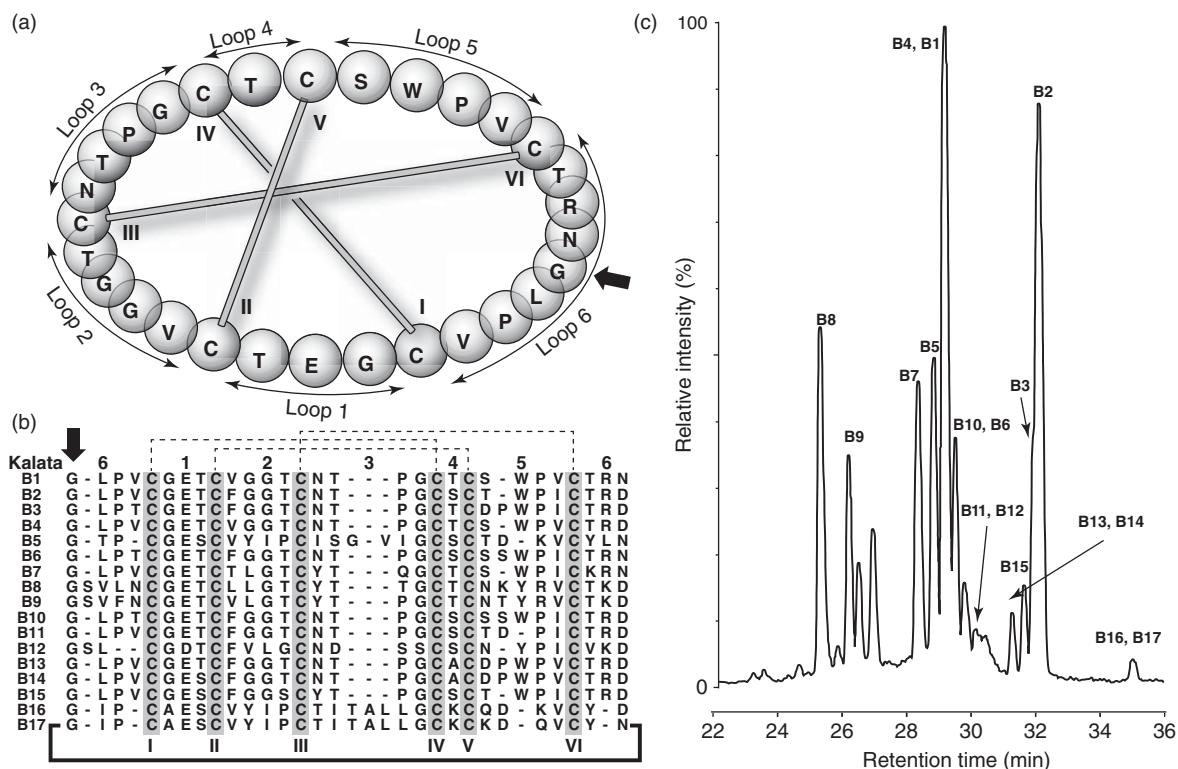
Figure 1 Isolation of cyclotides in *O. affinis*. (a) Schematic representation of the structure of kalata B1, showing the six loops and the position of disulfide-bonded cysteines (represented using Roman numerals) forming the cyclic cystine knot motif represented in all cyclotides. (b) Sequences of kalata B1–B17 extracted from *O. affinis*. Disulfide connectivity is shown using dotted lines. The proto-N-terminal amino acid of the cyclotides, as encoded in the precursor protein is indicated by the black arrow as in panel (a). (c) LC–MS trace for the crude extract of *O. affinis* aerial tissues. Peaks corresponding to kalata B1–17 are indicated (Plan *et al*., 2007).

Since the original discovery of kalata B1, more than 250 other cyclotides have been reported and it is estimated that there may be at least 50,000 members of the family, which would make it one of the largest known families of plant proteins (Gruber *et al*., 2008). An individual plant typically contains multiple cyclotides, as illustrated in Figure 1, which shows a high performance liquid chromatography (HPLC) trace of an extract from *O. affinis*, where 17 individual cyclotides have been identified and many others are known to be present. As can be seen from this HPLC trace, different cyclotides are expressed to different extents in a given tissue, and it is also known that the distribution of cyclotides from one tissue to another also varies. For example, a comparison of LC–MS (liquid chromatography–mass spectrometry) profiles of crude extracts obtained from different plant tissues of *Viola hederacea* from the Violaceae family demonstrated the diversity of constituent

cyclotides. Tissues examined included the leaves, petioles, flowers, pedicels, aboveground runners, belowground runners, bulbs, and roots (Trabi and Craik, 2004).

Following the discovery of the uterotonic activity of kalata B1, a number of other natural product screening studies were conducted in the mid-1990s, including one in which researchers from the National Cancer Institute discovered anti-HIV compounds circulin A and circulin B (reported as macrocyclic peptides) in a Tanzanian plant (Gustafson *et al*., 1994), and another wherein a group at Merck discovered a macrocyclic peptide with neurotensin antagonist activity in a South American plant (*Psychotria longipes*) (Witherup *et al*., 1994). Since then a wide range of other activities have been reported for natural cyclotides, including antimicrobial activity (Tam *et al*. 1999), anti-fouling activity (Göransson *et al*., 2004), cytotoxic activity (Lindholm *et al*., 2002),

as well as a range of biological activities involving pesticide functions, including insecticidal (Jennings *et al*., 2001, 2005), nematocidal (Colgrave *et al*., 2008, 2009), and anti-schistosomal activity (Malagón *et al*., 2013). The latter series of activities has led to the view that cyclotides are present in plants as natural defense agents, probably mainly directed against insect pests (Gruber *et al*., 2007).

It is currently not known why one plant should express a suite of upward of 100 cyclotides but it might be to target multiple pests or a strategy to avoid the development of resistance to an individual defense agent. In any case, an individual plant tissue can be regarded as containing a combinatorial library of cyclotides, each having a different sequence and bioactivity profile. This multiplicity presents some challenges from a bioanalytical perspective, including peak overlap during LC separations and large variations in dynamic range (i.e., in the concentration differences between most abundant and least abundant cyclotides in a given tissue).

There have been a number of recent reviews on cyclotides that have focused on various aspects of their discovery (Daly, Rosengren, and Craik, 2009; Gruber, 2010; Cemazar *et al*., 2012; Göransson *et al*., 2012), characterization (Cascales and Craik, 2010; Ireland *et al*., 2010; Kaas and Craik, 2010), biology (Bohlin *et al*., 2010; Franco, 2011; Pinto, Fensterseifer, and Franco, 2012b), functions (Craik, 2009, 2012; Henriques and Craik, 2012; Mondal and Gerlach, 2012), and applications (Daly and Craik, 2009; Dörnenburg, 2009, 2010; Henriques and Craik, 2010; Jagadish and Camarero, 2010; Smith, Daly, and Craik, 2011; Craik *et al*., 2012a). The first major review in the field was in 2004 when we surveyed the distribution of cyclotides in plants at a time when they had been found mainly in the Rubiaceae and Violaceae families (Craik *et al*., 2004). Since then, cyclotides have been reported in a large number of plant species and their distribution and evolution were recently reviewed (Koehbach *et al*., 2013). Also, recently reviewed have been their biosynthesis (Conlan *et al*., 2010; Aboye and Camarero, 2012; Göransson *et al*., 2012), methods for their three-dimensional structure determination (Daly *et al*., 2010, 2011), and approaches to their chemical (Gunasekera *et al*., 2006; Clark and Craik, 2010; Conibear and Craik, 2011; Craik and Conibear, 2011; Aboye and Camarero, 2012; Reinwarth *et al*., 2012) and recombinant syntheses (Gould *et al*., 2011). Several reviews have also reported on

the use of cyclotides as pharmaceutical templates (Henriques and Craik, 2010; Smith, Daly, and Craik, 2011; Craik *et al*., 2012a). In keeping with the theme of this Handbook, in this chapter, our focus is on the bioanalysis of cyclotides. We first provide a brief background on methods for their discovery and characterization.

2 DISCOVERY

2.1 Currently Known Cyclotides among Angiosperms

Cyclotides have been reported in five plant families, namely Violaceae (violet), Rubiaceae (coffee), Cucurbitaceae (cucurbit), Fabaceae (legume) (Poth *et al*., 2011), and Solanaceae (nightshade) (Poth *et al*., 2012). Also, tentative reports of cyclotide-like peptides in Poaceae (grass) (Mulvenna *et al*., 2006) and Apocynaceae (Gruber *et al*., 2008) families have been made, with a recent report detailing the characterization of acyclotides (homologous sequences that lack backbone cyclization) in the Poaceae family (Nguyen *et al*., 2013). Figure 2 illustrates the currently known cyclotide-containing plant families. To date, more than 250 cyclotide sequences have been characterized and documented in Cybase (www.cybase.org.au), a database dedicated to cyclotides and other naturally occurring cyclic peptides (Kaas and Craik, 2010). Although the full extent of cyclotide distribution is still unclear, it is estimated that there could be at least 50,000 cyclotides yet to be discovered (Gruber *et al*., 2008), making them one of the largest peptide families in the plant kingdom. Cyclotides appear to be ubiquitous in Violaceae (Simonsen *et al*., 2005; Ireland, Colgrave, and Craik, 2006; Burman *et al*., 2009), having been found in every species studied so far.

Cyclotides are expressed within ribosomally synthesized precursor proteins, which harbor one or more cyclotide domains. These can be broadly separated into those encoded by dedicated cyclotide genes and an intriguing example of where a cyclotide-encoding domain has replaced an albumin subunit. Various types of precursor proteins arise from dedicated cyclotide-encoding genes, and for Violaceae and most Rubiaceae cyclotides discovered to date, comprise an endoplasmic reticulum (ER) signal sequence, a prodomain, and one

		Order	Family	No. cyclotides sequenced	Estimated no of species
Basal angiosperms					
Magnoliids		Poales	Poaceae	9 (acyclic)	~18,000
Monocots		Celastrales			~1,200
Basal eudicots		Malpighiales	Violaceae	149	~16,000
		Oxalidales			~1,800
Rosids		Fabales	Fabaceae	24	>20,000
		Rosales			~7,500
		Cucurbitales	Cucurbitaceae	5	~2,300
Asterids		Gentianales	Rubiaceae	43	~16,000
		Lamiales			>23,000
		Solanales	Solanaceae	3	>4,000

Figure 2 Currently known cyclotide-containing families among angiosperms. The phylogenetic tree shows the relationship between the plant orders and the families. The cyclotide- and cyclotide-like-containing families are indicated in light gray and the corresponding orders to which they belong are shown in dark gray. The number of cyclotides sequenced in each family and the estimated number of species in each family are also indicated. There are many cyclotides common to both the Rubiaceae and Violaceae families. *Source*: C.W. Gruber (2010). Adapted from Wiley Periodicals, Inc.

or more cyclotide domains, each bracketed by an N-terminal repeat region (NTR) and a C-terminal region (CTR) (Kaas and Craik, 2010) as exemplified by cycloviolacin O8 and kalata B2 as shown in Figure 3a. Alternatively, they may be encoded within precursor proteins lacking the NTR, where the cyclotide domain directly follows the propeptide [e.g., Phyb A in Solanaceae (Poth *et al*., 2012) and some Rubiaceae cyclotides such as panitide L1 (Nguyen *et al*., 2012)]. Some precursors, including Oak 4 and TIPTOP 1 contain multiple cyclotide domains, which may code for different cyclotides or repeats of the same cyclotide. The genes encoding fabaceous cyclotides are different from the previous examples in that the cyclotide domain is situated within an otherwise intact albumin-1 gene, in place of the PA1b knottin described in all other Fabaceae albumin-1 genes (Poth *et al*., 2011). Trypsin-inhibitory cyclotides from Cucurbitaceae are encoded within precursor proteins that encode repeating cyclotide domains (up to seven) and uniformly also harbor an acyclotide-encoding subunit at their C-termini. This is exemplified by TIPTOP1, which encodes MCoTI-1 (*Momordica cochinchinensis* trypsin inhibitor 1) and MCoTI-2 alongside the linear MCoTI-5 (Mylne *et al*., 2012), as shown in Figure 3a, b.

The presence of a proto-C-terminal asparagine or aspartic acid residue has been demonstrated as integral for the cyclization of cyclotides (Jennings *et al*.,

2001; Saska *et al*., 2007; Gillon *et al*., 2008). The excision and cyclization reactions in their biosynthetic pathway are not yet fully understood; however, an asparaginyl endoproteinase (AEP) enzyme has been implicated. AEP is an endoprotease found in many plants that cleaves the C-terminus of the peptide bonds to Asn and Asp residues. Through this activity, it is considered crucial to the release of the CTR of the mature cyclotide from its precursor (Saska *et al*., 2007; Gillon *et al*., 2008), as well as ligation to the N-terminus (which is presumed to be liberated during an earlier cleavage event). The presumed cyclization of cyclotides by AEP is in keeping with similar findings on other cyclic plant peptides such as SFTI-1, a 14-amino-acid trypsin inhibitor isolated from sunflower seeds (Luckett *et al*., 1999). Recently, it has been reported to share the AEP-mediated maturation process (Mylne *et al*., 2011), as has MCoTI-2 (Mylne *et al*., 2011, 2012).

2.2 Extraction of Cyclotides from Plant Tissues

Cyclotides are commonly extracted from plant tissue using a combination of methanol and a hydrophobic solvent such as dichloromethane (DCM) to remove hydrophobic plant compounds such as lipids. The protocol adopted in our laboratory for the extraction of cyclotides from plant tissues is as follows:

(a)

(b)

Kalata B2	GLPVCG---ETCFGGT-CNTP--GCSCTWP-ICTR--D
Phyb A	G-IGCG---ESCVWIP-CVSAAIGCSCSNK-ICYR--N
Cter M	GLPTCG---ETCTLGT-CYVP--DCSCSWP-ICMK--N
Cycloviolacin O8	GTLPCG---ESCVWIP-CISSVVGCSCKSK-VCYK--N
MCoTI-1	G-GVCPKILQRCRRDSDCPGA---CICRGNGYCGSGSD
MCoTI-2	G-GVCPKILKKCRRDSDCPGA---CICRGNGYCGSGSD
MCoTI-5	Q-RACPRILKKCRRDSDCPGE---CICKGNGYCG----
Panitide L1	QLPICG---ETCVLGT-CYTP--GCRCQYP-ICVP---

Figure 3 (a) Schematic comparison of representative cyclotide precursor proteins. Cyclotides are encoded within precursor proteins typically comprising an endoplasmic reticulum signal sequence (ER signal, light green), an N-terminal propeptide region (gray). This is followed by 1–3 copies of the mature peptide (red), each bracketed by the N-terminal repeat (NTR, blue) and a C-terminal repeat region (yellow), with a short hydrophobic C-terminal tail at the end of the precursor (dark green). In the Fabaceae cyclotide gene, the cyclotide domain is contiguous with the ER signal sequence and encodes an albumin-1 a-chain of 53 amino acids (black). Orange regions depict the positions of knottin and acyclotide domains at the C-termini of TIPTOP1 and Panitide L1, respectively. Where multiple repeats of a cyclotide are present, the region is identified by an arrow with $n = 2$, where n is the number of repeated domains. (b) Alignment of representative cyclotides belonging to various plant families.

1. Harvest the plant material, and if required separate the material into their different component tissues (i.e., flowers, leaves, stems, and roots).
2. Using a mortar and pestle, grind the plant material to powder after precooling with liquid nitrogen.
3. Pour the powdered material into a large beaker and cover with methanol and an equal volume of DCM. Stir the sample overnight at room temperature.
4. Filter through a wire mesh in order to remove plant debris.
5. Transfer the filtrate to a separation flask, along with a small amount of water (one quarter of the filtrate volume) to aid in separation of the two solvent layers. Collect and concentrate the aqueous layer on a rotary evaporator.
6. Load the aqueous layer on a C18 flash column and elute with solvent A (Milli-Q H_2O with 0.05% TFA) mixed with increasing concentrations of solvent B [90% acetonitrile (ACN) in Milli-Q H_2O with 0.05% TFA]. Concentrate the individual fractions on rotary evaporator and lyophilize.
7. Dissolve dried samples in solvent A and analyze using LC–MS with a C18 analytical HPLC column. The sample components elute during a linear gradient of 5–80% solvent B (v/v). The elution of cyclotides typically occurs at >30% ACN (v/v).
8. Purify individual cyclotides using preparative reversed-phase high performance liquid chromatography (RP-HPLC), dissolving the samples in 100% solvent A. Subsequently, load the samples onto the C18 column and elute with a linear gradient of 5–80% solvent B and collect individual peaks. The collected fractions are analyzed by MS, and fractions containing masses

characteristic of cyclotides (e.g., 2.5–4 kDa) are retained.

9. Determine the sequences of the extracted cyclotides using the methods described in Section 2.3.

A recent report has detailed various extraction processes, taking into account solvent, time, and plant material (Yeshak *et al*., 2012).

2.3 Analysis of Cyclotides by Mass Spectrometry

Our standard protocol for the discovery of cyclotides (Gruber *et al*., 2008) first involves the identification of late-eluting peaks in HPLC analyses of plant extracts that have a mass of ~3 kDa. The next step is to narrow down a protein that contains six cysteine residues, which is done by observing a mass shift upon reduction and alkylation. Further tandem MS can be used for sequencing of possible hits identified using this protocol.

To reduce and alkylate a purified cyclotide or a mixture of cyclotides, the following method is used:

1. Prepare a 1 mg mL^{-1} peptide solution in 30 μL of 100 m*M* ammonium bicarbonate (pH 8.0).
2. Add 3 μL of freshly prepared 100 m*M* dithiothreitol and incubate at 60 °C for 1 h.
3. Add 3 μL of freshly prepared 250 m*M* iodoacetamide and incubate at RT in the dark for 30 min.
4. Desalt the sample using a Ziptip (Millipore).
    (i) Prepare a Ziptip by wetting with 70% ACN (v/v) and 0.5% formic acid (v/v).
    (ii) Wash the Ziptip with 0.5% formic acid (v/v).
    (iii) Load the sample on the Ziptip.
    (iv) Wash the Ziptip with 5% ACN (v/v) and 0.5% formic acid (v/v).
    (v) Elute the desalted peptides with 10 μL of 80% ACN (v/v) and 0.5% formic acid (v/v).
5. Mix the desalted sample in a 1 : 1 ratio (v/v) with matrix consisting of a saturated solution of α-cyano-4-hydroxycinnamic acid in 50% ACN (v/v) and 0.1% TFA with 5 m*M* ammonium phosphate, and analyze by MALDI-MS (matrix-assisted laser desorption/ionization mass spectrometry) (this step is further detailed in the following sections).

2.3.1 Sequencing of Cyclotides Following Enzymatic Digestion

Cyclotides can be sequenced using a tandem mass spectrometer equipped with either an ESI or MALDI ionization source. These can also be used as complementary techniques. Among the major differences between ESI and MALDI-MS are the method of sample introduction (solution compared with solid) and the observed charged state of the analyte ions in their resultant spectra (multiply charged and singly charged compared with singly charged). The multiply charged precursor ions produced in ESI-MS serve to enhance the yields of the fragment ions, such that typically, precursors of a higher charge state fragment more readily than those of a lower charge state. This propensity can be used to advantage during their sequence characterization. Multiply charged precursor ions also allow for analysis of peptides with higher molecular weights on spectrometers that have a limited mass-to-charge range.

The use of MALDI-MS to characterize cyclotides in plant extracts has been widely reported, and a general approach to sample preparation for sequencing is as follows:

1. *Preparation of MALDI Matrix*: A saturated solution of α-cyano-4-hydroxycinnamic acid (CHCA) is prepared by dissolving the matrix in 50% ACN and 0.1% TFA with 5 m*M* ammonium phosphate to give a final concentration of 5 mg mL^{-1}. The solution is vortexed and sonicated in a water bath for several minutes and then centrifuged at room temperature. The supernatant is then used in the preparation of samples for MALDI-TOF MS experiments.
2. *Sample Preparation*: An aliquot of isolated native cyclotide solution (obtained by RP-HPLC) or digests is diluted using 50% ACN and 0.1% TFA, combined with the CHCA matrix solution in a 1 μL : 1 μL ratio, and mixed thoroughly. One microliter of the sample/CHCA mixture is then spotted onto the target plate and dried under a gentle stream of nitrogen gas.
3. *Sample Analysis (Example)*: MALDI-TOF MS analysis is carried out in positive ion reflectron mode on a TOF–TOF mass spectrometer. Multiple positions on the MALDI spot are randomly sampled with 100 shots for each position, and the spectra accumulated over 2000 shots, in order to provide representative spectra. The low mass gate

is set to 500 Da, and data are collected over a range of 800–5000 Da, with focus mass of 3000 Da. Spectra are externally calibrated as described by Saska *et al*. (2008) by spotting CHCA matrix 1 : 1 with the ProteoMass MSCL1 peptide and protein MALDI-MS calibration kit calibration mixture (Sigma) diluted to 1 : 400.

4. *Analysis of Peptide Sequences*: The linear peptide fragments are analyzed using MALDI-TOF/TOF and the spectra analyzed in order to determine the peptide sequences.

Before the sequencing of cyclotides via MS/MS, reduction of their disulfide bonds and cleavage of their circular peptide backbone must first be performed, as this information cannot be determined while in their native folded state. The following protocol is used in our lab for the reduction, alkylation, and digestion of cyclotides (Ireland, Colgrave, and Craik, 2006; Plan *et al*., 2007; Wang *et al*., 2007):

1. Dissolve, reduce, and alkylate the peptide as described in Section 2.3 (steps 1–3).
2. Aliquot 30 μL of the sample mixture and desalt using a Ziptip. To confirm reduction and alkylation of cyclotides, compare the MALDI-MS spectra of native versus reduced and alkylated forms (confirm + 348 Da mass shift).
3. Split sample to 3 × 10 μL aliquots.
4. To 10 μL, add 5 μL of endoGlu-C (40 ng μL^{-1}) and 5 μL of 100 mM ammonium bicarbonate, and incubate at 37 °C for 3 h.
5. Separately, to 10 μL, add 5 μL of trypsin (40 ng μL^{-1}) and 5 μL of 100 mM ammonium bicarbonate, and incubate at 37 °C for 1 h.
6. Separately, perform both enzyme digests together at 37 °C for 3 h, for example, 10 μL sample + 5 μL endoGlu-C + 5 μL trypsin.
7. Add an equal volume of 1% formic acid (v/v) (e.g., 20 μL each) to quench the digestion, and desalt each sample using Ziptip (Millipore). The digested samples can be stored at 4 °C before analysis.
8. Analyze the fragments resulting from the digestions using MALDI-MS (confirm + 18 Da mass shift following linearization).
9. Acquire tandem mass spectra for each of the peptide fragments and then analyze and assign b- and y-ions using the protocol described by Medzihradszky (2005) to determine the sequence of the peptide analytes. In each case,

ensure that the entire sequence has been mapped. An example of cyclotide analysis is shown in Figure 4.

10. Use amino acid analysis to confirm peptide composition. Different enzymes may be utilized to help determine the sequence of the cyclotide.

For confident *de novo* sequencing of novel cyclotides, it is vital that extensive sequence coverage be attained by whichever ionization source has been utilized to collect tandem MS data. In our laboratory, typically, MALDI-TOF/TOF is used as an initial check for peptide fragments resulting from enzymatic digestion of cyclotides. However, if the spectra from MALDI-TOF/TOF experiments do not provide sufficient sequence coverage, ESI can be used as a complementary technique. In a sequence-dependent manner, some cyclotides fragment more extensively in MALDI-TOF/TOF experiments, whereas others fragment and provide superior sequence coverage in ESI. As stated, fragmentation is dependent on the sequence features of each peptide analyte, and therefore, we recommend using both ionization techniques to provide the complementary and comprehensive information necessary for confident assignments during *de novo* sequencing.

2.3.2 Case Study: Analyzing Cyclotide Sequences in Mixtures Using MALDI TOF/TOF

The approach to characterization of novel cyclotide sequences discussed in the previous section has been adopted by many researchers investigating cyclotides. Recently, Gruber's laboratory described the characterization of cyclotides from *Viola ignobilis* (Hashempour *et al*., 2013), a species of the violet family from Iran, which was recently reported to contain cyclotides (Hashempour *et al*., 2011). The analysis involved the use of MALDI-TOF MS and tandem MS with a selection of complementary enzymatic digests to attain full peptide sequence coverage by assembling and aligning peptide sequences. They refer to this approach in their report as *sequence fragment assembly*, and it closely mirrors the approach previously described in Section 2.3.1.

A frequently encountered problem using this approach for characterizing cyclotides is the presence of multiple peptides in the sample being analyzed even after separation by RP-HPLC. Typically, this
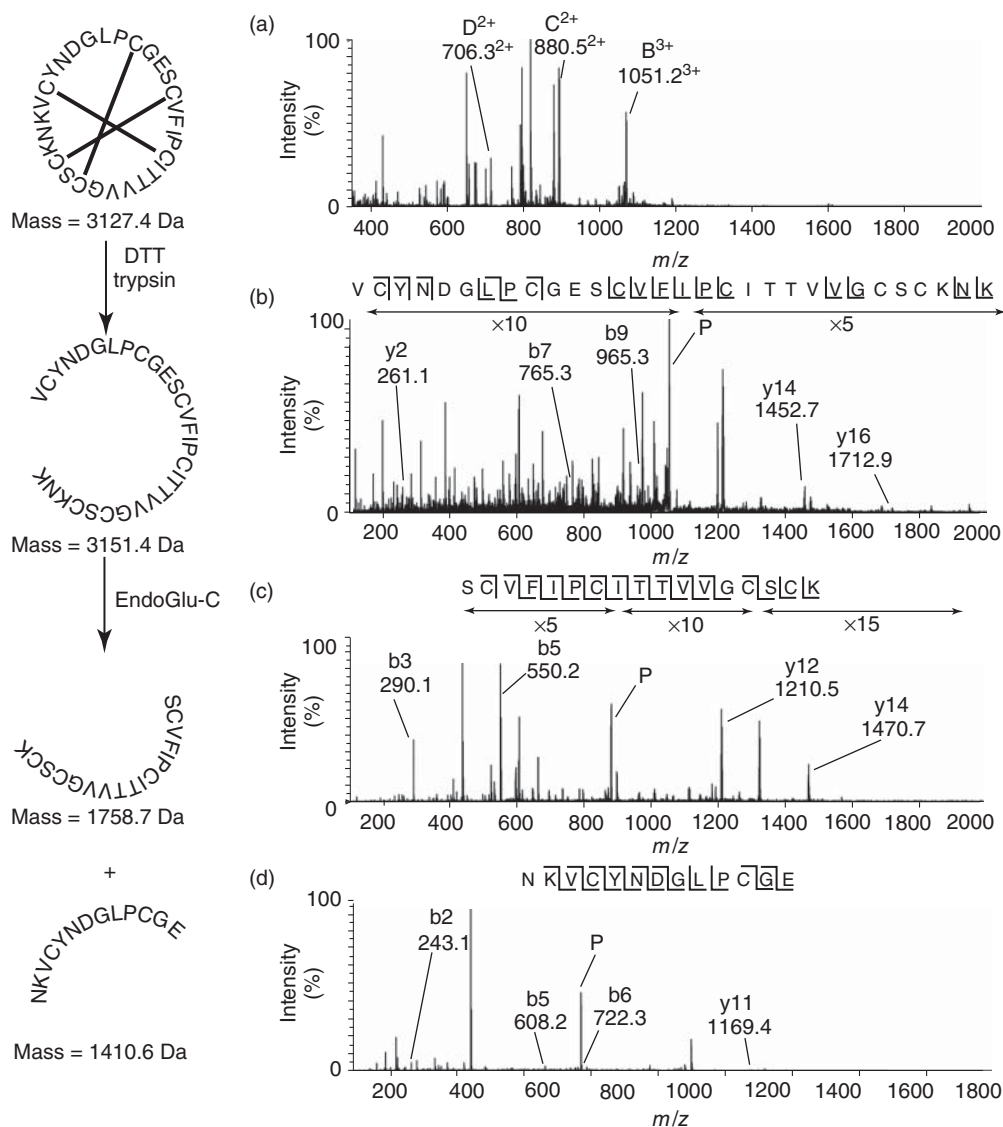
Figure 4 Nanospray MS sequencing analysis of cyclotide Cter H. On the left-hand side of the diagram, the digestion scheme and theoretical masses of proteolytic fragments for Cter H is shown. (a) TOF-MS spectrum of combined trypsin and endoproteinase Glu-C digest products, with cyclotide fragments indicated. The full-length linearized Cter H is represented by B^{3+}. Peptides produced following dual enzymatic cleavage of the cyclic precursor by trypsin and endoproteinase Glu-C are indicated by peaks labeled C^{2+} and D^{2+}. (b) MS/MS spectrum of precursor m/z 1051.2^{3+} (3150.6 Da). (c) MS/MS spectrum of precursor m/z 880.5^{2+} (1759.0 Da). (d) MS/MS spectrum of precursor m/z 706.3^{2+} (1410.6 Da). The precursor peaks in panels B, C, and D are indicated with "P".

multiplicity means that subsequent purification steps are necessary before sequence characterization, and this may be problematic as each purification step results in sample loss. Gruber and colleagues (Hashempour *et al*., 2013) sought to develop a protocol to perform MALDI-TOF/TOF experiments

for characterization of cyclotide sequences using semipure cyclotide fractions containing two or more cyclotides. This protocol involved a combination of single or combined enzyme digestions (to acquire fragments encompassing entire peptide sequences) with tandem MS. Using this approach, 13

sequences from *V. ignobilis* were characterized, and the laborious isolation challenges were overcome.

2.4 Analysis of Cyclotides at the Nucleic Acid Level

Analysis at the nucleic acid level is also a suitable approach for the discovery of novel cyclotides. The screening methodology involves the isolation of RNA for the preparation of cDNA that can be used to characterize novel cyclotide sequences and their respective precursor proteins. In our laboratory, we use a simple method for RNA isolation, which first involves the extraction of ground plant tissue using phenol-chloroform and then precipitation of RNA using lithium chloride. We have developed this method for use at a 1.5 mL tube scale (Mylne *et al.*, 2006, 2010, 2011; Trabi *et al.*, 2009; Qin *et al.*, 2010).

The protocol for the 1.5 mL tube scale RNA extraction in our laboratory is as follows:

1. Harvest the plant material and snap-freeze in liquid nitrogen, storing at $-80\,^\circ$C. It is also possible to store the whole plant material using RNAlater (Ambion).
2. Grind the plant material to a fine powder in a precooled mortar and pestle (liquid nitrogen) or a 1.5 mL plastic tube using a micropestle if there is only a small amount of material. Fine glass beads (e.g., Sigma, catalog number G4649) can be added to produce a fine tissue powder. Tissue powder should not be thawed directly and so liquid nitrogen should be used to pre-chill any equipment (e.g., micropestles, spatulas, and tubes) that comes in contact with the frozen tissue powder.
3. Prepare a mixture of 0.25 mL of phenol and 0.5 mL of homogenization buffer (0.1 M Tris, pH 8.0, 5 mM EDTA, 0.1 M sodium chloride, 0.5% SDS) in a 1.5 mL tube for each sample and just before use, add 5 μL of 2-mercaptoethanol. Heat this mixture to 60 $^\circ$C in a fume cupboard.
4. Add the aforementioned heated mixture to circa 0.3 mL of frozen tissue powder and shake well for 15 min. Use of high quality 1.5 mL tubes is recommended to prevent leakage, and no more than 0.3 mL of tissue powder should be added so as to avoid overfilling the tube.
5. Add 0.25 mL of chloroform to the mixture and shake for a further 15 min.

6. Centrifuge the mixture at top speed in a bench-top centrifuge (14,000 $\times$ g) at room temperature for 10 min. This step will generate a large amount of insoluble interphase. Avoiding this interphase, transfer the upper aqueous layer to a new tube. Add an equal volume of 1 : 1 phenol : chloroform mixture to the tube.
7. Shake the tube for 10 min and then centrifuge for 10 min, as described in the previous step. The supernatant obtained here must then be transferred to a new RNAse-free tube. To avoid any RNAse contamination, particular care must be taken during subsequent steps.
8. Add 0.1 vol of 3 M sodium acetate and 0.8 vol of isopropanol to the supernatant (typically 0.5 mL) to precipitate all nucleic acids.
9. Incubate this at $-80\,^\circ$C for 15 min (may freeze solid) and then centrifuge at top speed at 4 $^\circ$C for 30 min.
10. Pour off the supernatant and invert the tube on a tissue to dry for approximately 10 min and then dissolve the DNA/RNA pellet in 0.5 mL of water.
11. To selectively precipitate the higher molecular weight RNA, add 0.5 mL of 4 M lithium chloride (i.e., final concentration: 2 M) and incubate the sample at 4 $^\circ$C overnight.
12. Centrifuge the sample at top speed at 4 $^\circ$C for 30 min and gently pipette off the supernatant. The RNA pellet obtained at this stage is often colorless and may not be visible.
13. To this RNA pellet, add 1 mL of 80% ethanol and spin the tube at top speed at 4 $^\circ$C for 5 min. A micropipette is generally used to gently remove the wash, and the tube is inverted to drain the pellet.
14. Allow the RNA pellet to dry before dissolving it in 50 μL of water. There may still be some insoluble material present after dissolution of the pellet. Therefore, centrifuge for 5 min at room temperature and transfer the RNA-containing supernatant to a new tube.
15. Quantify the RNA using a UV (ultraviolet) spectrophotometer by its A_{260}. The commonly used measure of RNA sample quality is the $A_{260/280}$ ratio, wherein 1.6–1.8 is the acceptable range and 1.8–2.0 is preferred. To assess the integrity of the RNA sample, we recommend running the sample through electrophoresis. RNA of suitable integrity for cloning experiments will show characteristic bands (typically 4–7 dominant bands); however, for root or seed RNA, which has less

plastid RNA, only two bands may be observed. If no bands are observed, this indicates that the RNA has degraded.

16. To reverse transcribe RNA into cDNA, we typically use oligo(dT) primer in the Invitrogen SuperScript kit and follow the manufacturer's instructions. Several kits are currently available, which can be selected at the user's discretion. In case of 3′ rapid amplification of cDNA ends (RACE), we have successfully used JM38 (5′-GAG CAA CGT CAC GAA AGA AGC GTT TTT TTT TTT TTT T-3′) (Mylne *et al*., 2010). Contaminating genomic DNA can be removed from the total RNA before this step using DNAse. While not strictly necessary for RACE PCR, it may be useful if a user has encountered extra bands in downstream PCR that are not the gene of interest. To amplify full-length transcripts by 5′and 3′ RACE, we have used the SMART RACE cDNA Amplification kit (CLONTECH, catalog number 634923) (Mylne *et al*., 2011).

3   CYCLOTIDE CHARACTERIZATION

Characterization of cyclotides can help provide an understanding of not only their biosynthetic pathway but also their structure and function. The previous section detailed the sequence characterization (i.e., discovery) of cyclotides in both their nucleotide and peptide forms. Here, we present methods for their biophysical characterization. This includes a description of how to use MALDI imaging techniques to analyze cyclotide distribution, the benefits of nuclear magnetic resonance (NMR) techniques for structural analysis of cyclotides, and using membranolytic assays to investigate cyclotide function.

3.1   MALDI Imaging to Identify Novel Cyclotides in the Solanaceae Family

In addition to its utilization for peptide identification and sequence characterization, MALDI-MS can be used to determine the spatial distribution of peptide and small-molecule analytes. MALDI mass spectrometric imaging (MSI) is a technique in which mass spectra are collected in a raster pattern across a section of tissue, generating an average mass spectrum. The signal intensity map can then be overlayed with an image of the sample to reveal the spatial distribution and relative abundance of individual analytes present in the sample section (Caprioli, Farmer, and Gile, 1997). This technique has been used as an academic research tool in the study of disease pathology in animal and human tissues, as well as having applications in medical diagnosis (Ljungdahl *et al*., 2001; Castellino, Groseclose, and Wagner, 2011; Gemoll, Roblick, and Habermann, 2011; Greer, Sturm, and Li, 2011; Menger *et al*., 2011). In 2012, Poth and colleagues described the characterization of cyclotide expression and localization in petunia using a combination of MALDI-MSI and LC-ESI-MS analysis of leaf region extracts (Poth *et al*., 2012).

Detection and sequencing of the cyclotides and acyclotides from *Petunia* x *hybrida* was carried out using the approach described in Section 2. This study (Poth *et al*., 2012) was the first to report cyclotides from the Solanaceae family. To analyze the distribution of cyclotides in petunia leaves, a leaf tissue section was analyzed using MALDI-MSI, and a number of peptide masses that fell within the diagnostic range for cyclotides were observed. One of these signals corresponded to cyclotide Phyb A (*m/z* 3069) and appeared to correlate to the vascular features observed on the prepared leaf section (Figure 5). In addition, signals that did not correspond to characterized cyclotides were detected during the analysis. Signals at *m/z* 3110 and 3346 were abundant in areas of the leaf section distinct from *m/z* 3069.

In an analyte-specific manner, either differential or ubiquitous localization of peptide signals was observed in the MALDI-MSI experiment. This demonstrated that the intensities of nonlocalized analytes were mostly uniform across regions comprising cells of varying sizes, as well as that other analytes that did not locate to a specific region did not cluster together. Altogether, the results demonstrated the nonuniform distribution of the major cyclotide mass, consistent with its colocation with the vasculature structures of the leaf. This work validated MALDI-MSI as a suitable technique for providing information on the spatial relative abundance of peptide analytes *ex planta*.
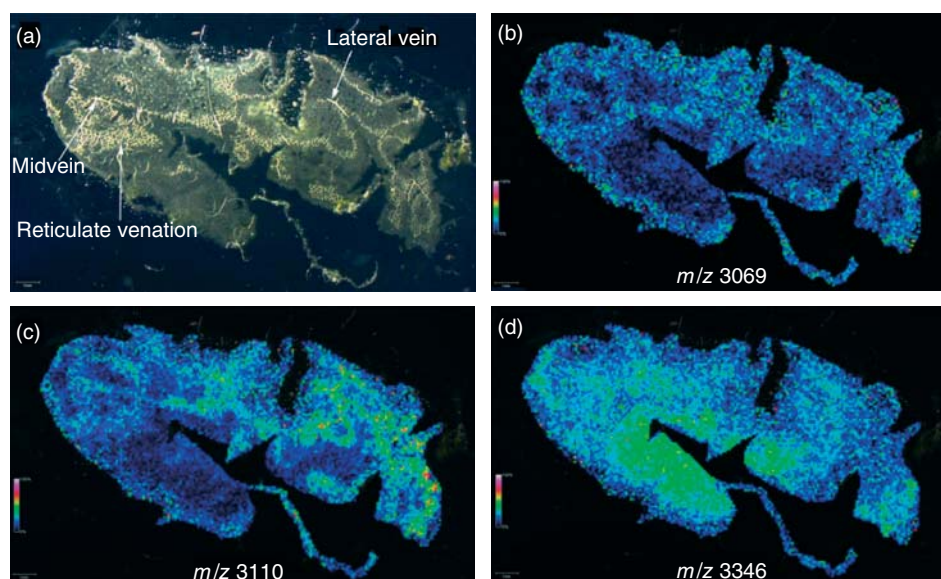
Figure 5 MALDI-MSI analysis of a petunia leaf, displaying localization of three distinct *m/z* signals, one of which corresponds to the major cyclotide, Phyb A. (a) Dark field microscope image of a paradermal cryosection of *P*. x *hybrida* leaf, indicating vascular features. (b) Localization of cyclotide Phyb A, with *m/z* 3069 ± 5 Da. (c) Localization of *m/z* 3110 ± 5 Da. (d) Localization of *m/z* 3346 ± 5 Da.

3.2 Structural Analysis of Cyclotides Using NMR

NMR is a valuable tool for the characterization of cyclotides and is able to provide information complementary to that obtained from MS. NMR is a particularly useful technique for the structural analysis of cyclotides, being the preferred method of analysis owing to cyclotides' small size, high solubility, and characteristic well-defined structures (Daly, Rosengren, and Craik, 2009). Of the 23 cyclotide structures deposited into the PDB, 22 have been solved by NMR and 1 by X-ray crystallography. We have previously noted that one difficulty associated with the NMR analysis of cyclotides comes from the close packing of the cystines in the cystine knot, which may lead to mistaken conclusions if the disulfide connectivity is inferred from only intercystine NOEs (Nuclear Overhauser effect) (Skjeldal *et al*., 2002; Craik *et al*., 2012b). Detailed approaches for defining cyclotide disulfide connectivity, both chemical and NMR have previously been described (Göransson and Craik, 2003; Rosengren *et al*., 2003). Generally, the procedures for NMR analysis of cyclotides are comparable to those for other disulfide-rich peptides. A detailed overview has recently been described in *Methods of Enzymology* (Clark and Craik, 2012) and so is not repeated here.

3.3 Membranolytic Assays of Cyclotides

Cyclotides have been shown to interact with model membranes and to disturb lipid bilayers (Kamimori *et al*., 2005; Shenkarev *et al*., 2006; Svangard *et al*., 2007; Henriques *et al*., 2011). This was demonstrated by the triggering of electrophysiological responses in lipid bilayers and inducing leakage of lipid vesicle contents (Huang *et al*., 2009). Membrane-binding and membrane disruptive abilities of cyclotides are dependent on membrane composition and correlate with bioactivity (Henriques *et al*., 2011; Sando *et al*., 2011; Henriques and Craik, 2012). For this reason, membrane interaction studies are widely used to characterize the mechanism of action of cyclotides. This can be done *in vitro* using surface plasmon resonance (SPR) and fluorescence methodologies. We recently described these procedures (Craik *et al*., 2012b).

4 QUANTIFICATION OF CYCLOTIDES

Various methods for the quantitation of cyclotides have been developed, enabling their study in complex matrices to answer questions about their ecological roles *in planta*, as well as to assess their potential

as biopharmaceutical scaffolds. In this section, we review and compare published methods for their quantification in plants via MALDI-TOF MS (Saska *et al*., 2008) or a combination of MALDI-TOF MS and LC-ESI-MS approaches (Ovesen *et al*., 2011; Gründemann *et al*., 2012). The bioanalysis and quantification of cyclotides in rat serum is also discussed (Colgrave, Jones, and Craik, 2005).

4.1 Detection Methods in Plants

To gain a greater understanding of the biosynthesis of cyclotides, an efficient and accurate means of quantifying cyclotides in plant tissue is required. The large dynamic range over which cyclotides exist in nature (up to 14 mg g^{-1} dry plant weight per cyclotide) (Ovesen *et al*., 2011), along with their similar compact structures and multiplicity, means that relatively low resolution protein analysis techniques such as gel electrophoresis, immunoblotting, or staining are poor methods for cyclotide analysis and thus quantification in plant extracts. Given their convenient throughput and automated sample introduction, as well as high resolution and high sensitivity, MALDI-TOF MS and LC-ESI-MS are the preferred methods for quantification of cyclotides. Examples of both approaches are discussed herein.

4.1.1 Quantification of Cyclotides in Plants Using MALDI-TOF MS

Until recently, MALDI-TOF MS was not commonly used for quantification of peptides in complex mixtures because of poor spot-to-spot reproducibility and hence variation in signal ion intensities. However, MALDI has been successfully applied in the quantification of peptides (Saska *et al*., 2008), amino acids (Zabet-Moghaddam, Heinzle, and Tholey, 2004), oligonucleotides (Zhang and Gross, 2000), and natural products identified in food sources (Camafeita *et al*., 1997; Wang, Sporns, and Low, 1999) and is favored for its rapid, sensitive, and robust analyses. To overcome the issue of reproducibility, the cyclotide signal can be normalized to an internal standard (IS) peptide, which is introduced to the sample before MS analysis (Saska *et al*., 2008). For example, although kalata B2 has been shown to be a reliable IS when quantifying kalata B1 in animal biofluids, it

cannot be used for quantification in plant extracts of *O. affinis* because it occurs in natural high abundance in this plant. Sample preparation and analysis for quantification of kalata B1 in *O. affinis* extracts has been reported (Saska *et al*., 2008) and is described in the following sections, with particular attention paid to the selection of an appropriate IS. Selection of an optimized matrix and application method are critical aspects for the production of uniform cocrystals, and together with data acquisition methods designed to overcome the difficulties surrounding poor spot-to-spot reproducibility and signal degradation can contribute significantly to the robustness of MALDI-TOF in quantification.

4.1.1.1 Sample Preparation
Calibration solutions of kalata B1 were prepared from a stock solution of 0.5 mg mL^{-1} dissolved in Milli-Q water and diluted further in 50% ACN and 0.1% TFA to give solutions with concentrations over a range of 1–400 μg mL^{-1}. Standard solutions were produced through 10-fold dilution in solvent and combined in a 1 : 1 ratio with the IS solution, consisting of the Calibration Mixture 2 (Sequazyme peptide Mass Standards Kit; Applied Biosystems), comprising five peptides. The IS was prepared as a 1/200 or 1/400 dilution in 50% ACN and 0.1% TFA to give a final concentration of 125 fmol μL^{-1} angiotensin I, 125 fmol μL^{-1} ACTH(1–17), 93.75 fmol μL^{-1} ACTH(18–39), 187.5 fmol μL^{-1} ACTH(7–38), and 218.75 fmol μL^{-1} of bovine insulin. The MALDI matrix and plant extract were prepared as described in Section 2, with a 1 : 1 ratio of matrix to calibrated solution or extract containing the IS and spotted onto the MALDI plate.

4.1.1.2 Sample Analysis
MALDI-TOF MS was carried out in positive ion reflector mode using a 200-Hz frequency-tripled Nd:YAG (neodymium-doped yttrium aluminum garnet) laser at 355 nm. Fifty spectra at 30 randomly selected positions were collated per spot over 1000–5000 Da and accumulated into a single spectrum. Calibration curves were produced by plotting the area of the kalata B1 peaks relative to the area of the IS signals, across the range of different cyclotide concentrations. Relative area was determined using the area of the first three peaks in the isotopic clusters of the analyte and the IS protein.

A good correlation was observed between the relative area of the peptide signal and the concentration of

the cyclotide as long as the mass difference between kalata B1 and the IS did not exceed circa 1 kDa. These results demonstrated that within a restricted mass range, peptides structurally unrelated to the peptide of interest can be successfully applied as ISs, and that this should be considered when designing quantification experiments for cyclotides.

To quantify cyclotide production in plant extracts, it is necessary to confirm that the correlation between cyclotide and the IS remains consistent in the extract samples as observed in the calibration solutions. This can be done by generation of a standard curve using samples of known concentrations from plant extracts. Once confirmed, the cyclotides can be quantified in unknown samples. This is exemplified by Saska *et al*. (2008), who were able to use this approach to quantify transgenic kalata B1 production in *Nicotiana benthamiana*.

Although cyclotides have been extensively studied over the last two decades, key aspects of their biosynthesis in plants remain undescribed. For example, the subcellular location of the ligation and cyclization reactions in cyclotide biosynthesis in plant cells was until recently undetermined. To investigate the subcellular location of cyclotide processing, an approach to quantify cyclotides in plant cells is desirable. An understanding of their biosynthetic pathway may lead to improved yields of cyclotides in transgenic plants.

Recently, subcellular targeting of cyclotides in plant cells was reported by Anderson and colleagues (Conlan *et al*., 2011). In this work, a combination of confocal microscopy and MALDI-TOF MS techniques were used to examine the targeting of green fluorescent protein (GFP)-tagged constructs expressed transiently in *N. benthamiana*. Nine constructs were made to identify the regions from the *Oak1* precursor in *O. affinis* that are responsible for the subcellular targeting of kalata B1. Through this approach, they were able to demonstrate that GFP-tagged cyclotide precursors accumulate in living plant cell vacuoles (Conlan *et al*., 2011). This suggests that excision and cyclization processes take place in the vacuole where cyclotides are later stored. Overall, these findings represent a significant step toward understanding of the biosynthetic production of cyclotides and also to the quantification of cyclotides in living plant cells.

4.1.2 Quantification of Cyclotides in Plants Using ESI-MS

In addition to being a method of choice for cyclotide discovery and characterization, ESI-MS is also commonly used for peptide quantification. ESI-MS can be coupled to liquid-based separation tools, allowing for the analysis of complex mixtures with minimal pretreatment of samples. Quantification of cyclotides in plant material using RP-LC-ESI-MS has been reported by Ovesen and coworkers (2011), who sought to develop a simple and rapid method of analysis. This method was developed using kalata B1, kalata B2, and cycloviolacin O2 as reference peptides. These peptides span a range of net charges, hydrophobicities and cyclotide subfamilies, and the analyses described were conducted using an instrument readily available in most research laboratories. Following development of a robust quantification method, it was applied in the quantification of cyclotides in a collection of samples including crude extracts of *O. affinis* and *Viola odorata*, as well as violet tea and candied violets. The described method was demonstrated as capable of quantifying cyclotides in complex samples with a limit of detection (LOD) of 1.7–$4.0\ \mu g\,L^{-1}$ (Ovesen *et al*., 2011).

4.1.2.1 Sample Preparation

Stock solutions of the reference peptides (1 $mg\,mL^{-1}$) were prepared in 25% methanol and 0.1% formic acid and standards were prepared by dilution of stock aliquots using 25% methanol and 0.1% formic acid to give concentrations over the range 0.05–$10\ \mu g\,mL^{-1}$. Cyclotides were quantified in dried plants, tea extracts, and candied violet extracts. Each sample was crushed into fine powder, and 0.1 g was extracted with 70% methanol and 0.1% formic acid and sonicated for 1 h. Following centrifugation of the samples, the supernatant was collected. This extraction process was repeated an additional time for each sample. Methanol phases were combined, and before injection of the sample on the LC–MS, the extract was diluted $1:2$ (v/v) with 0.1% formic acid to ensure that cyclotides would be retained on the column at the beginning of the run.

4.1.2.2 Determining Cyclotide Content

Six replicates for each concentration of the previously detailed reference peptides were analyzed in

this ESI-MS study. Specifically, the peptide was quantified by both method of standard addition (MOSA) to plant material and using a calibration curve independent of the samples. The responses of the reference peptides were fitted by linear regression and had r^2 values >0.968. Using these standard curves, the cyclotide content of the four samples was determined. No significant difference was observed in cyclotide content between using external calibration and using standard addition.

4.2   Detection Methods in Biological Fluids

The stable framework afforded by cyclotides, and thus their resistance to enzymatic degradation, has made them exciting candidates for drug design and development (Henriques and Craik, 2010). As a prerequisite to using this framework for medicinal applications, the pharmacokinetics of cyclotides must first be understood. To achieve this, a sensitive, specific, and reliable method for detecting and quantifying cyclotides in plasma is required. Colgrave et al. (2005) have reported an approach to quantify kalata B1 in rat plasma using MALDI-TOF MS. In that study, rats were administered kalata B1 intravenously, and over a 3-h period, blood samples were taken for analysis. Kalata B2 was selected as an IS.

4.2.1   Dosing of Rats and Sample Preparation

Rats were fasted overnight with access to water before administration of 3 mg kg^{-1} of kalata B1 over a 2-min period. Blood samples (ca. 300 μL) were taken over 3 h from the tip of the tail. The blood was collected into an Eppendorf tube containing 30 μL of 500 IU mL^{-1} of heparin. The samples were centrifuged and 100 μL of serum removed for analysis. Calibration solutions and plasma samples were spiked with kalata B2 as the IS (final concentration: 1 μg mL^{-1}). A twofold excess of ACN was added to precipitate large proteins out of solution, and the samples were shaken and centrifuged before removal of the supernatant for analysis via MS.

4.2.2   Sample Analysis

Calibration curves were first constructed using analyte peak areas from ESI LC–MS of plasma samples containing known concentrations of kalata B1. The LOD in plasma using LC–MS was determined to be 0.5 μg mL^{-1}, which is not sufficiently sensitive for in vitro pharmacokinetic studies. Therefore, MALDI-TOF was used for quantification (samples were prepared for MALDI analysis as described in Section 2.3.1). Calibration curves were constructed using summed peak area ratios of the analyte to the IS versus the theoretical concentration, and each concentration was analyzed in triplicate. MALDI-TOF MS displayed superior sensitivity to LC–MS with an LOD for kalata B1 of 0.05 μg mL^{-1}, which is sufficient for cyclotide quantification in plasma for the purposes of determining its pharmacokinetic profile. However, at concentrations above 10 μg mL^{-1}, the detector was saturated, with analyte signal intensities no longer displaying a linear correlation to concentration. This problem can be overcome by dilution of analyte plasma with a known volume of blank plasma before bioanalysis. Calibration curves were also derived using maximum signal intensity relative to the known concentrations of kalata B1. Linear regression analysis showed an r^2 of 0.9995 when using peak area and 0.9931 for signal intensity for individual curves, and these values improved when using the triplicate values. Unknown sample concentrations were extrapolated from the calibration curve as illustrate in Figure 6.

4.3   Quantification of Cyclotides Expressed in Plant Cell Cultures

The production of cyclotides in plant cell cultures represents an attractive alternative method to their chemical synthesis. Large-scale cultures can in principle be used as biological factories to produce high quality products in controlled conditions and offer the advantage of controlled supply in addition to being renewable and environmentally friendly (Dörnenburg, 2008). Over the last decade, efforts have been made to express cyclotides in cell cultures with the view of using them for drug development and in pharmaceutical applications.

Dörnenburg and colleagues have reported on cyclotide expression in plant cell cultures (Seydel et al., 2007; Dörnenburg, Frickinger, and Seydel,

Figure 6 MALDI-TOF spectra used in the analysis of standard solutions to produce a calibration curve for kalata B1. (a) MALDI-TOF MS spectrum obtained from the analysis of the 1 μg mL^{-1} standard solution of kalata B1. (b) Plasma sample "blank" spiked with the internal standard, kalata B2. (c) 0.05 μg mL^{-1} plasma standard solution with a signal-to-noise ratio of >3 : 1. This is the lowest concentration detectable using this method. (d) MALDI-TOF calibration curve for kalata B1 derived using the relative sum of the peak areas of kalata B1 and internal standard (IS) kalata B2. *Source*: M.L. Colgrave, et al (2005). Reproduced from Elsevier.

2008; Dörnenburg, 2010). The focus has been on production of the cyclotide kalata B1 from *O. affinis*, initiating cell lines from leaf, stem, and root explants of axenic plants (Dörnenburg, 2009). Cell cultures have also been initiated through germination of decontaminated seeds in aseptic conditions on solid agar-containing media (Seydel and Dörnenburg, 2006). Young shoot tips of the plantlets, which contained higher levels of kalata B1, were then used to generate cell cultures (Dörnenburg, 2010). Research into irradiation and agitation, bioreactor technology, and priming and elicitation strategies have been investigated as strategies to improve product yields (Dörnenburg, 2010). Productivity of kalata B1 cultures vary greatly as the growth conditions of the culture are altered (Dörnenburg, 2009), and process productivity of up to 0.5 mg of kalata B1 L^{-1} day^{-1} has been reported (Seydel *et al.*, 2007). Using bioreactor technology to scale up the synthesis, process productivity using a Medusa-type bioreactor on a 25 L scale was estimated to be 21 mg of kalata B1 per day (Dörnenburg, 2009). Although this technology is not yet high yielding enough for commercial applications, high yielding stable transformed cell lines may provide a valuable means for large-scale development, making this a promising method for cyclotide production.

5 CONCLUSION

In this chapter, we described several methods commonly used to study cyclotides. Methods for the discovery of cyclotides are well established and involve a combination of chemical treatments (reduction/alkylation/enzyme digestion) with tandem MS techniques. We found this to be a reliable approach to identify the presence of cyclotides and to elucidate

their sequences. The presence of the CCK motif in cyclotides requires that additional steps be taken in their isolation and characterization compared to linear peptides. One of the main challenges associated with using MS for cyclotide characterization is the absence of an N- or C-terminus and their stable cystine knotted structure, which together resist fragmentation.

Although protocols for the discovery, characterization, and quantification of peptides from plant tissue have been established, the development of new methods, such as those that can be applied in the characterization of molecular events in cyclotide biosynthesis, will address pertinent questions in the field of cyclotide biology. New methods for the characterization of cyclotides are continually being developed and will allow us to expand our understanding of distribution, evolution, and the biosynthetic pathway of cyclotides and how they function. NMR has long been established as the preferred technique for the structural analysis of cyclotides, and the procedures used do not differ greatly from those used for other disulfide-rich peptides. Strong correlations between cyclotide bioactivity and membrane-binding activity have also led to a focus on the development of membranolytic assays. The continued development of accurate and efficient methods for quantitation of cyclotides in plant tissues and biofluids will also aid a greater understanding of cyclotide pharmacokinetics, which will be important for the application of cyclotides in drug design and development.

6 RELATED ARTICLES

Extraction Methodologies: General Introduction; NMR as Analytical Tool for Crude Plant Extracts; Proteomics and Its Research Techniques in Plants

ACKNOWLEDGMENTS

REFERENCES

Aboye, T. L., and Camarero, J. A. (2012) *J. Biol. Chem.*, **287**, 27026–27032.

Bohlin, L., Göransson, U., Alsmark, C., *et al*. (2010) *Phytochem. Rev.*, **9**, 279–301.

Burman, R., Gruber, C. W., Rizzardi, K., *et al*. (2009) *Phytochemistry*, **71**, 13–20.

Camafeita, E., Alfonso, P., Mothes, T., *et al*. (1997) *J. Mass Spectrom.*, **32**, 940–947.

Caprioli, R. M., Farmer, T. B., and Gile, J. (1997) *Anal. Chem.*, **69**, 4751–4760.

Cascales, L., and Craik, D. J. (2010) *Org. Biomol. Chem.*, **8**, 5035–5047.

Castellino, S., Groseclose, M. R., and Wagner, D. (2011) *Bioanalysis*, **3**, 2427–2441.

Cemazar, M., Kwon, S., Mahatmanto, T., *et al*. (2012) *Curr. Top. Med. Chem.*, **12**, 1534–1545.

Chan, L. Y., Gunasekera, S., Henriques, S. T., *et al*. (2011) *Blood*, **118**, 6709–6717.

Clark, R. J., and Craik, D. J. (2010) *Biopolymers*, **94**, 414–422.

Clark, R. J., and Craik, D. J. (2012) *Methods Enzymol.*, **503**, 57–74.

Clark, R. J., Daly, N. L., and Craik, D. J. (2006) *Biochem. J.*, **394**, 85–93.

Colgrave, M. L., and Craik, D. J. (2004) *Biochemistry*, **43**, 5965–5975.

Colgrave, M. L., Jones, A., and Craik, D. J. (2005) *J. Chromatogr. A*, **1091**, 187–193.

Colgrave, M. L., Kotze, A. C., Huang, Y. H., *et al*. (2008) *Biochemistry*, **47**, 5581–5589.

Colgrave, M. L., Kotze, A. C., Kopp, S., *et al*. (2009) *Acta Trop.*, **109**, 163–166.

Conibear, A. C., and Craik, D. J. (2011) *Isr. J. Chem.*, **51**, 908–916.

Conlan, B. F., Gillon, A. D., Craik, D. J., *et al*. (2010) *Biopolymers*, **94**, 573–583.

Conlan, B. F., Gillon, A. D., Barbeta, B. L., *et al*. (2011) *Am. J. Bot.*, **98**, 2018–2026.

Craik, D. J. (2009) *Trends Plant Sci.*, **14**, 328–335.

Craik, D. J. (2012) *Toxins*, **4**, 139–156.

Craik, D. J., and Conibear, A. C. (2011) *J. Org. Chem.*, **76**, 4805–4817.

Craik, D. J., Daly, N. L., Bond, T., *et al*. (1999) *J. Mol. Biol.*, **294**, 1327–1336.

Craik, D. J., Daly, N. L., Mulvenna, J., *et al*. (2004) *Curr. Protein Pept. Sci.*, **5**, 297–315.

Craik, D. J., Swedberg, J. E., Mylne, J. S., *et al*. (2012a) *Expert Opin. Drug Dis.*, **7**, 179–194.

Craik, D. J., Henriques, S. T., Mylne, J. S., *et al*. (2012b) *Methods Enzymol.*, **516**, 37–62.

Daly, N. L., and Craik, D. J. (2009) *Fut. Med. Chem.*, **1**, 1613–1622.

Daly, N. L., Rosengren, K. J., and Craik, D. J. (2009) *Adv. Drug Delivery Rev.*, **61**, 918–930.

Daly, N. L., Chen, B., Nguyencong, P., *et al*. (2010) *Aust. J. Chem.*, **63**, 771–778.

Daly, N., Rosengren, K. J., Troeira Henriques, S., *et al*. (2011) *Eur. Biophys. J.*, **40**, 359–370.

Dörnenburg, H. (2008) *Biotechnol. Lett.*, **30**, 1311–1321.

Dörnenburg, H. (2009) *Biotechnol. J.*, **4**, 632–645.

Dörnenburg, H. (2010) *Biopolymers*, **94**, 602–610.

Dörnenburg, H., Frickinger, P., and Seydel, P. (2008) *J. Biotechnol.*, **135**, 123–126.

Franco, O. L. (2011) *FEBS Lett.*, **585**, 995–1000.

Garcia, A. E., and Camarero, J. A. (2010) *Curr. Mol. Pharmacol.*, **3**, 153–163.

Gemoll, T., Roblick, U. J., and Habermann, J. K. (2011) *Mol. Med. Rep.*, **4**, 1045–1051.

Gillon, A. D., Saska, I., Jennings, C. V., *et al*. (2008) *Plant J.*, **53**, 505–515.

Göransson, U., and Craik, D. J. (2003) *J. Biol. Chem.*, **278**, 48188–48196.

Göransson, U., Sjogren, M., Svangard, E., *et al*. (2004) *J. Nat. Prod.*, **67**, 1287–1290.

Göransson, U., Burman, R., Gunasekera, S., *et al*. (2012) *J. Biol. Chem.*, **287**, 27001–27006.

Gould, A., Ji, Y., Aboye, T. L., *et al*. (2011) *Curr. Pharm. Des.*, **17**, 4294–4307.

Gran, L. (1970) *Medd. Nor. Farm. Selsk.*, **32**, 173–180.

Greer, T., Sturm, R., and Li, L. (2011) *J. Proteomics*, **74**, 2617–2631.

Gruber, C. W. (2010) *Biopolymers*, **94**, 565–572.

Gruber, C. W., Ma, Č., Anderson, M. A., *et al*. (2007) *Toxicon*, **49**, 561–575.

Gruber, C. W., Elliott, A. G., Ireland, D. C., *et al*. (2008) *Plant Cell*, **20**, 2471–2483.

Gründemann, C., Koehbach, J., Huber, R., *et al*. (2012) *J. Nat. Prod.*, **75**, 167–174.

Gunasekera, S., Daly, N. L., Anderson, M. A., *et al*. (2006) *IUBMB Life*, **58**, 515–524.

Gunasekera, S., Foley, F. M., Clark, R. J., *et al*. (2008) *J. Med. Chem.*, **51**, 7697–7704.

Gustafson, K. R., Sowder, R. C. I., Henderson, L. E., *et al*. (1994) *J. Am. Chem. Soc.*, **116**, 9337–9338.

Hashempour, H., Ghassempour, A., Daly, N. L., *et al*. (2011) *Protein Pept. Lett.*, **18**, 747–752.

Hashempour, H., Koehbach, J., Daly, N., *et al*. (2013) *Amino Acids*, **44**, 581–595.

Henriques, S. T., and Craik, D. J. (2010) *Drug Discov. Today*, **15**, 57–64.

Henriques, S. T., and Craik, D. J. (2012) *ACS Chem. Biol.*, **7**, 626–636.

Henriques, S. T., Huang, Y. H., Rosengren, K. J., *et al*. (2011) *J. Biol. Chem.*, **286**, 24231–24241.

Huang, Y. H., Colgrave, M. L., Daly, N. L., *et al*. (2009) *J. Biol. Chem.*, **284**, 20699–20707.

Ireland, D. C., Colgrave, M. L., and Craik, D. J. (2006) *Biochem. J.*, **400**, 1–12.

Ireland, D. C., Clark, R. J., Daly, N. L., *et al*. (2010) *J. Nat. Prod.*, **73**, 1610–1622.

Jagadish, K., and Camarero, J. A. (2010) *Biopolymers*, **94**, 611–616.

Jennings, C., West, J., Waine, C., *et al*. (2001) *Proc. Natl. Acad. Sci. U. S. A.*, **98**, 10614–10619.

Jennings, C. V., Rosengren, K. J., Daly, N. L., *et al*. (2005) *Biochemistry*, **44**, 851–860.

Kaas, Q., and Craik, D. J. (2010) *Biopolymers*, **94**, 584–591.

Kamimori, H., Hall, K., Craik, D. J., *et al*. (2005) *Anal. Biochem.*, **337**, 149–153.

Koehbach, J., Attah, A. F., Berger, A., *et al*. (2013) *Biopolymers*, **100**, 438–452.

Lindholm, P., Göransson, U., Johansson, S., *et al*. (2002) *Mol. Cancer Ther.*, **1**, 365–369.

Ljungdahl, A., Hanrieder, J., Falth, M., *et al*. (2001) *PLoS One*, **6**, e25653.

Luckett, S., Garcia, R. S., Barker, J. J., *et al*. (1999) *J. Mol. Biol.*, **290**, 525–533.

Malagón, D., Botterill, B., Gray, D., *et al*. (2013) *Biopolymers*, **100**, 461–470.

Medzihradszky, K. F. (2005) Peptide sequence analysis, in *Methods in Enzymology*, ed. A. L. Burlingame, Academic Press, Oxford, pp. 209–244.

Menger, R. F., Stutts, W. L., Anbukumar, D. S., *et al*. (2011) *Anal. Chem.*, **84**, 1117–1125.

Mondal, D., and Gerlach, S. (2012) *Chron. Young Sci.*, **3**, 169–177.

Mulvenna, J. P., Mylne, J. S., Bharathi, R., *et al*. (2006) *Plant Cell*, **18**, 2134–2144.

Mylne, J. S., Barrett, L., Tessadori, F., *et al*. (2006) *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 5012–5017.

Mylne, J. S., Wang, C. K., van der Weerden, N. L., *et al*. (2010) *Biopolymers*, **94**, 635–646.

Mylne, J. S., Colgrave, M. L., Daly, N. L., *et al*. (2011) *Nat. Chem. Biol.*, **7**, 257–259.

Mylne, J. S., Chan, L. Y., Chanson, A. H., *et al*. (2012) *Plant Cell*, **24**, 2765–2778.

Nguyen, G. K. T., Lim, W. H., Nguyen, P. Q. T., *et al*. (2012) *J. Biol. Chem.*, **287**, 17598–17607.

Nguyen, G. K. T., Lian, Y., Pang, E. W. H., *et al*. (2013) *J. Biol. Chem.*, **288**, 3370–3380.

Ovesen, R. G., Göransson, U., Hansen, S. H., *et al*. (2011) *J. Chromatogr. A*, **1218**, 7964–7970.

Pinto, M. F. S., Almeida, R. G., Porto, W. F., *et al*. (2012a) *Evid. Based Compl. Alt.*, **17**, 40–53.

Pinto, M. S., Fensterseifer, I. M., and Franco, O. (2012b) Plant cyclotides: an unusual protein family with multiple functions, in *Plant Defence: Biological Control*, eds. J. M. Mérillon and K. G. Ramawat, Springer, Netherlands, pp. 333–344.

Plan, M. R. R., Göransson, U., Clark, R. J., *et al*. (2007) *Chem. Bio. Chem.*, **8**, 1001–1011.

Poth, A. G., Colgrave, M. L., Lyons, R. E., *et al*. (2011) *Proc. Natl. Acad. Sci. U. S. A.*, **108**, 10127–10132.

Poth, A. G., Mylne, J. S., Grassl, J., *et al*. (2012) *J. Biol. Chem.*, **287**, 27033–27046.

Poth, A. G., Chan, L. Y., and Craik, D. J. (2013) *Biopolymers*, **100**, 480–491.

Qin, Q., McCallum, E. J., Kaas, Q., *et al*. (2010) *BMC Genomics*, **11**, 111.

Reinwarth, M., Nasu, D., Kolmar, H., *et al*. (2012) *Molecules*, **17**, 12533–12552.

Rosengren, K. J., Daly, N. L., Plan, M. R., *et al*. (2003) *J. Biol. Chem.*, **278**, 8606–8616.

Saether, O., Craik, D. J., Campbell, I. D., *et al*. (1995) *Biochemistry*, **34**, 4147–4158.

Sando, L., Henriques, S. T., Foley, F., *et al*. (2011) *Chem. Bio. Chem.*, **12**, 2456–2462.

Saska, I., Gillon, A. D., Hatsugai, N., *et al*. (2007) *J. Biol. Chem.*, **282**, 29721–29728.

Saska, I., Colgrave, M. L., Jones, A., *et al*. (2008) *J. Chromatogr. B*, **872**, 107–114.

Seydel, P., and Dörnenburg, H. (2006) *Plant Cell Tissue Organ Cult.*, **85**, 247–255.

Seydel, P., Gruber, C., Craik, D., *et al*. (2007) *Appl. Microbiol. Biotechnol.*, **77**, 275–284.

Shenkarev, Z. O., Nadezhdin, K. D., Sobol, V. A., *et al*. (2006) *FEBS J.*, **273**, 2658–2672.

Simonsen, S. M., Sando, L., Ireland, D. C., *et al*. (2005) *Plant Cell*, **17**, 3176–3189.

Skjeldal, L., Gran, L., Sletten, K., *et al*. (2002) *Arch. Biochem. Biophys.*, **399**, 142–148.

Smith, A. B., Daly, N. L., and Craik, D. J. (2011) *Expert Opin. Ther. Pat.*, **21**, 1657–1672.

Svangard, E., Burman, R., Gunasekera, S., *et al*. (2007) *J. Nat. Prod.*, **70**, 643–647.

Tam, J. P., Lu, Y. A., Yang, J. L., *et al*. (1999) *Proc. Natl. Acad. Sci. U. S. A.*, **96**, 8913–8918.

Trabi, M., and Craik, D. J. (2004) *Plant Cell*, **16**, 2204–2216.

Trabi, M., Mylne, J. S., Sando, L., *et al*. (2009) *Org. Biomol. Chem.*, **7**, 2378–2388.

Wang, J., Sporns, P., and Low, N. H. (1999) *J. Agric. Food Chem.*, **47**, 1549–1557.

Wang, C. K. L., Colgrave, M. L., Gustafson, K. R., *et al*. (2007) *J. Nat. Prod.*, **71**, 47–52.

Witherup, K. M., Bogusky, M. J., Anderson, P. S., *et al*. (1994) *J. Nat. Prod.*, **57**, 1619–1625.

Yeshak, M. Y., Burman, R., Eriksson, C., *et al*. (2012) *Phytochem. Lett.*, **5**, 776–781.

Zabet-Moghaddam, M., Heinzle, E., and Tholey, A. (2004) *Rapid Commun. Mass Spectrom.*, **18**, 141–148.

Zhang, L.-K., and Gross, M. (2000) *J. Am. Soc. Mass Spectrom.*, **11**, 854–865.

Part Four

Biological Analysis

Phenotyping

Qiaosheng Guo and Zaibiao Zhu

Institute of Chinese Medicinal Materials, Nanjing Agricultural University, Nanjing, Jiangsu, PR China

1 INTRODUCTION

Medicinal plants are the main source of natural medicines, and they are increasingly meaningful for renewable resources and provide ingredients that provide health benefits. Plant research deals with questions are responsible for medicinal plants' characteristics in order to gain high yield, high quality, and high safety, or stability under various environmental stress situations.

Owing to its medicinal and industrial importance, the demand for medicinal plants has steadily increased in the global market. To achieve this with existing resources, new plant characteristics need to be identified, quantified, and bred to obtain more productive plant varieties within existing environments. This will require a greater understanding of how the genetic makeup of plants determines their phenotype (visible traits) in high resolution and in high throughput.

Common plant morphological traits of interest include parameters such as main stem height, size and inclination, petiole length and initiation angle, and leaf width, length, inclination, thickness, area, and biomass. For ornamental plants, assessing external shape is a key issue to control visual quality and commercial value (Boumaza *et al.*, 2010), so are medicinal plants.

The phenotype can be seen as the combination of all the morphological, physiological, anatomical, chemical, developmental, and behavioral characteristics that, when put together, represent the individual organism. In daily life, it is the phenotype that is of importance when it comes to food production in crops, health profiles in humans, and husbandry or the ecological performance of organisms in nature. Therefore, phenotyping, which is the process of characterizing the phenotype, is as important as genotyping for establishing the relationship between genes and traits (Pieruschka and Poorter, 2012).

The conventional procedure to collect these data is usually time consuming, expensive, and not suitable for large-scale application, and the precision is also not high enough for monitoring complex plant system. A typical manual phenotypic analysis of 200 plants would require approximately 100 man-hours of work, which is impractical. In the light of the importance of gene discovery and agricultural crop improvement, the development of solutions to automate such a tedious task is imperative (Roy, Tucker, Tester, 2011; Paproki *et al.*, 2012).

Automation and establishment of high throughput systems in the life sciences have significantly progressed within the last decade, mainly in the field of drug discovery and development and in the field of animal behavior. One of the pioneering projects where phenotyping has played an essential role is, for example, the Mouse Clinic and the Human Phenome Project. These projects have revealed the role of the genes and the environment on development, morphology, physiology, metabolism, behavior, and pathology resulting in new disease-preventing

approaches. Analogous development in plant biology will provide opportunities for a revolution in understanding of plant performance, which will enable more efficient breeding strategies (Sato, Isobe, and Tabata, 2010; Pieruschka and Poorter, 2012).

In recent years, imaging-based, automated, noninvasive, and nondestructive high throughput plant phenotyping platforms have become popular tools for plant biology, underpinning the field of plant phenomics (Paproki *et al.*, 2012). With the recent development of multi-parallel and high throughput phenotyping technologies, there are several fast, large-scale evaluations of plant performance that aims to automatize and standardize the phenotyping process (Cabrera-Bosquet *et al.*, 2012; Paproki *et al.*, 2012). This chapter presents the currently used phenotyping techniques.

2 PHENOTYPING TECHNOLOGIES

2.1 Near-Infrared Spectroscopy (NIRS) and Plant Spectral Reflectance

Near-infrared spectroscopy (NIRS) is a chemometric technique that combines spectroscopy and mathematics to rapidly produce indirect, quantitative estimates of the concentration of OH-, NH-, CH-, or SH-containing compounds. Owing to the low cost, rapidity, high precision, and repeatability of this technique, NIRS has been largely implemented in routine laboratory analyses of grain and fodder quality traits in several crop species such as maize, wheat, sorghum, and soybean for assessing several plant traits including nitrogen, moisture, fiber, carbohydrates, amino acids, minerals, and a number of other plant compounds. However, more recently, the use of these techniques for indirect assessment of crop growth and yield performance under potential yield and stress conditions has been addressed. In fact, NIRS is not just limited to analyze any specific compound or metabolite relevant to adaptation to abiotic stresses such as the stable isotope signatures of carbon and oxygen in a fast, albeit nondestructive, and comparatively inexpensive manner; NIRS measurement of kernels has also been used to infer genotypic variability in wheat yield (Cabrera-Bosquet *et al.*, 2012).

Plant spectral reflectance (SR) or hyper SR spectroscopy is based on the different pattern of light reflectance on leaves at different wavelengths through the photosynthetically active radiation (PAR, 400–700 nm), near-infrared radiation (NIR, 700–1200 nm), and shortwave infrared (IR) (up to 2500 nm) regions of the electromagnetic spectrum. SR information from leaves or canopies is used to build vegetation indices (VIs), which are simple operations (e.g., ratios and differences) between SR data at given wavelengths. VIs are related to different plant characteristics such as photosynthetically active biomass, pigment content, and water status. Thus, VIs have been used to predict green biomass, leaf area, chlorophyll content, and yield in wheat and maize under field conditions. Other VIs such as the photochemical reflectance index (PRI) and the water index (WI) may allow inference in photosynthetic efficiency and plant water status (Cabrera-Bosquet *et al.*, 2012).

Both NIRS spectroscopy and plant SR techniques rely on the development of calibration models relating spectral information and reference data (i.e., traditional laboratory analyses) of the trait. Once calibration models have been successfully validated, they can be later employed in routine analyses to predict phenotypic values on external data sets using spectral data and further used in combination with environmental and genotypic data (Cabrera-Bosquet *et al.*, 2012).

NIRS on agricultural harvesters and canopy spectroradiometers can be mounted on tractors, decreasing the work force and costs and making the acquisition of spectral information in a large number of plots much faster and more precise. Aerial platforms are also very much suited for field phenotyping. These include different options such as remote-controlled helicopters and "policopters," airplanes, balloons, or cranes and conveyers (Montes, Melchinger, and Reif, 2007), Figure 1; Cabrera-Bosquet *et al.*, 2012).

In contrast to conventional sample-based methods, NIR on agricultural harvesters secures a good distribution of measurements within plots and covers substantially larger amounts of plot material. Consequently, agricultural harvesters equipped with NIRS reduce the sampling error and yield more representative measurements of the plot material. NIR on agricultural harvesters can also be used successfully to determine dry matter, starch, and crude protein contents in maize grain. In silage maize, the potential of this technology has also been reported for dry matter, starch, and soluble sugars. Measurements
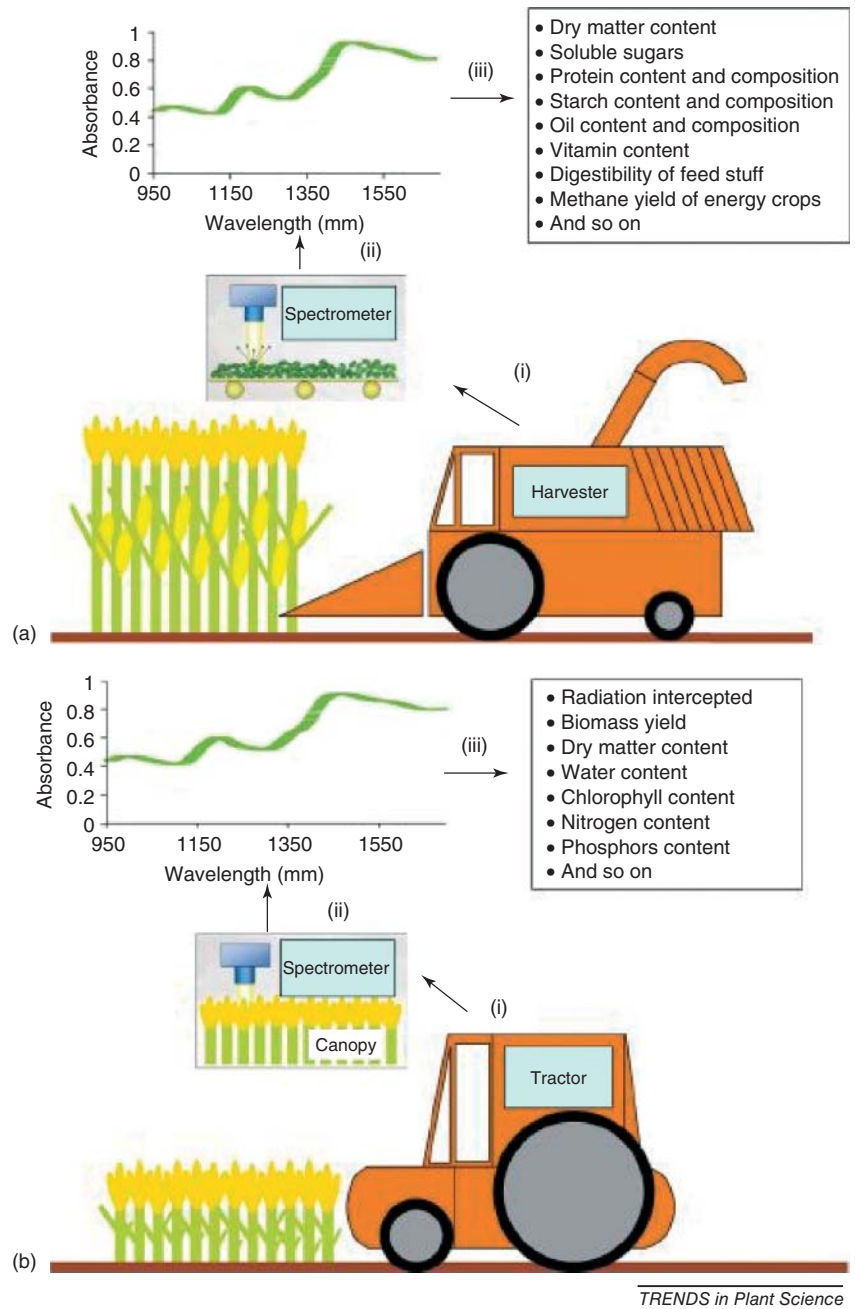
Figure 1 Description of novel phenotyping platforms: (a) near-infrared spectroscopy on agricultural harvesters. (i) The material is harvested, cleaned by air flow, and transported inside the harvesters to be measured with a mounted near-infrared spectrometer. (ii) Spectral measurements are collected while the harvested material is transported under the spectrometer. (iii) Phenotypic values are then predicted using calibration models that relate spectral and phenotypic information. (b) Spectral reflectance of plant canopy. (i) Tractor equipped with canopy spectral reflectance sensors for driving on the field. (ii) Measurement of canopy spectral reflectance. (iii) Prediction of phenotypic values using calibration models (Montes, Melchinger, and Reif, 2007). (*Source*: Reprinted from Trends Plant Sci., 12, Montes J.M., Melchinger A.E., and Reif J.C., Novel throughput phenotyping platforms in plant genetic studies, 433–436, Copyright 2007, with permission from Elsevier.)

of crude protein, digestibility, fiber contents, and energy-related traits can be used successfully to classify genotypes, but, for the precise quantitative trait evaluation, further technical improvements for an optimal sample presentation are crucial. Moreover, feasibility studies on rapeseed suggest that this technology could be used for quantitative determination of dry matter, crude protein, oil, and glucosinolate contents. In addition, amino acids, fatty acids, vitamins, and other quality components, which are currently measured by laboratory NIRS, could also be determined using field-based NIR (Montes, Melchinger, and Reif, 2007).

Although the use of NIRS and SR techniques as a high throughput phenotyping tool is very promising, its application is in its infancy and requires further technical and logistic advances (Cabrera-Bosquet *et al*., 2012). One limitation of the application of these techniques in high throughput phenotyping is handling the extensive data produced.

2.2   Three-Dimensional Laser Scanners or X-Ray Tomographs

Nowadays, three-dimensional (3D) laser scanners or X-ray tomographs make it possible to record a full 3D acquisition and reconstruction of entire plants. 3D laser scanners can provide access to the branching and shoot structure, whereas X-ray tomographs are especially well adapted to trace the 3D structure of the root. Such imaging systems provide access to quantitative morphological measurements that may be targeted beforehand or after the acquisition.

Root architecture plays important roles in plant water and nutrient acquisition. However, accurate modeling of the root system that provides a realistic representation of roots in the soil is limited by a lack of appropriate tools for the nondestructive and precise measurement of the root system architecture *in situ*. Fang, Yan, and Liao (2009) describe a root growth system in which the roots grow in a solid gel matrix that was used to reconstruct 3D root architecture *in situ* and dynamically simulate its changes under various nutrient conditions with a high degree of precision. A 3D laser scanner combined with a transparent gel-based growth system was used to capture 3D images of roots. The root system skeleton was extracted using a skeleton extraction method based on the Hough transformation, and mesh modeling using Ball-B spline was employed.

The findings of Fang, Yan, and Liao (2009) suggest that the growth system simulates a 3D growth environment similar to the root growth situation in the field. This approach solves the problems associated with opaque root system growth environments that cause difficulties for observation of the root system *in situ*. Moreover, the growth medium and the transparent cylinder are easy to make. They can be used to observe root system growth in many ways. The 3D laser scanner used in this study can measure root growth with high precision, up to 0.1 mm. Thus, it is suitable for capturing the growth of fine roots such as rice adventitious roots. Moreover, the scanner calibrates automatically before scanning objects, without a complex manual calibration. The combination of the 3D transparent growth system and the 3D laser scanner not only provides a new method for *in situ* observation of root architecture but also solves the problems of 3D image capture of soft and tiny fibrous root systems, which is harder to capture as a 3D image than more robust tap root systems (Figure 2) (Fang, Yan, and Liao, 2009).

However, when applied to large population of plants, complete 3D acquisitions can be time expensive for high throughput phenotyping and will also produce huge amount of data. In some biological contexts, full reconstruction of entire plants may not be necessary to characterize specific aspects of the morphology. Efforts to expand these investigations into the 3D structure remain constrained by low throughput that requires over 1 h to acquire a single root system, small scanning volumes, and limited quantification capabilities (Clark *et al*., 2011).

Clark *et al*. (2011) introduce a novel 3D imaging and software platform for the high throughput phenotyping of 3D root traits during seedling development (Figure 3). Rice (*Oryza sativa*) plants grown in a gellan gum system. Rotational image sequences consisting of 40 2D images were captured using an optically corrected digital imaging system. 3D root reconstructions were generated and analyzed using a custom-designed software, RootReader3D. This highly versatile phenotyping platform greatly improves throughput and reduces root system capture times to less than 5 min, while also advancing our phenotyping capacity beyond 2D whole root system traits into a range of 3D RSA (root system architecture) and root type-specific traits.
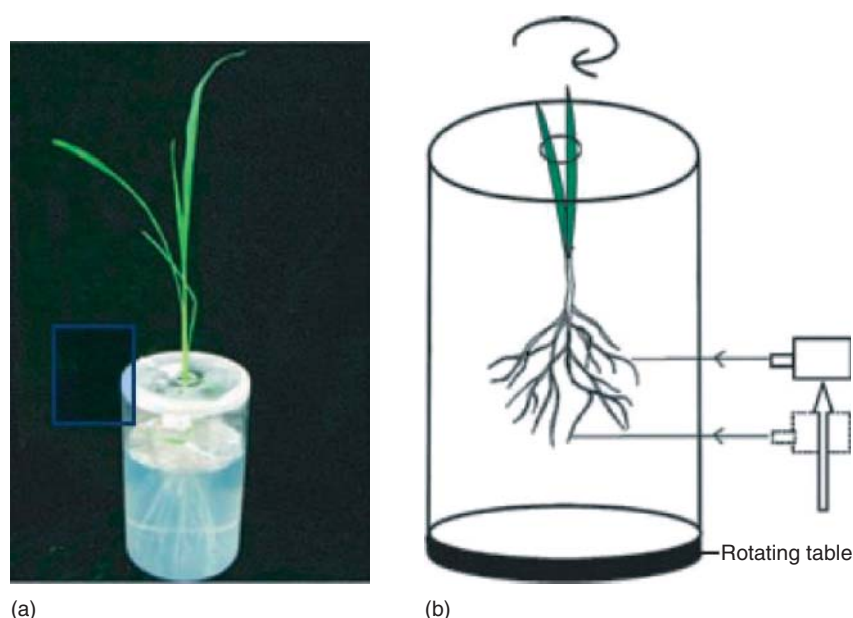
Figure 2 The 3D transparent growth system (a) and scanning strategy of the 3D laser scanner (b). (a) The 3D transparent growth system comprised a transparent cylinder filled with growth medium made from rice hydroponic solution with 0.15% phytagel. (b) The laser scanner scanned as it rotated the growth cylinder and moved the laser beam from bottom to top (Fang, Yan, and Liao, 2009). (*Source*: S.Q. Fang, *et al*. (2009). Reproduced by permission of John Wiley & Sons, Ltd.)

2.3 Low Cost RGB-Depth Camera for 3D Measurements

Depth cameras can be associated with conventional RGB imaging system producing, after registration, a four-component RGB-depth image (Chéné *et al.*, 2012). The accessibility of such RGB-depth imaging systems has recently increased with the introduction of low-cost RGB-depth originally designed for videogames. This opens new possibilities for low cost embedded image-processing vision machines (Gonzalez-Sanchez and Puig, 2011).

Chéné *et al.* (2012) used a Kinect Microsoft RGB-depth camera with the drivers and depth calibration procedure proposed by Microsoft. The Kinect produces depth images from the analysis of spatially structured lighting pattern. It is composed of two CMOS cameras and an IR light source. The first camera, equipped with a 400–800-nm bandpass filter, is dedicated to the RGB imaging. The second camera, equipped with an 850–1100-nm bandpass IR filter, provides the depth image. This system produces 640×480-pixel RGB-depth images coded with a 16-bit dynamic and acquired at a rate of 30 frames per second. The algorithm and applications

presented below are in no way specific to the camera used for illustration, but are relevant to depth camera in general. Motivations for the choice of the specific depth camera used are its small size, low weight, and low cost, which make it suitable for in-field embedded phenotyping and high throughput phenotyping when the system has to be replicated.

The first important step in the 3D analysis of the shoot of plants is the segmentation of the leaves in images. Chéné *et al.* (2012) propose to tackle this segmentation task with a single top-view image acquisition. As visible in Figure 4A, this is not an easy task with standard RGB images as leaves are usually poorly contrasted from one another and upper leaves may create shadow onto the lower leaves. The segmentation appears much easier from the depth image of Figure 4B. To this purpose, the segmentation algorithm of Figure 5 have been developed and implemented. The principle is inspired from the maximally stable extremal regions algorithm introduced in computer vision for the segmentation of single object over background in gray-level images. The algorithm of Figure 5 computes a scan in depth and detects an object to be segmented when a stable number of connected components are reached. Maximally
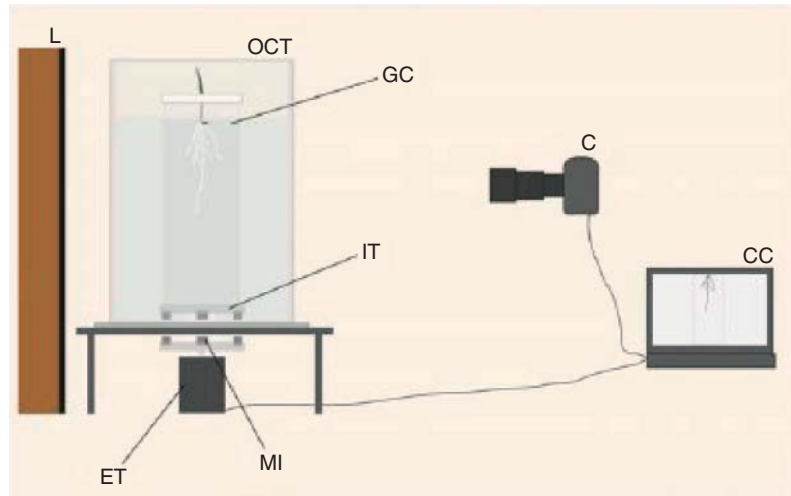
(a)

(b)   (c)

Figure 3   3D root growth and imaging system. (A) Schematic of the 3D imaging system used for capturing image sequences consisting of 40 2D images every 9° of rotation over a full 360° revolution. C, Camera; CC, computer controlling turntable and camera; ET, external turntable; GC, growth cylinder; IT, internal turntable; L, light box; MI, magnetic interface; and OCT, optical correction tank. (B) Growth cylinder containing gellan gum and a 10-d-old Azucena rice seedling. (C) Representative single 2D root system image from an image sequence captured with the 3D imaging system (Clark *et al*., 2011). (*Source*: R. Clark, *et al*. (2011). Reproduced by permission of American Society of Plant Biologists.)
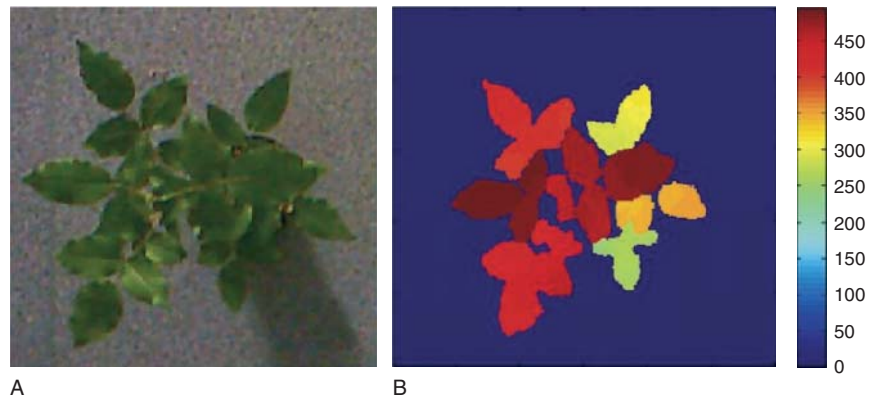
Figure 4 (A) Top view RGB image of a rosebush. (B) Same rosebush as (A) with a depth camera scaled in millimeter with ground as Chéné *et al.* (2012). (*Source*: Reprinted from Comp. Electr. Agr., 82, Chéné Y., Rousseau D., Lucidarme P., Bertheloot J., Caffier V., Morel P., Belin É., and Chapeau-Blondeau F, On the use of depth camera for 3D phenotyping of entire plants, 122–127, Copyright 2012, with permission from Elsevier.)

```
% Algorithm Input–Outputs
Input: Depth image;
Output: Segmented image and Total number of object;
% Geometric parameters
Minimum area = minimum area of a leaf;
Total depth = approximate plant height;
Depth step = to be chosen smaller than the typical distance between leaves;
% Algorithm parameter
Number of iterations = Round (Total depth / Depth step);

For i = 1 ® Number of iterations
    Observed image = Depth image between [Maximum depth - i ´ Depth step, Maximum depth];
    Detection of n(i) Object, i.e. connected component. in Observed image;
    For j = 1 ® n(i)
       If area of Object j ³ Mininmum area
          Storing spatial coordinates c_{i,j}(x,y) of Object j;
          If i ³ 1
             For k = 1 ® n(i−1)
                If c_{i,j}(x,y) = c_{i−1,k}(x,y)
                   Total number of object = Total number of object + 1;
                   Segmented image, pixels c_{i,j}(x,y) = Total number of object;
                   Depth image, pixels c_{i,j}(x,y) = 0;
                End If
             End For
          End If
       End If
    End For
End For
```

Figure 5 Segmentation algorithm of leaves from depth image (Chéné *et al*., 2012). (*Source*: Reprinted from Comp. Electr. Agr., 82, Chéné Y., Rousseau D., Lucidarme P., Bertheloot J., Caffier V., Morel P., Belin É., and Chapeau-Blondeau F, On the use of depth camera for 3D phenotyping of entire plants, 122–127, Copyright 2012, with permission from Elsevier.)

stable extremal regions are therefore extended here to depth images and adapted to multiple object segmentation.

Some prior knowledge on the plant to be acquired is required as input to the algorithm of Figure 5. These prior parameters are the plant height, the minimum expected area of a leaf, and the depth step. The plant height can be obtained automatically if a colored landmark is placed on the ground. This landmark can be localized in the RGB image. The distance between the ground and the closest pixel captured in the depth image provides the size of the plant.

The minimum expected area is assumed as biological prior, and the minimum depth step is limited by the depth resolution of 10 mm. As a result of the constitution of algorithm of Figure 5, any leaf captured by the depth image will be correctly segmented from the other leaves provided it is separated from the other leaves by a distance larger than the depth resolution of the camera. The algorithm of Figure 5 has been tested for illustration on the rosebush of Figure 4. Results are in good agreement with the ground truth. As seen in Figure 6, this agreement holds when the leaf is visible from the top view and when no partial occlusion divides the leaf into multiple objects in the depth image.

The algorithm and applications are relevant to depth camera in general. Motivations for the choice of the specific depth camera used in this report are its small size, low weight, and low cost, which make it suitable for in-field embedded phenotyping and high throughput phenotyping when the system has to be replicated. For large plants, one would probably have to work in outdoor conditions with sunlight. Strong IR radiations from sunlight can significantly degrade the measurements. In field conditions, acquisition with the depth camera used in this manuscript would thus have to be done at night or include a special modulation of the IR light of the depth camera to separate it from sunlight IR radiations (Chéné et al., 2012).

2.4 Fully Automated Digital Imaging Systems

Precisely measure the growth of a crop plant throughout its lifecycle using 3D color digital imaging, which is one of the key developments of the current HTPPs (high throughput phenotyping platforms) designed for potted plants (Cabrera-Bosquet et al., 2012).

During recent decade, a series of techniques have built on computer tomography, which enables temporally continuous, 3D monitoring of roots, soil material, and water content at high resolution (in the 100 mm range; Figure 7a). By providing a wealth of dynamic, structural, and functional data on roots in the soil environment, these techniques are well
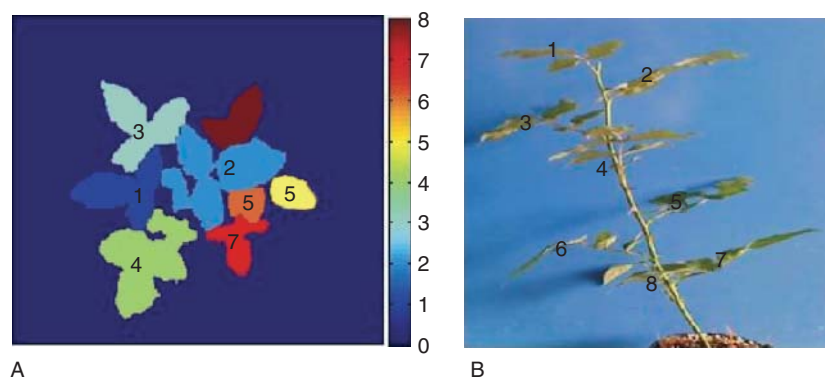
Figure 6   (A) Leaves segmentation result for the rosebush of Figure 4 with the algorithm of Figure 5. Each color corresponds to an object identified as a separate segmented leaf. Numbers in (A) represent the order of appearance of each ground-truth leaf from the top of the plant to the ground. Numbers in (B) come from a direct visual inspection by a human expert from the side view of the rosebush. Ground-truth leaves 1–4, 7, and 8 are correctly segmented. Ground-truth leaf 6, occluded by leaf 3, is not accessible from top view. Ground-truth leaf 5, partially occluded by leaf 2, is segmented in two objects in the depth image. The height of the plant is approximately 0.5 m and the ground-truth leaves are clearly separated by more than the depth resolution of 10 mm as requested by algorithm of Figure 5 (Chéné et al., 2012). (*Source*: Reprinted from Comp. Electr. Agr., 82, Chéné Y., Rousseau D., Lucidarme P., Bertheloot J., Caffier V., Morel P., Belin É., and Chapeau-Blondeau F, On the use of depth camera for 3D phenotyping of entire plants, 122–127, Copyright 2012, with permission from Elsevier.)
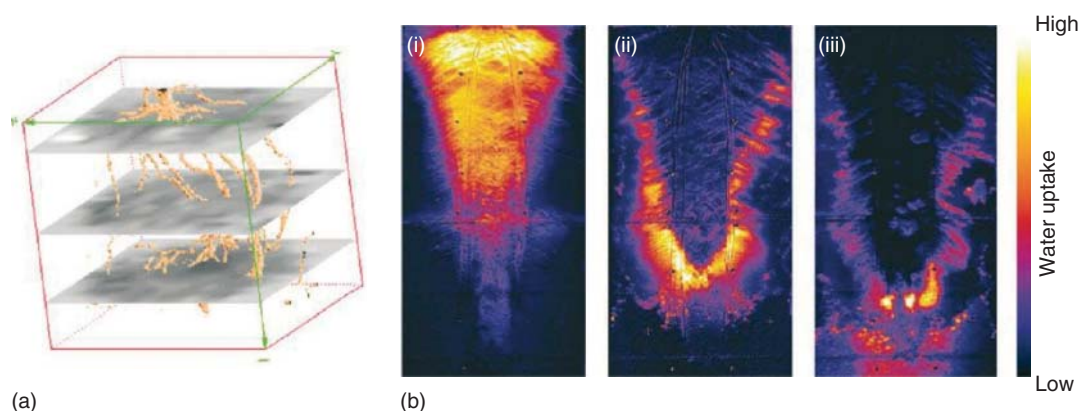
Figure 7 Noninvasive imaging of roots and their environment. (a) Magnetic resonance imaging (CT-MRI) of maize roots and soil water. Roots are indicated by the brown iso-surfaces. The three horizontal layers represent water content changes over a period of 2 weeks. (b) 2D noninvasive imaging of *Lupinus albus* roots and soil water. Three-day sequence (i–iii) illustrating the downward and laterally expanding front of water uptake, estimated from light transmission in a thin rhizotron. (*Source*: Reprinted from Trends Plant Sci, 12, de Dorlodot S., Forster B., Pagès L., Price A., Tuberosa R. and Draye X., Root system architecture: opportunities and constraints for genetic improvement of crops, 474–481, Copyright 2007, with permission from Elsevier.)

suited to the analysis of morphological responses of roots to soil heterogeneity and to support modeling work involving soil–root interactions. A current limitation for genetic studies is the duration of data acquisition (one to several hours depending on the soil volume and spatial resolution). This limitation is avoided using 2D imaging techniques exploiting X-ray absorption and light transmission (Figure 7b). Dynamic approaches to study RSA will eventually lead to more detailed considerations of how genetically regulated phenomena at the cell level translate to the growth and architecture of the whole root system. Kinematic analysis can be used to profile the contribution of various apical segments to the root elongation. This technique has been used in maize to split the apical root meristem in regions revealing high specificities of gene expression and offering new opportunities to dissect root architecture further into its underlying dynamic processes (de Dorlodot *et al*., 2007).

One limitation of the application of these techniques in high throughput phenotyping is handling the extensive data produced.

2.5 Thermal Images

IR thermography "visualizes" surface temperature distribution of an object by focusing the long wave radiation emitted by the object onto a temperature-sensitive detector – the temperature of the object determines how much radiation is emitted at what wavelength following Planck's law. Leaf temperature varies with transpiration rate, which is largely a function of stomatal conductance (Sirault, James, and Furbank, 2009). As leaf temperature is an indicator of stomatal conductance, automating the analysis of thermal images acquired with long-wave IR sensors have great potential for the development of high throughput screens, in particular, if the image capture can also be automated. Thermal imaging has been particularly valuable for screening for stomatal behavior, crop water stress, and water use (Munns *et al*., 2010).

For example, Sirault, James, and Furbank (2009) presented a new screening method for osmotic component of salinity tolerance in cereals using IR thermography. Thermal images of seedlings were acquired between 11:00 and 14:00 h, 3 days after imposing the salt treatment, in the controlled environment chamber using a ThermaCAM SC660 IR camera. In addition, a time-series analysis over 2 h was also acquired to study variation in leaf temperature across time. Thermal data from the seedlings were acquired at 20-s intervals. The SC660 IR camera uses a focal plane array, uncooled microbolometer with 640×480-detector elements, a spectral range of 7.5–13 µm, a thermal resolution of 0.045°C, and an accuracy of $\pm 1\%$. A 24°lens was mounted on the camera and the emissivity was deemed to be
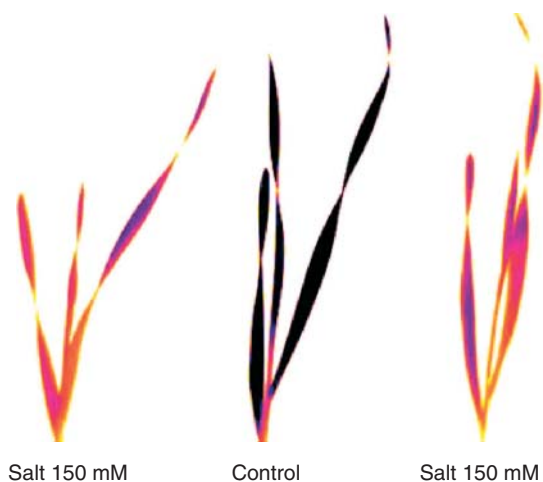
Figure 8 Infrared thermograph of the set-up for assessing leaf temperature differences simultaneously of salt-treated durum wheat seedlings relative to a nonsalt (control) durum wheat seedling. Thermographs were acquired 3 days after the salt stress was imposed. Air temperature was 24°C and RH as 50% (Sirault, James, and Furbank, 2009). (*Source*: X.R.R. Sirault, *et al.* (2009). Reproduced by permission of CSIRO Publishing.)

0.95. The IR camera was placed in the controlled environment chamber 2 h before the measurement series to allow the optics to reach thermal equilibrium with air temperature. The IR camera was positioned at a distance of 0.8 m, perpendicular to the seedlings.

In each image, two salt-treated seedlings were juxtaposed with a (nonsalt) control seedling to assess the difference in leaf temperature because of salinity, rather than determining absolute leaf temperature (Figure 8).

To determine environmental conditions that would maximize transpiration rate under controlled environment chamber conditions, an LI-6400 gas-exchange system was used to examine the influence of vapor pressure deficit (VPD) of the air on both conductance and transpiration rates.

Abaxial stomatal conductance measurements were obtained using an AP4 cycling porometer, three days after the final salt concentrations were reached. Abaxial stomatal conductance had previously been shown to be more sensitive to salinity than adaxial stomatal conductance of wheat grown in controlled environment chambers. Stomatal conductance measurements were made from the mid-portion of the most recently expanded leaf (leaf 2) immediately following IR measurements.

Images were stored directly onto a hard drive as 14-bit resolution files using ThermaCAM Researcher

Pro 2.9. IR images were converted into matrices of raw temperature data in degree Celsius (°C) ($T_{480,640}$ (R)) using ThermaCAM Researcher Pro for the development of the segmentation algorithm.

Temperature matrices $T_{480,640}$ (R) were transformed into 8-bit resolution, gray-level images for viewing in MATLAB, using a custom-built function "tmp2img.m" (Figure 9a). The function scales and normalizes the temperature data using corresponding transformation.

Seedlings were automatically identified in the gray image using Otsu's thresholding method, which assumes only two populations of pixels in an image, that is, background and foreground – the foreground corresponding to the seedling.

To obtain a homogeneous background, the thermographs of the seedlings were acquired against a bronze-colored acrylic. Acrylic does not transmit IR radiation and its different emissivity results in an apparent temperature 2° hotter than the air temperature. This provided a clear threshold value, $i_{threshold}$, for separating background and foreground pixels in the IR images (Figure 9b). A binary matrix, $B_{480,640}$ (R), was computed according to corresponding transformation.

Each plant in the binary matrix was then labeled with an integer value using the "bwlabel" function of the image-processing toolbox. Visual observations of the resulting images demonstrated that the thresholding approach was accurate and effective (Figure 9d). The binary matrix B was then used as a mask to derive the temperature of the two seedlings in the thermograph only. By multiplying arrays $T_{480,640}$ (R) and $B_{480,640}$(R) according to array arithmetic rules, element-by-element multiplication and temperature values for the seedlings were derived.

A structure identifying the location and the number of pixels for each labeled seedling was defined in MATLAB, allowing automatic calculation of the average temperature for each labeled seedling (Figure 9d). The difference between the control and salt-treated seedlings was then computed. Each seedling was represented by a minimum of 3000 independent, thermally calibrated data points.

Microbolometer-based thermal imaging sensors are improving rapidly in spatial resolution and these can be mounted on platforms above the crop on model aircraft, helium balloons, or manned aircraft; however, the speed of acquisition gained by raising the height of the imaging sensor above the crop obviously reduces spatial resolution. The complexities
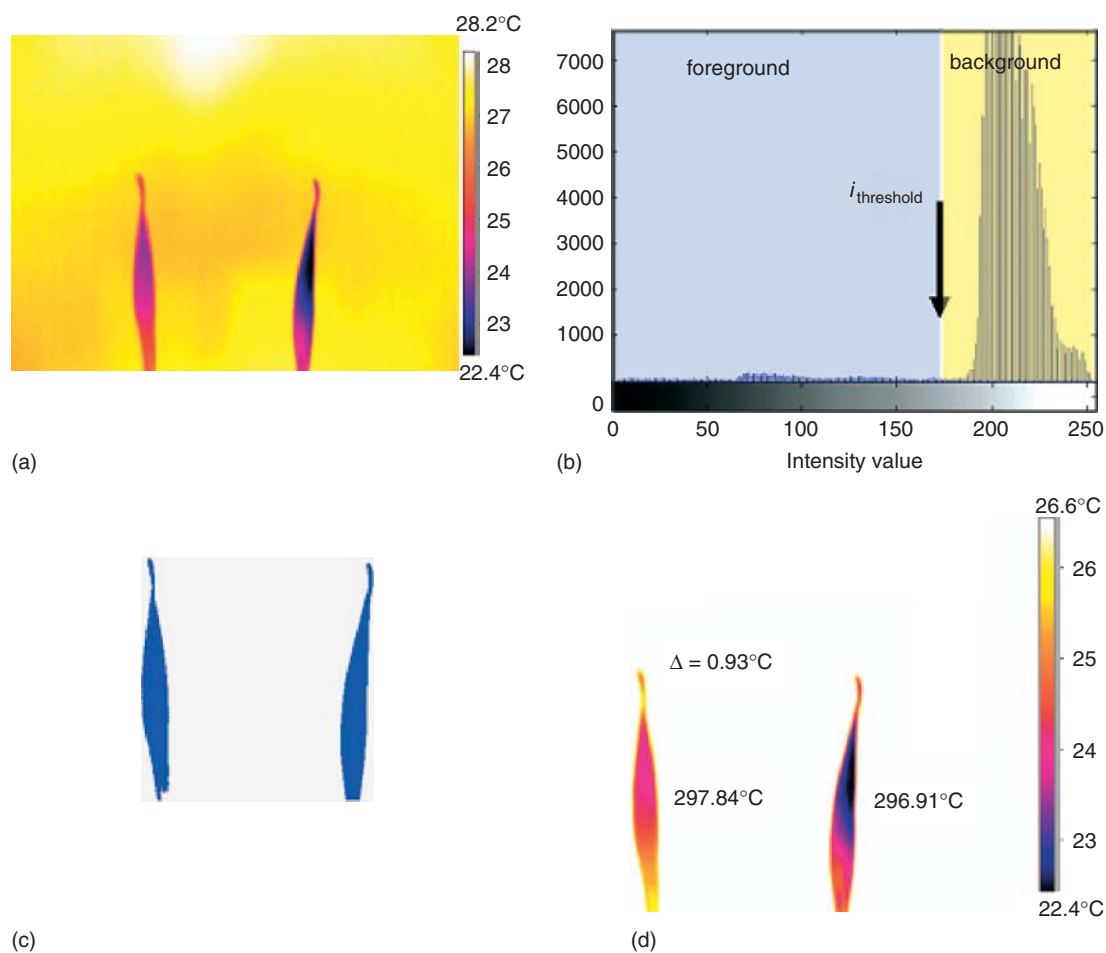
Figure 9 Segmentation algorithm for analyzing infrared thermographic images: (a) infrared thermograph of Himalaya barley in 150-mM NaCl (left) v. nonsalt "control" (right) visualized as an 8-bit gray image (iron bow palette); (b) histogram of gray-level values with $i_{threshold}$ computed according to Otsu's method; (c) binary image of (a); and (d) result of array multiplication to apportion temperature data to seedlings, 3 days after the salt stress was imposed. Air temperature was 24°C and RH was 50% (Sirault, James, and Furbank, 2009). (*Source*: X.R.R. Sirault, *et al.* (2009). Reproduced by permission of CSIRO Publishing.)

of diurnal variation of the radiation load on the crop, angle of view, and solar angle must still be accounted for to obtain biologically meaningful and reproducible results (Furbank and Tester, 2011).

Such screens early in plant development allow many thousands of lines to be assessed rapidly and at low cost relative to techniques requiring measurements across the whole lifecycle (Furbank and Tester, 2011). The main benefits of the thermal imaging screening protocol are its precision, noninvasiveness, and speed, compared with other methods used to measure stomatal conductance, such as porometry (Sirault, James, and Furbank, 2009).

However, several factors such as surface emittance, changes in ambient temperature and humidity, and background thermal radiation could introduce errors in the measurement of absolute leaf temperature. This issue was overcome by measuring relative temperature differences between salt-stressed and control seedlings simultaneously. In addition, to minimize fluctuations in environmental factors, the experiments had better be completed in controlled environment chambers and care was also considered for variability of leaf angles. By screening seedlings rather than more advanced plants, leaf angles were no longer randomly distributed,

providing a comparable surface to the IR camera (Sirault, James, and Furbank, 2009).

2.6 Chlorophyll Fluorescence Imaging (CFI)

Chlorophyll fluorescence imaging (CFI) is a non-destructive method that can be applied repeatedly during plant growth allowing the physiological state of plants to be measured over the course of the plants' lives and in response to different treatments. As a rapid, noncontact, optical technique, CFI is well suited for robotic imaging procedures. The leaf area of whole plants can be measured simultaneously in a process that requires only a few seconds per measurement. CFI provides parameters that describe many aspects of the operation and regulation of photosystem II, from which valuable information about other physiological processes can be inferred (Harbinson *et al*., 2012).

In particular, the availability of high throughput chlorophyll fluorescence measurements opens the door to a range of strategies for obtaining the physiological data necessary for the gene localization procedure. The Φ_{PSII} parameter derived from fluorescence is a useful proxy for the light use efficiency for CO_2 fixation and will make possible the localization of genetic factors for photosynthetic parameters. Once identified, these loci can be exploited in marker-assisted breeding for improved photosynthetic properties.

Photosynthesis is a complex trait whose developmental and acclamatory responses are poorly understood genetically, and that offers the potential for further yield increases. The availability of CFI makes it possible to analyze the genetic determinants of the numerous properties of the photosynthetic process. These include those properties for which there is as yet no physiological model, such as the ability of photosynthesis to tolerate environmental fluctuations (Figure 10) (Harbinson *et al*., 2012).

High throughput phenotyping of chlorophyll fluorescence is a new and promising approach to be implemented in the next step of marker-assisted breeding (Harbinson *et al*., 2012). However, the application of CFI still requires development. It is currently easy to measure individual small plants *in situ* – the Φ_{PSII} of around 103 plants can be measured several times per day with one camera in this way. Working with larger plants is more difficult as

Figure 10 A robotic high throughput system for imaging small plants, such as *Arabidopsis thaliana*, using visible light, near-infrared, and chlorophyll fluorescence. In this design, the camera is attached to a robotic *x*–*y* movement system that allows the imaging head (which includes LEDs to supply the measuring light for the various imaging channels, and a camera with a filter wheel for wavelength-specific imaging) to be moved to a randomly chosen group of plants. The plants are grown hydroponically in individual rock-wool blocks that are securely fixed beneath a black plastic cover plate that also provides a good background for the plant images. The plants are watered using an ebb and flood system and the complete system is installed in a growth cabinet for good environmental control (Harbinson *et al*., 2012). (*Source*: Reprinted from Curr. Opin. Biotech., 23, Harbinson J., Prinzenberg A.E., Kruijer W., and Aarts M.G., High throughput screening with chlorophyll fluorescence imaging and its use in crop improvement, 221–226, Copyright 2012, with permission from Elsevier.)

they need to be moved to an imaging station for measurement, which disturbs the plants and consumes time. Images of larger plants also need to be converted to a 3D form, which is not yet a routine procedure. The results from larger plants are also more difficult to interpret, especially if the plant is not measured in its growth environment, because of the complex acclamatory responses of photosynthesis to shading. In addition, more comprehensive descriptions of photosynthetic properties, such as irradiance or temperature responses, would require longer measurements and moving the plant to an imaging station.

2.7 Light Curtains (LCs) and Spectral Reflectance (SR) Sensors

A HTPP employing light curtains (LCs) and SR sensors is designed mounted on a tractor and evaluated its performance under field conditions.

Employing LCs is a new technique to determine the profile of plants. The concept of sensor fusion has been proposed for crop-weed distinction based on measurements of various geometrical, optical, and physical properties of plants. It relies on the combined use of different but complementary sensor techniques with varying selectivity of plant characteristics. Accordingly, a combination of LC and SR sensors may also be used for investigations of biomass. Montes *et al.* (2011) found that that the described high throughput, nondestructive phenotyping platform based on LC and SR sensors has a great potential for early biomass determination in field trials of maize and other row-crops.

3   AUTOMATIC PHENOTYPING PLATFORMS

Plant phenotyping platforms are devices with the ability to automatically generate and collect data on the phenotype of plants, therefore, simplifying tasks that would otherwise be conducted manually (Pereyra-Irujo *et al.*, 2012). Automatic greenhouse systems for high throughput plant phenotyping allow the nondestructive screening of plants over a period of time by means of image acquisition techniques. During such screening, different images of each plant are recorded and must be analyzed by applying sophisticated image analysis algorithms (Hartmann *et al.*, 2011; Prasanna *et al.*, 2012).

To improve plant breeding, numerous experiments for large plant populations grown under strictly controlled environmental conditions (such as water availability, continuous lighting, and temperature conditions) are conducted. To determine the performance and the tolerance to different biotic and abiotic environmental conditions (e.g., quantification of the sensitivity to drying stresses), phenotypes should be analyzed noninvasively by imaging throughout their growth cycle. For this purpose, various fully automatic high throughput plant growth and phenotyping platforms have been developed. A technology called PHENOPSIS developed by Optimalog is used by the French National Institute for Agricultural Research (INRA) for *Arabidopsis thaliana*. The Research Center Jülich analyzes phenotypes of different plant species with the in-house GROWSCREEN system. Both techniques use a camera that is moved over the plants. The company

CropDesign developed the TraitMill™ platform, a high throughput technology that enables large-scale transgenesis and plant evaluation of *O. sativa*. In addition, the Plant Accelerator1 in Australia (http://www.plantaccelerator.org.au), CropDesign in Belgium (http://www.cropdesign.com), and the Leibniz Institute of Plant Genetics and Crop Plant Research in Germany (http://www.ipk-gatersleben.de/Internet) also have been coming on line to help increase the speed at which plants can be phenotyped (Finkle, 2009; Roy, Tucker, and Tester, 2011).

Among existing plant phenotyping platforms, PHENOPSIS and GROWSCREEN provide 2D image-processing-based semiautomated solutions for leaf phenotyping (leaf width, length, area, and perimeter) and root data monitoring (number of roots, root area, and growth rate). LAMINA, another 2D-image-based tool for leaf shape and size probing, proposes a leaf analysis for various plant species (Paproki *et al.*, 2012). Recent image-processing solutions, such as TraitMill and HTPheno, provide a more general plant analysis and measure information such as plant height, width, center of gravity, projected area, and biovolume and provide colorimetric analysis (e.g., greenness differences among plants).

Example of this type of platform is HTPheno for high throughput plant phenotyping (Figure 11) (Hartmann *et al.*, 2011). HTPheno is implemented as a plug-in for ImageJ, an open-source image-processing software. It provides the possibility to analyze color images of plants that are taken in two different views (top view and side view) during a screening. Within the analysis, different phenotypical parameters for each plant such as height, width, and projected shoot area of the plants are calculated for the duration of the screening.

ImageJ plug in HTPheno open-source pipeline to handle image data from plants. It has been developed in a modular way to allow the analysis of images from different phenotyping sources. HTPheno has been tested with the LemnaTec system, as other phenotyping platforms were not available to the authors. However, HTPheno is a generalized image analysis pipeline that can be adapted by the user to other phenotyping systems and therefore be used by a wide community of plant scientists. It is also possible to record the images independent from automated phenotyping systems that acquire all images of the individual plants.
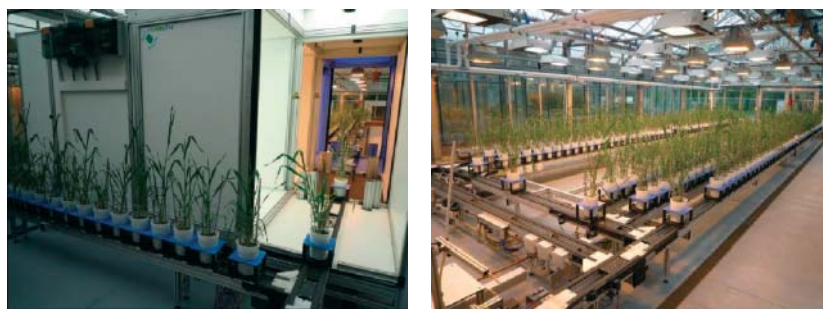
Figure 11   View of the automatic greenhouse system. Image acquisition device for images in visible, near-infrared, and ultraviolet spectra (left) and greenhouse device consisting of a conveyor belt system carrying 312 barley plants (right) (Hartmann *et al*., 2011). (*Source*: A. Hartmann, *et al*. (2011). Reproduced from BioMed Central Ltd.)

Such a manual image acquisition consists of a commercially available standard camera with a camera tripod that is positioned in some distance to the plant to reduce perspective distortion. Depending on the lens of the camera and its resolution, the perspective distortion could also be reduced by a preprocessing step in the image analysis pipeline. The plant should be arranged on a table that is in front of a unicolor wall. Light blue showed a good color choice as it is easily separable from colors in the plants. For an adequate illumination, it is recommended to install a light source above and on both sides of the plant. This setup enables the user to record images from different plants manually. Images from such a manual and low cost system can also be the source for computation of phenotypic parameters by HTPheno. To enable a versatile applicability of HTPheno to different high throughput phenotyping setups, Hartmann *et al*. (2011) decided to follow a modular approach: several configuration files allow the adjustment of the plug-in to the users' needs. With the HTPheno plug-in, it is possible to retrieve single images or a series of images from the local file system and to automatically analyze the colored images via color segmentation. The phenotypic data is determined on the basis of the segmented plant: for the side view, it is plant height, width, and projected shoot area; for the top view, it is *x*-extent, *y*-extent, projected shoot area, and the diameter of the plant. Images are finally illustrated in an image stack and the phenotypic data are composed in a result table that can be exported into various spreadsheet applications to derive meaningful diagrams.

Phenotyping technologies still have high initial costs, and consequently phenotyping platforms are not becoming widely accessible to potential users,

that is, researchers and breeders. Evidence of this limitation is clear in the fact that most platforms are currently located in large phenotyping facilities, usually one or few in each country (Kolukisaoglu and Thurow, 2010; Pereyra-Irujo *et al*., 2012).

4   CONCLUSION

High throughput phenotyping for medicinal plants will extend the standard approach by growing, measuring, and analyzing temporally thousands of medicinal plants. To date, the researches about high throughput plant phenotyping analysis focus on model plants or crop plants, few are conducted with medicinal plants. However, those literatures provide a foundation for medicinal plants high throughput phenotyping analysis.

The use of HTPPs reduces the work force and expenditure required for the determination of relevant traits. To date, the literature is distinctly dominated by 2D image-processing techniques for high throughput phenotyping of plants. The major limitation of these 2D solutions is the loss of crucial spatial and volumetric information (e.g., thickness, bending, rolling, and orientation) when transposing available data from 3D to 2D. The recent introduction of new tools for plant analysis based on explicit 3D reconstructions promises to increase potential of high throughput studies in terms of accuracy and exhaustiveness of the measured features, but available 3D solutions are currently focused on a specific organ (e.g., leaves or roots), tailored to a particular image acquisition system, and tend to be qualitative (or applied) rather than providing quantitative information and estimates of accuracy. Hence, a clear need

exists for a more generalized plant analysis based on increasingly explicit 3D models and in which the reliability of the measurements is questioned and quantitatively assessed (Paproki *et al.*, 2012).

In summary, HTPPs are in their infancy, and will succeed when high throughput platforms for specific traits of medicinal plants become reliable in accuracy and capacity and become affordable.

ACKNOWLEDGMENT

REFERENCES

Boumaza, R., Huché-Thélier, L., Demotes-Mainard, S., *et al.* (2010) *Food Qual. Prefer*, **21**, 987–997.

Cabrera-Bosquet, L., Crossa, J., von Zitzewitz, J., *et al.* (2012) *J. Integr. Plant Biol.*, **54**, 312–320.

Chéné, Y., Rousseau, D., Lucidarme, P., *et al.* (2012) *Comp. Electr. Agr.*, **82**, 122–127.

Clark, R., Maccurdy, R., Jung, J., *et al.* (2011) *Plant Physiol.*, **156**, 455–465.

de Dorlodot, S., Forster, B., Pagès, L., *et al.* (2007) *Trends Plant Sci.*, **12**, 474–481.

Fang, S. Q., Yan, X. L. and Liao, H. (2009) *Plant J.*, **60**, 1096–1108.

Finkle, E. (2009) *Science*, **325**, 380–381.

Furbank, R. T. and Tester, M. (2011) *Trends Plant Sci.*, **16**, 635–644.

Gonzalez-Sanchez, T. and Puig, D. (2011) *Electron Lett.*, **53**, 697–698.

Harbinson, J., Prinzenberg, A. E., Kruijer, W., *et al.* (2012) *Curr. Opin. Biotech.*, **23**, 221–226.

Hartmann, A., Czauderna, T., Hoffmann, R., *et al.* (2011) *BMC Bioinformat.*, **12**, 148–156.

Kolukisaoglu, U. and Thurow, K. (2010) *Plant Sci.*, **178**, 476–484.

Montes, J. M., Melchinger, A. E. and Reif, J. C. (2007) *Trends Plant Sci.*, **12**, 433–436.

Montes, J. M., Technow, F., Dhillon, B. S., *et al.* (2011) *Field Crop Res.*, **121**, 268–273.

Munns, R., James, R. A., Sirault, X. R. R., *et al.* (2010) *J. Exp. Bot.*, **61**, 3499–3507.

Paproki, A., Sirault, X., Berry, S., *et al.* (2012) *BMC Plant Biol.*, **12**, 63–75.

Pereyra-Irujo, G. A., Gasco, E. D., Peirone, L. S., *et al.* (2012) *Funct. Plant Biol.*, **39**, 905–913.

Pieruschka, R. and Poorter, H. (2012) *Funct. Plant Biol.*, **39**, 813–820.

Prasanna, B. M., Araus, J. L., Crossa, J. *et al.* (2012) High-throughput and precision phenotyping in cereal breeding programs, *Cereal Genomics*, 2nd edn, Springer.

Roy, S. J., Tucker, E. J. and Tester, M. (2011) *Curr. Opin. Plant Biol.*, **14**, 232–239.

Sato, S., Isobe, S. and Tabata, S. (2010) *Curr. Opin. Plant Biol.*, **13**, 146–152.

Sirault, X. R. R., James, R. A. and Furbank, R. T. (2009) *Funct. Plant Biol.*, **36**, 970–977.

Identification of Medicinal Plants Using DNA Barcoding

Xiaohui Pang and Shilin Chen

The National Engineering Laboratory for Breeding of Endangered Medicinal Materials, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, PR China

1 INTRODUCTION

Medicinal plants have been used in the health care system in China for thousands of years (Gong and Sucher, 2002; Marian *et al*., 2006). The correct identification of these plants is a prerequisite for their safe application. Traditionally, medicinal plants are discriminated mainly using morphological characteristics by experienced experts. However, an expert may take many years to identify these plants proficiently, and only a few taxonomists have accurately distinguished more than 1000 species (Hebert *et al*., 2003). Moreover, the number of trained experts has been decreasing along with the increasing workload on identification in China. Therefore, the traditional method is insufficient to meet the demands of identification in hospitals, drug stores, businesses, and customs. The number of adulterants and substitutes on the market has been increasing recently, which leads to variation in the quality of medicinal plants (Shaw *et al*., 1997; Tomlinson *et al*., 2000). Thus, the identification of medicinal plants is a great challenge, which poses safety hazards to both the traditional Chinese medicine industry and consumers. Hence, a rapid and accurate method for authentication of these plant species is needed.

DNA barcoding is a novel method for species identification that has potential to resolve the above-mentioned difficulties (Hebert *et al*., 2003). This method is based on sequence differences within short and standardized DNA regions, with a primary goal of authenticating known species as well as discovering new ones (Hebert *et al*., 2003; Mitchell, 2008; Moritz and Cicero, 2004). DNA barcoding is an effective and inexpensive technique that could promote the work of specialists and make the authentication of species accessible to non-specialists (Chase *et al*., 2005; Mallet and Willmott, 2003; Schindel and Miller, 2005; Taberlet *et al*., 2007). Furthermore, it is independent of both morphological characteristics and expert experiences, which makes the results more rapid, subjective, and accurate. Consequently, DNA barcoding has rapidly become the focus of research, and it has been recognized as a powerful tool for species identification (Hajibabaei *et al*., 2005; Hebert and Gregory, 2005).

The determination of universal DNA barcodes is the important precondition of the application of DNA barcoding technique. A portion of the mitochondrial cytochrome *c* oxidase 1 (*CO1*) gene is accepted as a universal barcode for identifying animal lineages (Elias *et al*., 2007; Hebert *et al*., 2004; Tavares and Baker, 2008; Ward *et al*., 2005). Consensus is lacking regarding plant DNA barcode, which has limited the application of DNA barcoding for a long time (Kress and Erickson, 2007; Lahaye *et al*., 2008;

Sass *et al*., 2007; Song *et al*., 2009). Recently, the Consortium for the Barcode of Life Plant Working Group (2009) recommended the combination of *matK* + *rbcL* as the standard plant barcode. However, the group conceded that the barcode is far from perfect because of the low identification rate. Chen *et al*. (2010) compared seven DNA regions from medicinal plants, and they proposed the ITS2 (second internal transcriber spacer) locus as a universal DNA barcode for medicinal plants. Consequently, the effectiveness of ITS2 for identifying species has been tested in a wide range of plant taxa (China Plant BOL Group, 2011; Gao *et al*., 2010a,b; Luo *et al*., 2010; Pang *et al*., 2010, 2011; Pang, Luo, and Sun, 2012). At present, the ITS2 region is a useful barcode for identifying medicinal plants. In this study, we introduce the DNA barcoding technique and determine its usage in discriminating medicinal plants, as well as its advantages and limitations.

2　USING DNA BARCODING IN IDENTIFYING MEDICINAL PLANTS

This study describes a protocol for using the ITS2 barcode to identify medicinal plants. The protocol provides an example of the use of DNA barcoding. The specific steps are as follows.

2.1　DNA Isolation, Amplification, and Sequencing

Leaf tissues were rubbed in liquid nitrogen or in a FastPrep bead mill (Retsch MM400, Germany) for 1 min at a frequency of 30 times per second. The total genomic DNA was isolated using the Plant Genomic DNA Kit (Tiangen Biotech Co., China) according to the manufacturer's instructions. Universal primers for the ITS2 barcode and general polymerase chain reaction (PCR) condition were used as in previous studies (Chen *et al*., 2010; Chiou *et al*., 2007; Sass *et al*., 2007). The sequence of the forward primer was 5′-ATGCGATACTTGGTGTGAAT-3′, and the sequence of the reverse primer was 5′-GACGCTTCTCCAGACTACAAT-3′. The reaction conditions were as follows: 94°C for 5 min; 94°C for 30 s, 56°C for 30 s, and 72°C for 45 s of 40 cycles; and 72°C for 10 min. PCR amplification was performed in 25 μL reaction mixtures containing approximately 30 ng of genomic DNA template, 1 × PCR buffer without magnesium chloride ($MgCl_2$), 2.0 mM $MgCl_2$, 0.2 mM of each deoxyribonucleotide triphosphate, 0.1 μM of each primer (synthesized by Sangon Co., China), and 1.0 U *Taq* DNA Polymerase (Biocolor Bio-Science & Technology Co., China), using a Peltier Thermal Cycler PTC0200 (BioRad Lab, Inc., USA). Purified PCR products were sequenced in both directions with the primers used for PCR amplification on a 3730XL sequencer (Applied Biosystems, USA).

2.2　Sequence Assembly and Quality Control

Sequence assembly and quality control were performed as in a previous study (Chen *et al*., 2010). The original forward and reverse sequences were assembled using a CodonCode Aligner 3.0 (CodonCode Co., USA) to estimate the quality of the generated sequence traces. Base calling was carried out using the Phred program (version no. 0.020425.c). The quality values (QVs) were defined for three levels: low quality (0–19 QV), medium quality (20–30 QV), and high quality (>30 QV). The sequences showing >2 bases with a QV below 20 QV in a 20-base window were trimmed. The forward and reverse reads had a minimum length of 100 bp, a minimum average QV of 30, and post-trim lengths of >50% of the original read length. In addition, the assembled contig showed a minimum average QV score of 40% and >50% overlap in the alignment of the forward and reverse reads. All complete ITS2 sequences were retrieved according to Keller *et al*. (2009).

2.3　Sequence Alignment, Genetic Analysis, and Species Identification

The ITS2 sequences were aligned using Clustal W, and Kimura 2-Parameter distances were calculated using PAUP4b10 (Florida State University, USA). Three parameters, namely, all interspecific distance, theta prime, and minimum interspecific distance, were used to evaluate interspecific differences, and three other parameters, namely, all intraspecific difference, theta, and coalescent depth, were used to assess intraspecific variations (Chen *et al*., 2010; Lahaye *et al*., 2008; Meier, Zhang, and Ali,

2008; Meyer and Paulay, 2005). The intraspecific differences and interspecific variations of congeneric species were calculated using a custom Perl script, as described in previous studies (Chen *et al*., 2010; Meier, Zhang, and Ali, 2008; Meyer and Paulay, 2005). All ITS2 sequences of medicinal plants were used as query sequences. Two methods of species identification, namely, the basic local alignment search tool 1 and the nearest distance method, were performed as described in a previous study (Ross, Murugan, and Li, 2008).

3 ADVANTAGES AND LIMITATIONS OF DNA BARCODING

The DNA barcode project establishes a standard system for the authentication of all eukaryotic species (Hebert *et al*., 2003; Miller, 2007). DNA barcoding is cost- and time-effective in that it achieves accurate and fast identifications in a wide range of species based on a public reference database, which results in automated species identification and is very important in distinguishing species in campaigns with large samples and in large surveys with goals of detecting unknown species (Ball and Armstrong, 2006; Rusch *et al*., 2007). Morphological data gathering is laborious and time-consuming, and in some instances, the data can be completely confusing or impossible to obtain (Evans, Wortley, and Mann, 2007; Huang *et al*., 2007; Litaker *et al*., 2007). By contrast, acquisition of molecular data is rapid. Moreover, DNA barcoding attains absolute superiority in certain cases, such as in determining species without clear morphological traits, matching adults with immature samples, and identifying damaged organisms or fragments of food and stomach extracts. The DNA barcode and primers are universal, which makes DNA barcoding a method of choice over molecular-based methods. Despite the advantages of DNA barcoding, some limitations also exist, in which the value of their mention is an important prerequisite for DNA barcoding to extract DNA successfully. DNA extraction is easy for fresh materials. However, it may be difficult for some materials with degraded DNA. Moreover, sampling shortages across taxa could sometimes lead to ambiguous identification, which then limits DNA barcoding. Therefore, an important consideration is in choosing sufficient individuals that cover the majority of the existing diversity to represent each taxon during the reference database construction phase.

4 CONCLUSION

A rapid and accurate method to identify herbal materials is very important to ensure their safe usage. Accordingly, this goal can be best achieved by DNA barcoding. At present, DNA barcoding is widely recognized and applied in the field of identification of medicinal plants. The ability of the ITS2 barcode in identification of medicinal plant species has been tested in a wide range of taxa. We believe that DNA barcoding will have increasing applications in identifying medicinal plants in the near future, which will greatly improve their safe use.

5 RELATED ARTICLES

Selection, Identification, and Collection of Plants; Identification and Characterization of Hydroxycinnamates of Six *Galium* Species from the Rubiaceae Family; Transcriptome Analysis of Medicinal Plants with Next-Generation Sequencing Technologies; High Throughput Screening of Vegetal Natural Substances; Quality Assessment of Herbal Drugs and Medicinal Plant Products; Identification and Characterization of Trimeric Proanthocyanidins of Two Members of the *Rhododendron* Genus (*Ericaceae*) by Liquid Chromatography Multi-Stage Mass Spectrometry.

REFERENCES

Ball, S. L. and Armstrong, K. F. (2006) *Can J. For. Res*, **36**, 337–350.

CBOL Plant Working Group (2009) *Proc. Natl. Acad. Sci. U.S. A.*, **106**, 12794–12797.

Chase, M. W., Salamin, N., Wilkinson, M., *et al*. (2005) *Phil. Trans. R Soc. B*, **360**, 1889–1895.

Chen, S. L., Yao, H., Han, J. P., *et al*. (2010) *PLoS ONE*, **5**, e8613.

China Plant BOL Group (2011) *Proc. Natl. Acad. Sci. U. S. A.*, **108**, 19641–19646.

Chiou, S. J., Yen, J. H., Fang, C. L., *et al*. (2007) *Planta Med.*, **73**, 1421–1426.

Elias, M., Hill, R. I., Willmott, K. R., *et al*. (2007) *Proc. R. Soc. Lond. B Biol. Sci.*, **274**, 2881–2889.

Evans, K. M., Wortley, A. H., and Mann, D. G. (2007) *Protist*, **158**, 349–364.

Gao, T., Yao, H., Song, J. Y., *et al*. (2010a) *J. Ethnopharmacol.*, **130**, 116–121.

Gao, T., Yao, H., Song, J. Y., *et al*. (2010b) *BMC Evol. Biol.*, **10**, 324.

Gong, X., and Sucher, N. J. (2002) *Phytomedicine*, **9**, 478–484.

Hajibabaei, M., deWaard, J. R., Ivanova, N. V., *et al*. (2005) *Phil. Trans. R Soc. B*, **360**, 1959–1967.

Hebert, P. D. N., and Gregory, T. R. (2005) *Syst. Biol.*, **54**, 852–859.

Hebert, P. D. N., Cywinska, A., Ball, S. L., and deWaard, J. R. (2003) *Proc. Royal Soc. B Biol. Sci.*, **270**, 313–321.

Hebert, P. D. N., Penton, E. H., Burns, J. M., *et al*. (2004) *Proc. Natl. Acad. Sci. U. S. A.*, **101**, 14812–14817.

Huang, J., Qin, X. Q., Sun, Z. J., *et al*. (2007) *Pedobiologia*, **51**, 301–309.

Keller, A., Schleicher, T., Schultz, J., *et al*. (2009) *Gene*, **430**, 50–57.

Kress, W. J., and Erickson, D. L. (2007) *PLoS ONE*, **2**, e508.

Lahaye, R., van der Bank, M., Bogarin, D., *et al*. (2008) *Proc. Natl. Acad. Sci. U. S. A.*, **105**, 2923–2928.

Litaker, R. W., Vandersea, M. W., Kibler, S. R., *et al*. (2007) *J. Phycol.*, **43**, 344–355.

Luo, K., Chen, S. L., Chen, K. L., *et al*. (2010) *Sci. China*, **53**, 701–708.

Mallet, J., and Willmott, K. (2003) *Trends Ecol. Evol.*, **18**, 57–59.

Marian, F., Widmer, M., Herren, S., *et al*. (2006) *Forsch Komplementärmed*, **13**, 70–77.

Meier, R., Zhang, G. Y., and Ali, F. (2008) *Syst. Biol.*, **57**, 809–813.

Meyer, C. P., and Paulay, G. (2005) *PLoS Biol.*, **3**, 2229–2238.

Miller, S. E. (2007) *Proc. Natl. Acad. Sci. U. S. A.*, **104**, 4775–4776.

Mitchell, A. (2008) *Aust. J. Entomol.*, **47**, 169–173.

Moritz, C., and Cicero, C. (2004) *Public Library Sci. Biol.*, **2**, 279–354.

Pang, X. H., Song, J. Y., Zhu, Y. J., *et al*. (2010) *Planta Med.*, **76**, 1784–1786.

Pang, X. H., Song, J. Y., Zhu, Y. J., *et al*. (2011) *Cladistics*, **27**, 165–170.

Pang, X., Luo, H., and Sun, C. (2012) *Plant Biol.*, **14**, 839–844.

Ross, H. A., Murugan, S., and Li, W. L. S. (2008) *Syst. Biol.*, **57**, 216–230.

Rusch, D. B., Halpern, A. L., Sutton, G., *et al*. (2007) *Plos Biol.*, **5**, 398–431.

Sass, C., Little, D. P., Stevenson, D. W., *et al*. (2007) *PLoS ONE*, **2**, e1154.

Schindel, D. E., and Miller, S. E. (2005) *Nature*, **435**, 17.

Shaw, P. C., Ngan, F. N., But, P. P. H., *et al*. (1997) *J. Food Drug Anal.*, **5**, 273–284.

Song, J. Y., Yao, H., Li, Y., *et al*. (2009) *J. Ethnopharmacol.*, **124**, 434–439.

Taberlet, P., Coissac, E., Pompanon, F., *et al*. (2007) *Nucleic Acids Res.*, **35**, e14.

Tavares, E. S., and Baker, A. J. (2008) *BMC Evol. Biol.*, **8**, 81.

Tomlinson, B., Chan, T. Y., Chan, J. C., *et al*. (2000) *J. Clin. Pharmacol.*, **40**, 451–456.

Ward, R. D., Zemlak, T. S., Innes, B. H., *et al*. (2005) *Proc. Natl. Acad. Sci. U. S. A.*, **360**, 1847–1857.

Transcriptome Analysis of Medicinal Plants with Next-Generation Sequencing Technologies

Shilin Chen and Hongmei Luo

The National Engineering Laboratory for Breeding of Endangered Medicinal Materials, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, PR China

1   INTRODUCTION

Plants, especially medicinal plants, are most promising sources for the production of valuable natural products, and many are also the major sources of natural medicines that belong to secondary metabolites possessing pharmacological properties (Newman and Cragg, 2007). However, the transcriptomes/genomes of most of the medicinal plants are not available (De Luca and Laflamme, 2001; Oksman-Caldentey and Saito, 2005). Transcriptome analysis is a promising approach to illustrate the biological processes, such as the biosynthesis of bioactive compounds, growth and development, and the genetic diversity of medicinal plants (Jennewein *et al*., 2004; Lin *et al*., 2011; Murata *et al*., 2008). In addition, the identification of key genes and genetic markers related to biosynthesis of the secondary metabolites will facilitate the fast-track breeding of medicinal plants (Graham *et al*., 2010).

Transcriptomics study, often used interchangeably with "gene expression profiling," refers to the *transcriptome analysis of a cell*, *cell type*, or *an organism*, which is growing/developing at specific stages or under different conditions. Transcriptome includes the complete set of messenger ribonucleic acid (mRNA) and/or noncoding ribonucleic

acid (ncRNA) transcripts generated by a particular cell, cell type, or organism (Morozova, Hirst, and Marra, 2009). The earlier attempts to understand large-scale transcriptomes of tissues or organisms benefit from the application of DNA microarray and expressed sequence tag (EST) approaches. DNA microarray techniques can simultaneously analyze the expression levels of thousands of genes and provide reliable data for comparing the expression ratio of the same transcript under different conditions. However, the ability of DNA microarray methods to detect novel transcripts and the ability to study the coding sequence of detected transcripts were very limited. EST approaches have provided a rapid and cost-effective alternative route for the purposes of gene discovery. ESTs, generated from cDNA libraries, represent portions of the species transcriptome. Transcriptome sequencing studies have evolved from detecting the sequences of individual cDNA clones (Stone *et al*., 1985) to comprehensive attempts to construct cDNA libraries. There are several distinct types of cDNA libraries, including full-length cDNA sequence libraries, normalized libraries, subtracted libraries, and suppressive subtractive hybridization (SSH) libraries. The construction systems of these cDNA libraries have been introduced in detail (Parkinson, 2009).

Table 1   Comparison of several NGS technologies used for transcriptome sequencing applications[a].

Sequencing platforms	Roche/454 GS FLX + instrument	Illumina/Miseq	ABI/SOLiD5500xl	PacBio
Libraries	Fragment	Fragment	Fragment	Fragment
	Mate-pair	Mate-pair	Mate-pair	
Template preparation	Emulsion PCR	Bridge PCR	Emulsion PCR	Single molecule
NGS chemistry	Pyrosequencing	Sequencing-by-synthesis with reversible terminators	Sequencing-by-ligation	Real-time
Read length (bases)	700–800	150	35 or 50	~1000
Run time	1 day	20.7–24 h	8–10 days	<24 h
Sequencing throughput (per run)	600–700 Mb	3.7–4.6 Gb	300 Gb	—

[a] Data tabulated on April 1, 2012.

Despite the power of the EST approach to detect the expression of thousands of genes simultaneously, the cloning-based amplification and Sanger sequencing for ESTs are lengthy, labor-intensive, and costly processes (Hall, 2007). Next-generation sequencing (NGS) technologies circumvent the cloning requirements by taking advantage of a highly efficient *in vitro* DNA amplification method.

NGS technologies, including second-generation sequencing (SGS) and third-generation sequencing (TGS) platforms, have revolutionized transcriptomics by providing opportunities for deep and multidimensional examinations of cellular transcriptomes in which high throughput transcript data are obtained at a single-base resolution (Morozova, Hirst, and Marra, 2009). The comparison of NGS technologies used for transcriptome sequencing is depicted in Table 1. Second-generation sequencers include 454/Roche, Solexa/Illumina, and SOLiD/ABI (supported oligonucleotide ligation and detection/Applied Biosystems). Third-generation sequencers, including Helicos from Heliscope, PacBio from Pacific Biosciences, and Nanopore from Oxford Nanopore Technologies (ONT), can sequence single-molecule templates directly with no PCR (polymerase chain reaction) amplification. The application of NGS technologies has accelerated the investigations on transcript profiling in various tissues with different physiological and pathological conditions, leading to the mining of novel genes and biological processes. NGS has sparked a revolution in gene expression analysis, providing the capability to characterize transcriptomes from virtually any organism or cell type with unprecedented resolution on a massive scale. At present, the transcriptome sequences based on the NGS technique provide a vast amount of data for the identification of biosynthetic

genes. In this chapter, we focus on the application of NGS technologies for transcriptome analysis and discuss the progress in the study of transcriptomes of medicinal plants.

2   SECOND-GENERATION SEQUENCING TECHNOLOGIES

SGS platforms, which produce large amounts (typically millions) of short DNA sequence reads from 454/Roche, Solexa/Illumina, and SOLiD/ABI, have provided unprecedented opportunities for high throughput transcriptome sequencing.

2.1   454 Sequencing

The pyrosequencing approach, a sequencing-by-synthesis technique measuring the release of inorganic pyrophosphate (PP_i) by chemiluminescence, is the first alternative to Sanger sequencing used in SGS technologies (Nyren, Pettersson, and Uhlen, 1993; Ronaghi *et al*., 1996). The first successful pyrosequencing system was developed by 454 Life Sciences and commercialized by Roche as the Genome Sequencer (GS) 20, which can produce over 20 million base pairs in just over 4 h (Margulies *et al*., 2005). The GS20 was replaced during 2007 by the GS FLX model, capable of producing over 100 million base pairs of sequence in a similar amount of time. After the introduction of the Titanium chemistry in 2008, this technology has advanced to produce more than 400 Mb, and the DNA sequence reads with an average length of 300–500 bp. Most recently, the GS Junior System was released to the market, together with the reagents of GS Junior Titanium. Thus, the length of reads can reach up to 600–800 bp,

which is similar to the sequence length generated from ABI3730 sequencer.

The 454 technology applied emulsion PCR for DNA amplification *in vitro* (Tawfik and Griffiths, 1998). Emulsion PCR isolates individual DNA molecules along with primer-coated beads in aqueous droplets within an oil phase. In emulsion PCR, individual DNA fragment-carrying streptavidin beads, generated from shearing the DNA and attaching the fragments to the beads using adapters, are populated and captured into separate emulsion droplets. The droplets act as individual amplification micro reactors, which contain beads, template, PCR components, and primers, producing $\sim 10^7$ clonal copies of a unique DNA template per bead (Margulies *et al*., 2005). Each template-containing bead is subsequently transferred into a well of a picotiter plate and the pyrosequencing reaction is carried out. The four nucleotides are added to the sequencing reaction one at a time and the addition of the correct nucleotide is detectable by light produced by a chemiluminescent enzyme present in the reaction mixture. As the amount of light produced is proportional to the number of incorporated nucleotides, the pyrosequencing approach is prone to errors in the detection of homopolymers resulting from incorrect estimation of the length of the homopolymeric sequence stretches.

2.2 SOLiD Sequencing

The ABI SOLiD system was available in 2006. The earlier version of SOLiD 3 was capable of producing 50 gigabases (Gb) of data per run in 2009. The newer SOLiD 4, which was released subsequently in 2010, can generate about 100 Gb data. At the end of 2010, the newest SOLiD5500xl was available and can produce 30–45 Gb of sequence data per 1-day run. The SOLiD system can be used for the detection of gene expression and presumably used for resequencing where comparison to a reference and high read redundancy enables the identification of a sequence error.

In contrast to the sequencing-by-synthesis approach discussed earlier, the SOLiD system uses a sequencing-by-ligation approach, first published by the Church laboratories as the "polony sequencing techniques" (Shendure *et al*., 2005). The DNA amplification for analysis on the SOLiD instrument starts with an emulsion PCR single-molecule amplification step. After the amplification, the products are subsequently transferred onto a glass surface where the sequence is inferred indirectly by sequential rounds of hybridization and ligation events. The SOLiD system used 16 dinucleotide combinations labeled by four different fluorescent dyes such that one dye labels four dinucleotides. Using the four-dye encoding scheme, each base is effectively probed twice and determined by analyzing the fluorescence generated from the two successive ligation reactions. The two-base encoding scheme allows for the distinction of a sequencing error from a true sequence polymorphism.

2.3 Solexa Sequencing

The first Solexa sequencer, the Genome Analyzer, was launched in 2006 and can produce 1 Gb in a single run. Solexa was acquired by Illumina in early 2007. The Genome Analyzer, Hiseq2000, Hiseq1000, HiScanSQ, and Miseq instruments are available. In comparison to the earlier system of Genome Analyzer, the Hiseq and Miseq can produce 30 and 60 billion reads per run, respectively. The internal advantages of library automation, low input DNA, capture hybridization multiplexing, and analysis pipeline improvements including integration of variant calling and annotation tools contribute to the higher capacity of the Illumina HiSeq platform. MiSeq integrates all the steps of the Illumina sequencing process, including cluster generation and paired-end (PE) sequencing, and reads into a fully automated "hands-off" workflow. The workflow includes the steps of primary and secondary data analysis, including variant calling and *de novo* assembly. All these advantages of MiSeq contribute to the generation of data from 2×250-bp PE reads with <1% sequence error, and it is anticipated that 2×400-bp PE reads will be available this year (Mason and Elemento, 2012). After unceasing refinements and optimization, the newest 2011 generation of Illumina SBS technology-based instruments produces up to 600 Gb per run.

The Illumina/Solexa system also uses sequencing-by-synthesis approach similar to that used in 454 sequencing technology (Bentley, 2006). The sequencing templates are immobilized on a solid surface, where the solid-phase amplification occurs.

After the single-stranded DNA fragments are attached to a flow cell, the solid-phase bridge amplification of single-molecule DNA templates is performed and the clusters of identical copies of each DNA molecule are generated. The templates are sequenced in a massively parallel manner using four fluorescently labeled nucleotide analogs as reversible sequencing terminators in the process. The template sequence of each cluster is deduced and revealed by reading off the color of its fluorescent label at each successive nucleotide addition step. The shorter reads generated from Solexa sequencer lead the system to not resolve the short sequence repeats (Bentley, 2006).

3   TRANSITIONING FROM SGS TO THIRD-GENERATION SEQUENCING (TGS)

Ion Torrent's semiconductor sequencer sits between the SGS and the TGS categories. LIFE Technology's Personal Genome Machine (PGM) promises fast and low cost (<$1000 per run) sequencing solutions. The Ion Torrent Personal Genome Machine (ION Torrent PGM) is a powerful bench-top sequencing platform that has rapidly evolved and takes approximately 2 h. Ion Torrent platform has obtained a variety of significant developments, including the development of new capabilities, the automation, streamline, and optimization of the ion sequencing process from sample preparation to data generation, the test and evaluation of new protocols (such as longer reads, Ion Express, One Touch, and enrichment automation), the application of the technology to a variety of projects, and others.

Ion Torrent's sequencing instrument is the state-of-the-art semiconductor technology that is employed to create a high density array of micro-machined wells (Schadt, Turner, and Kasarskis, 2010b). Nucleotides flow one-by-one across a well, followed by the measure of incorporation of a nucleotide by taking advantage of the molecular changes that occur. Ion Torrent is capable of directly translating chemical signals into digital information. No light detection (such as costly and sensitive optics) is involved in Ion Torrent, which differs from other platforms. Instead, Ion Torrent detects the change in pH produced by the release of a hydrogen ion when the new nucleotide incorporation occurs. An accurate template-to-sphere

ratio (also known as *dilution factor*) is essential for a successful sequencing run. Therefore, the major advantages of Ion Torrent are the low price of the machine and reagents and the rapid sequencing technology, leading to the production of sequencing data with unprecedented speed, scalability, and low cost. However, this technology, like all current SGS technologies, requires PCR amplification of the DNA template, as well as termination events halting sequencing after the incorporation of each nucleotide. In addition, the read length is limited to that of current SGS platforms.

4   THIRD-GENERATION SEQUENCING TECHNOLOGIES

In contrast to the SGS technologies, the TGS performs single-molecule DNA sequencing to a further level of scale and dramatically reduces the costs. The preparation of sample for single-molecule sequencing does not involve adding adaptors to cDNA and is lacking cloning, amplification, or ligation. In particular, TGS technologies offer advantages over current sequencing technologies including faster turnaround time, longer read length to enhance *de novo* assembly, small amounts of starting materials, and higher throughput (Schadt, Turner, and Kasarskis, 2010b). The throughput of single-molecule sequencing has improved about 10-fold per year over the last few years (Eid *et al*., 2009; Greenleaf and Block, 2006; Harris *et al*., 2008).

4.1   The Heliscope Single Molecule Sequencer

The first demonstration of single-molecule sequencing was published in 2003 (Braslavsky *et al*., 2003), and the first commercial release of a single-molecule sequencing instrument was the Heliscope Single Molecule Sequencer (Helicos Biosciences) in 2008 (Bowers *et al*., 2009; Harris *et al*., 2008; Lipson *et al*., 2009; Tessler, Reifenberger, and Mitra, 2009). The Heliscope system uses the nucleotide analogs, referred to as *virtual terminator nucleotides*, which reduce the processivity of DNA polymerase as terminators in the sequencing processes. The high densities of unamplified single-molecule DNA templates are captured on the flow cell surface and can be extended asynchronously, aligned with Cy3

labels attached at both ends of the DNA molecule. After the addition of the Cy5-labeled nucleotides to the reaction at a time, the detection of incorporated nucleotides is captured by TIRF (total internal reflection fluorescence). Therefore, the Heliscope instrument works by imaging individual DNA molecules. Several significant advantages of Heliscope over SGS technologies include no PCR requirement for sequencing, the digital measurement of homopolymer sequences, and a 1000-fold improvement in parallelism (Pushkarev, Neff, and Quake, 2009). However, owing to the single-molecule nature of this system, the raw read error rates are at or over 5%, although the consensus read accuracy is generally >99%, resulting from the high fold coverage generated by the highly parallel nature of this technology.

4.2 The PacBio *RS*

The PacBio *RS* was commercialized on sale by the US Pacific Biosciences in April of 2011. The PacBio *RS* combines a novel single-molecule sequencing technology, based on a novel single-molecular real-time (SMRT) technology, and presents a unique opportunity for DNA sequencing with read lengths of several kilobases (Schadt, Turner, and Kasarskis, 2010b). The key strength of PacBio's SMRT sequencing technology and RS system is the long read lengths, compared to other SGS technologies (i.e., 454, Illumina and SOLiD). The maximum read length offered by 454 is about 1 kb, 150 bp by Illumina, and even shorter by SOLiD. Average read length from RS system with standard sequencing protocol is about 2 kb using version 2 chemistry in conjunction with 45-min moves, with 5% of reads longer than 4 kb. Therefore, its benefits include long read lengths and rapid turnaround times. Pacific Biosciences will soon introduce an improved sequencing chemistry that provides mean read lengths of up to 2700 bases using 75-min continuous sequencing. PacBio has been employed successfully for genome (low complexity genomic region such as extremely high GC, long repeats, rearrangement, and gene fusion) and transcriptome sequencing (Eid *et al*., 2009; Flusberg *et al*., 2010; Song *et al*., 2012).

SMRT DNA sequencing monitors nucleotide incorporation events in real time, involving imagination of the continuous incorporation of dye-labeled nucleotides during DNA synthesis, and results in a wealth of kinetic information in addition to the extraction of the primary DNA sequence. The single DNA polymerase molecules are attached to the bottom surface of individual zero-mode waveguide (ZMW) detectors. The utility of ZMW detectors can provide sequence information and leads to the resolution of observing a DNA polymerase working in real time and detecting the incorporation of a single nucleotide taken from a large pool of potential nucleotides during DNA synthesis. The SMRT sequencing system requires minimal amount of template and reagent to carry out a run. The sample preparation is simple for SMRT sequencing, consisting of fragmentation of the DNA into desired lengths and blunting the ends, following by the ligation of hairpin in adaptors and the sequencing (Travers *et al*., 2010).

Despite the significant advantages of SMRT sequencing, a number of challenges still remain. For example, the raw read error rates are also in excess of 5%, like the Helicos system. In addition, the modest throughput of SMRT sequencing will not initially match the production of SGS. Finally, the data form of SMRT differs from SGS data, leading to the requirement of a more advanced probabilistic modeling to exploit more information on the chemical and structural nature of nucleotide sequences than current sequencing technologies.

4.3 Nanopore

ONT intends to commercialize devices based on the successful implementation of "strand" nanopore sequencing in this year. Nanopore sequencing directly reads the sequence of individual DNA molecules, such as the long read data from real genomic samples with competitive accuracy (Branton *et al*., 2008). Nanopores are small holes, created either by inserting porous proteins into membranes (biological nanopores) or by creating holes in solid-state materials (synthetic nanopores). The potential of nanopores is used as part of an electronics-based technology for DNA/RNA sequencing starting with extremely small amounts of input materials. Therefore, an ion current passing through a pore can be modulated by nucleotides moving through the pore.

5 THE LIMITATIONS OF NGS TECHNOLOGIES

The NGS technologies constitute various strategies including template preparation, sequencing and imaging, and sequence assembly. The main goals of sequencing technology improvements of NGS platforms are to achieve longer read lengths, higher sequence accuracy, and greater throughput. SGS produces high accuracy and higher throughout at the cost of read length; the read lengths currently achievable by 454 sequencing technology are up to 300–700 bp, whereas the Solexa and SOLiD instruments generated shorter than 35–200 bp reads. However, the third generation of sequencing platform, such as PacBio, offers very long read length and low GC bias, but at the cost of modest throughput and low accuracy. A limitation of short-read sequence data is the difficulty in *de novo* sequence assembly. PE sequencing approaches have the potential to improve the utility of short reads for *de novo* sequencing assembly. Both ends of a fragment of defined size in PE sequencing are sequenced to provide more information about the fragment based on the PE sequencing methods.

The multistep scheme of library construction and gel-based size selection limits the sequence-ready library preparation significantly. Moreover, preparation of libraries of a number of samples is extremely time consuming. In addition, in order to generate large number of DNA molecules from templates, PCR amplification is required. Furthermore, errors in the template sequence and amplification bias can be introduced in the amplification process. Therefore, improvements in library preparation and chemistry, such as utility of a novel set of reversible terminators for reducing the high error rates in NGS and speeding cycle times, are imperative.

The massively high throughput sequence data generated from NGS technologies challenge data transfer, storage and quality control, informatics operations to align or assembly read data, and laboratory information systems (Schadt *et al.*, 2010a). The ability to analyze and assemble this data is currently limited by a lack of dedicated bioinformatics tools that are designed to cope with the *de novo* assembly and alignment of the shorter reads. This introduces unique analytical challenges and opportunities for improvement in computational analysis based on the advantages in bioinformatics.

6 APPLICATION OF NGS TECHNOLOGIES IN TRANSCRIPTOME ANALYSIS OF MEDICINAL PLANTS

NGS technologies are driving advantages in transcriptome study of medicinal plants, aiming at gene discovery, gene expression profile analysis, protein-coding gene annotation, and genetic marker identification. A unique feature of high throughput transcriptome data generated from NGS technologies is the sufficient sequencing depth and the versatility of the data, which can simultaneously be analyzed to provide insight into the gene expression level, single-nucleotide variations, and transcript rearrangement.

6.1 Discovery of Genes Associated with the Biosynthesis of Bioactive Compounds

The identification of the biosynthetic pathways of bioactive compounds, which are complex and including multiple cross-talks of molecular regulation, is the focus of the research on medicinal plants. However, the genomic/transcriptomic data of most medicinal plants are not available to date compared with that of model plants, limiting the identification and elucidation of the biosynthetic pathways of important bioactive compounds (secondary metabolites) in medicinal plants. Transcriptome analysis, in comparison to genome sequencing, is feasible and has been a key area of biological inquiry.

EST approaches, a typical transcriptome analysis method, has provided a rapid and cost-effective alternative route toward the isolation of gene-encoding specific enzymes and (or) regulatory factors involved in secondary biosynthetic pathways (Ohlrogge and Benning, 2000). The mining of ESTs (sequenced using Sanger method) from specific tissues or organs has revealed the presence of transcripts involved in the biosynthesis of many important bioactive compounds, such as artemisinin (isoprenoid), paclitaxel (diterpenoid), vinblastine, and vindoline (monoterpenoid indole alkaloid, MIA) (Jennewein *et al.*, 2004; Murata *et al.*, 2008; Teoh *et al.*, 2006). However, owing to the high cost of Sanger method (Sanger, Nicklen, and Coulson, 1977) and the complexity of the process including cDNA library construction and clone isolation, EST analysis for nonmodel

plants (including most of the medicinal plants) is not feasible.

The presence of NGS technologies provides opportunities for large-scale analysis of transcriptomes from medicinal plants with the goal of identification of key genes associated with the biosynthesis of bioactive compounds. The NGS techniques have been used to analyze the transcriptomes in the root of *Panax* (Araliaceae) plants, including *American ginseng* (Sun *et al.*, 2010), *Panax ginseng* (Chen *et al.*, 2011), and *Panax notoginseng* (Luo *et al.*, 2011). As a result of this process, a handful of the candidate CYP450 and glycosyltransferase genes involved in the biosynthesis of ginsenosides were found. The transcriptome sequencing of *Huperzia serrata* and *Phlegmariurus carinatus*, the ancient ferns, also led to the discovery of candidate CYP450 genes involved in lycopodium alkaloid biosynthesis (Luo *et al.*, 2010a). At present, the transcriptomes of a number of medicinal plants have been sequenced and analyzed using NGS technologies, such as *Glycyrrhiza uralensis* (Li *et al.*, 2010a), opium poppy (Desgagne-Penix *et al.*, 2010), *Artemisia annua* (Wang *et al.*, 2009a), *Camptotheca acuminate* (Sun *et al.*, 2011), and *Salvia miltiorrhiza* (Hua *et al.*, 2011). Genes involved in the biosynthesis of glycyrrhizin, benzylisoquinoline alkaloids, artemisinin, camptothecin, tanshinome, and salvianolic acid have been discovered from these transcriptome data. The studies on the discovery of secondary metabolite biosynthetic genes in medicinal plants by transcriptome data analysis are summarized in Table 2.

In addition to the discovery of genes associated with the biosynthesis of bioactive compounds, the novel genes have been detected in the presence of the transcriptomes of many plant species (Alagna *et al.*, 2009; Lin *et al.*, 2011; Luo *et al.*, 2010a; Novaes *et al.*, 2008; Vega-Arreguin *et al.*, 2009). For instance, one such study generated 406,044 reads from *A. annua* and identified novel transcripts in this species (Wang *et al.*, 2009a).

6.2 Gene Expression Profiling

The abundance of mRNA in different cell types or under different conditions could be quantified by microarray or tag sequencing methods [e.g., SAGE (serial analysis of gene expression) and MPSS (massively parallel signature sequencing)] in the past decade. Microarray assay detects the hybridization intensity of cDNA and its complementary oligonucleotide probes to estimate the abundance of a particular mRNA species (Schena *et al.*, 1995).

Tag sequencing method uses sequencing of cDNA fragments (typically the 3′ or 5′ end of mRNA) followed by counting the number of times a fragment has been sequenced. SAGE was the first reported tag sequencing method used for the analysis of gene expression profiling (Velculescu *et al.*, 1995). The MPSS method generates small fragment signatures of mRNA to determine the expression profiles (Brenner *et al.*, 2000). To date, SAGE and MPSS have been used in conjunction with NGS technologies, such as LongSAGE (Nielsen, Hogh, and Emmersen, 2006; Saha *et al.*, 2002) and DeepSAGE (Nielsen, Hogh, and Emmersen, 2006). A novel method termed *robust analysis of 5′ transcript ends* (*5′-RATE*) has been developed (Gowda *et al.*, 2006). Another novel tag-based method based on the 454 technology is 3′-UTR sequencing, which focuses on sequencing unique fragments found at the 3′-UTR of genes to allow for distinction of closely related transcripts (Eveland, McCarty, and Koch, 2008).

RNA-Seq technology based on Illumina sequencing platform has become widely utilized as a tool to understand the transcriptome of a given experimental system in the area of transcript profiling, discovery of novel transcripts, and identification of alternative splicing events (Guttman *et al.*, 2009; Maher *et al.*, 2009; Mortazavi *et al.*, 2008; Wang, Gerstein, and Snyder, 2009b). This method can gain information about the RNA content and transcriptional status of an experimental system. Methods for making NGS-specific RNA-Seq libraries typically comprise preparing rRNA (ribosomal ribonucleic acid)-depleted RNA, following by RNA fragmentation, cDNA synthesis, 5′ and 3′ adaptor ligation, and multiple clean-up steps. The use of PCR amplification in cDNA preparation and poly(A) + mRNA as the starting material is the most common factor for RNA-Seq method.

RNA-Seq is becoming the *de facto* gold standard high throughput approach for transcriptome analysis, gradually replacing microarray as sequencing costs drop down. However, RNA-Seq library preparation is labor intensive, time consuming, and prone to technician errors to automate the RNA-Seq library construction process. More importantly, validation of the quality of the sequenced transcripts is critical.

Table 2   The transcriptome analysis of medicinal plants related to secondary metabolite biosynthesis.

Organisms	Main bioactive compounds	Analysis of organs/tissues	Platforms	References
Artemisia annua	Artemisinin	Glandular trichome	454	Wang *et al*. (2009a)
Bupleurum chinense	Saikosaponins	Roots	454	Sui *et al*. (2011)
Camptotheca acuminata	Camptothecin	Leaves	454	Sun *et al*. (2011)
Catharanthus roseus	Vinblastine and vincristine	Leaves, roots	Sanger	Shukla *et al*. (2006)
Catharanthus roseus	Vinblastine and vincristine	Cell suspensions	Sanger	Rischer *et al*. (2006)
Cistus creticus	Labdane-type diterpenes	Trichomes	Sanger	Falara *et al*. (2008)
Fritillaria cirrhosa	Steroidal alkaloid	Bulbs	Sanger	Sun *et al*. (2011)
Ginkgo biloba	Ginkgolide and flavonoid	Leaves	454	Lin *et al*. (2011)
Glycyrrhiza uralensis	Glycyrrhizin	Roots, stems, leaves	454	Li *et al*. (2010a)
Glycyrrhiza uralensis	Glycyrrhizin	Thickened roots, a rhizome-like organ	Sanger	Sudo *et al*. (2009)
Gynostemma pentaphyllum	Gynosaponins	Root, leaves	454	Subramaniyam *et al*. (2011)
Huperzia serrata	Huperzine A	Leaves	Sanger	Luo *et al*. (2010b)
Huperzia serrata and *Phlegmariurus carinatus*	Lycopodium alkaloid	Roots, stems, leaves	454	Luo *et al*. (2010a)
Momordica charantia	Fatty acids	Seeds	454	Yang *et al*. (2010)
Opium poppy	Benzylisoquinoline alkaloids	Cell cultures	454	Desgagne-Penix *et al*. (2010)
Panax ginseng	Ginsenoside	Roots	454	Chen *et al*. (2011)
Panax ginseng	Ginsenoside	Hairy roots	454	Choi *et al*. (2005)
Panax ginseng	Ginsenoside	Leaves	Sanger	Kim *et al*. (2006)
Panax ginseng	Ginsenoside	Hairy roots, roots	Sanger	Sathiyamoorthy *et al*. (2010)
Panax ginseng	Ginsenoside	Seedlings	Sanger	Sathiyaraj *et al*. (2011)
Panax quinquefolius	Ginsenoside	Roots	454	Sun *et al*. (2010)
Panax quinquefolius	Ginsenoside	Roots, flowers, leaves	Sanger	Wu *et al*. (2010)
Panax notoginseng	Ginsenoside	Roots	454	Luo *et al*. (2011)
Perilla frutescens	Anthocyanins	Leaves	Sanger	Yamazaki *et al*. (2008)
Salvia fruticosa	Terpenoids, flavonoids	Glandular trichomes	Sanger	Chatzopoulou *et al*. (2010)
Siraitia grosvenorii	Mogrosides	Fruits	Solexa	Tang *et al*. (2011)
Salvia miltiorrhiza	Tanshinones, salvianolic acids	Plantlets, flowers, stems, leaves, roots	Solexa	Hua *et al*. (2011)
Salvia miltiorrhiza	Tanshinones, salvianolic acids	Roots	454	Li *et al*. (2010b)
Taxus chinensis	Taxol	Cell cultures	Solexa	Qiu *et al*. (2009)
Taxus cuspidata	Taxol	Needles	454	Wu *et al*. (2011)
Taxus mairei	Taxane	Roots, stems, leaves	Solexa	Hao *et al*. (2011)

6.3   Protein-Coding Gene Annotation Using Transcriptome Data

Protein-coding gene annotation sheds light on the essential information of genomes. Genome annotation includes the discovery of novel noncoding RNA genes, regulatory elements that monitor temporal or spatial gene expression, and splice variants, as well as the illustration of the transcription start and polyadenylation sites and exon–intron structures. In detail, ESTs or full-length complementary deoxyribonucleic acid (FLcDNA) sequences are aligned and mapped to reference genome sequences, illustrating the presence of exons, introns, exon junctions, and transcription boundaries for genes. ESTs

or FLcDNAs generated from Sanger sequencing method provide an accurate and effective means for the annotation of many protein-coding genes (Hillier *et al*., 1996; Seki *et al*., 2002). The limitation of Sanger sequencing is the low coverage of transcriptome, resulting in the annotation only for most abundantly expressed genes.

The NGS technologies can produce a large amount of sequence data to address the gap in the protein-coding gene annotation being limited only for abundant genes based on Sanger sequencing. Particularly, owing to the longer read length and sequence accuracy compared to that generated by other NGS technologies, ESTs produced by 454 sequencing can be effectively used for *de novo*

assembly of transcriptomes (Chen *et al.*, 2011; Lin *et al.*, 2011; Luo *et al.*, 2011; Sun *et al.*, 2010, 2011). The shorter reads produced by Illumina, SOLiD, and HeliScope create challenges for *de novo* sequencing assembly leading to the limitations on gene annotation. However, the much deeper sequencing capacity of NGS technologies contributing to the data of these short reads has been successfully used to identify novel exons and novel splicing events in species with available reference genome sequences (Cloonan *et al.*, 2008; Morin *et al.*, 2008).

In addition, the reference genome sequences are not available for most of the medicinal plants. In such cases, annotations based on EST sequencing can be performed by comparative analysis of the derived ESTs with reference genomes of related species. The data generated from gene expression profiling studies that use high throughput sequencing can also provide protein-coding gene annotation information. For instance, Reid *et al.* (2008) reported the annotation information about transcriptional start sites and other molecular events involving the 5′ end of transcripts using the 5′-RATE method (Reid *et al.*, 2008). The annotation of protein-coding genes is important for studies of dynamic genetic architectures, revealing critical developmental pathways and enabling our understanding of evolution.

6.4 Noncoding RNA and Novel Small RNA Discovery

Small ncRNAs are RNA molecules that are crucial regulators of development and cell fate determination and not translated into protein products. These 18- to 30-nucleotide-long RNA molecules include transfer ribonucleic acid (tRNA), rRNA, micro ribonucleic acid (miRNA) and small interfering ribonucleic acid (siRNA). Among these ncRNAs, miRNA and siRNA are similar in size and both are involved in posttranscriptional regulation of gene expression in a wide range of organisms (Filipowicz, Bhattacharyya, and Sonenberg, 2008; Gowda, Li, and Wang, 2007).

The high throughput miRNA profiling has been analyzed using microarray-based approaches, which are not suitable for the identification of novel miRNAs (Meyers *et al.*, 2006). High throughput sequencing of small RNAs has a profound influence on ncRNA research. MPSS technologies have been applied for the discovery of novel miRNA and siRNA

genes (Lu *et al.*, 2005). The disadvantages of this approach include the high complexity and cost of the MPSS technology, as well as the short reads corresponding to only a portion of the small RNA molecules.

NGS technologies provide several advantages for ncRNA sequencing studies over MPSS method, including the decreased procedural complexity and cost and the increasing high throughput and depth of coverage. In addition to novel miRNA discovery, these NGS-sequencing-based approaches also enable to detect variants in the mature miRNA length, miRNA editing events, and miRNA–target RNA pairs (German *et al.*, 2008; Reid *et al.*, 2008). Small RNA profiling studies using NGS technologies (e.g., 454 sequencing) have characterized ncRNAs in many species (Aravin *et al.*, 2006; Axtell *et al.*, 2006; Yao *et al.*, 2007; Zhao *et al.*, 2007). Wu *et al.* (2012) characterized a large number of messenger ribonucleic acid-like noncoding ribonucleic acid (mlncRNA) based on 454 high throughput sequencing and detected the importance of mlncRNAs in secondary metabolism and stress response in *Digitalis purpurea* (Wu *et al.*, 2012). The high throughput of Illumina and SOLiD allows for the generation of deep miRNA libraries, providing even deeper coverage of small RNAs than the 454 platform. Qiu *et al.* (2009) reported the effect of the taxoid elicitor methyl jasmonate (MJ) on the miRNA expression of *Taxus chinensis* using Illumina sequencing system (Qiu *et al.*, 2009). A novel class of small RNAs, termed *Piwi-interacting ribonucleic acid* (*piRNA*), has been discovered, which was benefited from the utility of 454 sequencing technology in small RNA sequencing studies (Girard *et al.*, 2006; Houwing *et al.*, 2007; Lau *et al.*, 2006).

6.5 Genetic Marker Discovery and Selection

Genetic markers are generated from the variations of nucleotide sequences of individual species. The two typical classes of genetic markers are simple sequence repeats (SSRs) and single nucleotide polymorphisms (SNPs). SNPs, also termed as the *third-generation markers*, have been widely used in genetic studies of many plants (Gupta, Roy, and Prasad, 2001). ESTs are a large resource for the identification of SSRs and SNPs. EST-derived SSRs and SNPs from expressed sequences that are

Table 3 The identification of genetic markers for medicinal plants based on transcriptome analysis.

Organisms	Genetic markers	Platforms	References
Artemisia annua	SNP	454	Graham *et al.* (2010)
Epimedium sagittatum	SSR	454	Zeng *et al.* (2010)
Ginkgo biloba	SSR	454	Lin *et al.* (2011)
Huperzia serrata	SSR	Sanger and 454	Luo *et al.* (2010a, b)
Phlegmariurus carinatus	SSR	454	Luo *et al.* (2010a)
Panax quinquefolius	SSR	Sanger	Wu *et al.* (2010)
Salvia miltiorrhiza	SSR	454	Li *et al.* (2010b)
Taxus cuspidata	SSR	454	Wu *et al.* (2011)

more conserved than the nongenic sequences are potentially tightly linked with functional genes that may control certain important agronomic characters (Bozhko *et al.*, 2003; Graham *et al.*, 2010). EST-derived SSRs and SNPs have been proven useful for detecting the signature of genetic diversity and population structure of plants (Ching and Rafalski, 2002; Vasemagi, Nilsson, and Primmer, 2005; von Stackelberg, Rensing, and Reski, 2006). EST-derived SSRs and SNPs have been identified in the medicinal plants that are summarized in Table 3. Graham *et al.* (2010) used deep sequencing on the transcriptome of *A. annua* to identify the genes and genetic markers (SNP) that control artemisinin yield for fast-track breeding (Graham *et al.*, 2010). Zeng *et al.* (2010) identified 2810 EST-SSRs from the EST dataset and characterized these EST-SSRs in a traditional Chinese medicinal plant, *Epimedium sagittatum* (Zeng *et al.*, 2010). Many SSRs have been detected by mining ESTs from *Ginkgo biloba* (Lin *et al.*, 2011), *H. serrata* (Luo *et al.*, 2010a, b), *P. carinatus* (Luo *et al.*, 2010a), *Panax quinquefolius* (Wu *et al.*, 2010), *S. miltiorrhiza* (Li *et al.*, 2010b), and *Taxus cuspidate* (Wu *et al.*, 2011) (Table 3). These EST-SSRs are powerful resources for further studies on taxonomy, molecular breeding, genetics, genomics, and secondary metabolism in medicinal plants.

7 EXAMPLES OF 454 SEQUENCING FOR TRANSCRIPTOME ANALYSIS

To illustrate the process of transcriptome analysis of medicinal plants using NGS technologies, an example of 454 sequencing and data analysis for transcriptome analysis was present. First, total RNA was isolated and mRNA was purified using RNeasy Plus Mini Kit (QIAGEN, CA) and Oligotex® mRNA Midi Kit (QIAGEN, CA), respectively. Second, the cDNA was synthesized from the purified poly (A) RNA according to the manufacturer's instructions provided with Clontech's SMART cDNA synthesis kit (Clontech, USA). Thus, the cDNA samples were generated as the starting material for 454 sequencing, followed by GS FLX Titanium cDNA Rapid Library Preparation and GS FLX Titanium Sequencing according to the manufacturer's instructions provided with (454 Life Sciences, Roche).

Approximately 5 μg of amplified cDNA was sheared by nebulization to produce random fragments of approximately 300–800 bp in length for 454 sequencing. The fragmented cDNA samples were assessed by gel electrophoresis to evaluate the effectiveness of the process. Oligonucleotide adaptors were ligated to the fragmented cDNA samples according to standard procedures. The DNA fragments were then denatured to generate single-stranded DNA that was amplified by emulsion PCR for sequencing. The sequencing of the libraries was performed on a 454 GS FLX Titanium sequencing platform (454 Life Sciences, Roche).

8 CONCLUDING REMARKS

NGS technologies have been rapidly improving quality, read length, and throughput to make transcriptome sequencing and association study possible in a very cost-effective manner. Continued improvement and development of sample preparation protocols and data analysis tools have been significantly helping in transcriptome analysis. Recent advances in NGS open doors to conservation genomic/transcriptomic studies of wildlife species, including most of the medicinal plants for which only limited genetic resources exist. Characterization of the transcriptome has previously revealed many active genes with spatiotemporal regulation and evolutionary significance.

REFERENCES

Alagna, F., D'Agostino, N., Torchia, L., *et al*. (2009) *BMC Genomics*, **10**, 399.

Aravin, A., Gaidatzis, D., Pfeffer, S., *et al*. (2006) *Nature*, **442**, 203–207.

Axtell, M. J., Jan, C., Rajagopalan, R., *et al*. (2006) *Cell*, **127**, 565–577.

Bentley, D. R. (2006) *Curr. Opin. Genet. Dev.*, **16**, 545–552.

Bowers, J., Mitchell, J., Beer, E., *et al*. (2009) *Nat. Methods*, **6**, 593–595.

Bozhko, M., Riegel, R., Schubert, R., *et al*. (2003) *Mol. Ecol.*, **12**, 3147–3155.

Branton, D., Deamer, D. W., Marziali, A., *et al*. (2008) *Nat. Biotechnol.*, **26**, 1146–1153.

Braslavsky, I., Hebert, B., Kartalov, E., *et al*. (2003) *Proc. Natl. Acad. Sci. U. S. A.*, **100**, 3960–3964.

Brenner, S., Johnson, M., Bridgham, J., *et al*. (2000) *Nat. Biotechnol.*, **18**, 630–634.

Chatzopoulou, F. M., Makris, A. M., Argiriou, A., *et al*. (2010) *Plant Cell Rep.*, **29**, 523–534.

Chen, S., Luo, H., Li, Y., *et al*. (2011) *Plant Cell Rep.*, **30**, 1593–1601.

Ching, A., and Rafalski, A. (2002) *Cell. Mol. Biol. Lett.*, **7**, 803–810.

Choi, D. W., Jung, J., Ha, Y. I., *et al*. (2005) *Plant Cell Rep.*, **23**, 557–566.

Cloonan, N., Forrest, A. R., Kolle, G., *et al*. (2008) *Nat. Methods*, **5**, 613–619.

De Luca, V., and Laflamme, P. (2001) *Curr. Opin. Plant Biol.*, **4**, 225–233.

Desgagne-Penix, I., Khan, M. F., Schriemer, D. C., *et al*. (2010) *BMC Plant Biol.*, **10**, 252.

Eid, J., Fehr, A., Gray, J., *et al*. (2009) *Science*, **323**, 133–138.

Eveland, A. L., McCarty, D. R., and Koch, K. E. (2008) *Plant Physiol.*, **146**, 32–44.

Falara, V., Fotopoulos, V., Margaritis, T., *et al*. (2008) *Plant Mol. Biol.*, **68**, 633–651.

Filipowicz, W., Bhattacharyya, S. N., and Sonenberg, N. (2008) *Nat. Rev. Genet.*, **9**, 102–114.

Flusberg, B. A., Webster, D. R., Lee, J. H., *et al*. (2010) *Nat. Methods*, **7**, 461–465.

German, M. A., Pillay, M., Jeong, D. H., *et al*. (2008) *Nat. Biotechnol.*, **26**, 941–946.

Girard, A., Sachidanandam, R., Hannon, G. J., *et al*. (2006) *Nature*, **442**, 199–202.

Gowda, M., Li, H., Alessi, J., *et al*. (2006) *Nucleic Acids Res.*, **34**, e126.

Gowda, M., Li, H., and Wang, G. L. (2007) *Nat. Protoc.*, **2**, 1622–1632.

Graham, I. A., Besser, K., Blumer, S., *et al*. (2010) *Science*, **327**, 328–331.

Greenleaf, W. J., and Block, S. M. (2006) *Science*, **313**, 801.

Gupta, P. K., Roy, J. K., and Prasad, M. (2001) *Curr. Sci.*, **80**, 524–536.

Guttman, M., Amit, I., Garber, M., *et al*. (2009) *Nature*, **458**, 223–227.

Hall, N. (2007) *J. Exp. Biol.*, **210**, 1518–1525.

Hao, D. C., Ge, G. B., Xiao, P. G., *et al*. (2011) *PLoS ONE*, **6**, e21220.

Harris, T. D., Buzby, P. R., Babcock, H., *et al*. (2008) *Science*, **320**, 106–109.

Hillier, L. D., Lennon, G., Becker, M., *et al*. (1996) *Genome Res.*, **6**, 807–828.

Houwing, S., Kamminga, L. M., Berezikov, E., *et al*. (2007) *Cell*, **129**, 69–82.

Hua, W., Zhang, Y., Song, J., *et al*. (2011) *Genomics*, **98**, 272–279.

Jennewein, S., Wildung, M. R., Chau, M., *et al*. (2004) *Proc. Natl. Acad. Sci. U. S. A.*, **101**, 9149–9154.

Kim, M. K., Lee, B. S., In, J. G., *et al*. (2006) *Plant Cell Rep.*, **25**, 599–606.

Lau, N. C., Seto, A. G., Kim, J., *et al*. (2006) *Science*, **313**, 363–367.

Li, Y., Luo, H. M., Sun, C., *et al*. (2010a) *BMC Genomics*, **11**, 268.

Li, Y., Sun, C., Luo, H. M., *et al*. (2010b) *Yao Xue Xue Bao*, **45**, 524–529.

Lin, X., Zhang, J., Li, Y., *et al*. (2011) *Physiol. Plant.*, **143**, 207–218.

Lipson, D., Raz, T., Kieu, A., *et al*. (2009) *Nat. Biotechnol.*, **27**, 652–658.

Lu, C., Tej, S. S., Luo, S., *et al*. (2005) *Science*, **309**, 1567–1569.

Luo, H., Li, Y., Sun, C., *et al*. (2010a) *BMC Plant Biol.*, **10**, 209.

Luo, H., Sun, C., Li, Y., *et al*. (2010b) *Physiol. Plant.*, **139**, 1–12.

Luo, H., Sun, C., Sun, Y., *et al*. (2011) *BMC Genomics*, **12**(Suppl 5), S5.

Maher, C. A., Kumar-Sinha, C., Cao, X., *et al*. (2009) *Nature*, **458**, 97–101.

Margulies, M., Egholm, M., Altman, W. E., *et al*. (2005) *Nature*, **437**, 376–380.

Mason, C. E., and Elemento, O. (2012) *Genome Biol.*, **13**, 314.

Meyers, B. C., Souret, F. F., Lu, C., *et al*. (2006) *Curr. Opin. Biotechnol.*, **17**, 139–146.

Morin, R., Bainbridge, M., Fejes, A., *et al*. (2008) *Biotechniques*, **45**, 81–94.

Morozova, O., Hirst, M., and Marra, M. A. (2009) *Annu. Rev. Genomics Hum. Genet.*, **10**, 135–151.

Mortazavi, A., Williams, B. A., McCue, K., *et al*. (2008) *Nat. Methods*, **5**, 621–628.

Murata, J., Roepke, J., Gordon, H., *et al*. (2008) *Plant Cell*, **20**, 524–542.

Newman, D. J., and Cragg, G. M. (2007) *J. Nat. Prod.*, **70**, 461–477.

Nielsen, K. L., Hogh, A. L., and Emmersen, J. (2006) *Nucleic Acids Res.*, **34**, e133.

Novaes, E., Drost, D. R., Farmerie, W. G., *et al*. (2008) *BMC Genomics*, **9**, 312.

Nyren, P., Pettersson, B., and Uhlen, M. (1993) *Anal. Biochem.*, **208**, 171–175.

Ohlrogge, J., and Benning, C. (2000) *Curr. Opin. Plant Biol.*, **3**, 224–228.

Oksman-Caldentey, K. M., and Saito, K. (2005) *Curr. Opin. Biotechnol.*, **16**, 174–179.

Parkinson, J. (2009) *Expressed Sequence Tags (ESTs) Generation and Analysis*, Humana Press, New York.

Pushkarev, D., Neff, N. F., and Quake, S. R. (2009) *Nat. Biotechnol.*, **27**, 847–850.

Qiu, D., Pan, X., Wilson, I. W., *et al*. (2009) *Gene*, **436**, 37–44.

Reid, J. G., Nagaraja, A. K., Lynn, F. C., *et al*. (2008) *Genome Res.*, **18**, 1571–1581.

Rischer, H., Oresic, M., Seppanen-Laakso, T., *et al*. (2006) *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 5614–5619.

Ronaghi, M., Karamohamed, S., Pettersson, B., *et al*. (1996) *Anal. Biochem.*, **242**, 84–89.

Saha, S., Sparks, A. B., Rago, C., *et al*. (2002) *Nat. Biotechnol.*, **20**, 508–512.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.*, **74**, 5463–5467.

Sathiyamoorthy, S., In, J. G., Gayathri, S., *et al*. (2010) *Mol. Biol. Rep.*, **37**, 3465–3472.

Sathiyaraj, G., Lee, O. R., Parvin, S., *et al*. (2011) *Mol. Biol. Rep.*, **38**, 2761–2769.

Schadt, E. E., Linderman, M. D., Sorenson, J., *et al*. (2010a) *Nat. Rev. Genet.*, **11**, 647–657.

Schadt, E. E., Turner, S., and Kasarskis, A. (2010b) *Hum. Mol. Genet.*, **19**, 227–240.

Schena, M., Shalon, D., Davis, R. W., *et al*. (1995) *Science*, **270**, 467–470.

Seki, M., Narusaka, M., Kamiya, A., *et al*. (2002) *Science*, **296**, 141–145.

Shendure, J., Porreca, G. J., Reppas, N. B., *et al*. (2005) *Science*, **309**, 1728–1732.

Shukla, A. K., Shasany, A. K., Gupta, M. M., *et al*. (2006) *J. Exp. Bot.*, **57**, 3921–3932.

Song, C. X., Clark, T. A., Lu, X. Y., *et al*. (2012) *Nat. Methods*, **9**, 75–77.

von Stackelberg, M., Rensing, S. A., and Reski, R. (2006) *BMC Plant Biol.*, **6**, 9.

Stone, E. M., Rothblum, K. N., Alevy, M. C., *et al*. (1985) *Proc. Natl. Acad. Sci. U. S. A.*, **82**, 1628–1632.

Subramaniyam, S., Mathiyalagan, R., Jun Gyo, I., *et al*. (2011) *Plant Cell Rep.*, **30**, 2075–2083.

Sudo, H., Seki, H., Sakurai, N., *et al*. (2009) *Plant Biotechnol.*, **26**, 105–107.

Sui, C., Zhang, J., Wei, J., *et al*. (2011) *BMC Genomics*, **12**, 539.

Sun, C., Li, Y., Wu, Q., *et al*. (2010) *BMC Genomics*, **11**, 262.

Sun, C., Sun, Y., Song, J., *et al*. (2011) *J. Med. Plants Res.*, **21**, 5307–5314.

Sun, Y., Luo, H., Li, Y., *et al*. (2011) *BMC Genomics*, **12**, 533.

Tang, Q., Ma, X., Mo, C., *et al*. (2011) *BMC Genomics*, **12**, 343.

Tawfik, D. S., and Griffiths, A. D. (1998) *Nat. Biotechnol.*, **16**, 652–656.

Teoh, K. H., Polichuk, D. R., Reed, D. W., *et al*. (2006) *FEBS Lett.*, **580**, 1411–1416.

Tessler, L. A., Reifenberger, J. G., and Mitra, R. D. (2009) *Anal. Chem.*, **81**, 7141–7148.

Travers, K. J., Chin, C. S., Rank, D. R., *et al*. (2010) *Nucleic Acids Res.*, **38**, e159.

Vasemagi, A., Nilsson, J., and Primmer, C. R. (2005) *Mol. Biol. Evol.*, **22**, 1067–1076.

Vega-Arreguin, J. C., Ibarra-Laclette, E., Jimenez-Moraila, B., *et al*. (2009) *BMC Genomics*, **10**, 299.

Velculescu, V. E., Zhang, L., Vogelstein, B., *et al*. (1995) *Science*, **270**, 484–487.

Wang, W., Wang, Y., Zhang, Q., *et al*. (2009a) *BMC Genomics*, **10**, 465.

Wang, Z., Gerstein, M., and Snyder, M. (2009b) *Nat. Rev. Genet.*, **10**, 57–63.

Wu, Q., Song, J., Sun, Y., *et al*. (2010) *Physiol. Plant.*, **138**, 134–149.

Wu, Q., Sun, C., Luo, H., *et al*. (2011) *Planta Med.*, **77**, 394–400.

Wu, B., Li, Y., Yan, H., *et al*. (2012) *BMC Genomics*, **13**, 15.

Yamazaki, M., Shibata, M., Nishiyama, Y., *et al*. (2008) *FEBS J.*, **275**, 3494–3502.

Yang, P., Li, X., Shipp, M. J., *et al*. (2010) *BMC Plant Biol.*, **10**, 250.

Yao, Y., Guo, G., Ni, Z., *et al*. (2007) *Genome Biol.*, **8**, R96.

Zeng, S., Xiao, G., Guo, J., *et al*. (2010) *BMC Genomics*, **11**, 94.

Zhao, T., Li, G., Mi, S., *et al*. (2007) *Genes Dev.*, **21**, 1190–1203.

FURTHER READINGS

Cloonan, N., Forrest, A. R., Kolle, G., *et al*. (2008) *Nat. Methods*, **5**, 613–619.

De Luca, V., and Laflamme, P. (2001) *Curr. Opin. Plant Biol.*, **4**, 225–233.

Filipowicz, W., Bhattacharyya, S. N., and Sonenberg, N. (2008) *Nat. Rev. Genet.*, **9**, 102–114.

Flusberg, B. A., Webster, D. R., Lee, J. H., *et al*. (2010) *Nat. Methods*, **7**, 461–465.

German, M. A., Pillay, M., Jeong, D. H., *et al*. (2008) *Nat. Biotechnol.*, **26**, 941–946.

Morozova, O., Hirst, M., and Marra, M. A. (2009) *Annu. Rev. Genomics Hum. Genet.*, **10**, 135–151.

Mortazavi, A., Williams, B. A., McCue,, K., *et al*. (2008) *Nat. Methods*, **5**, 621–628.

Microarray

Chang Liu, Haimei Chen, Jianqin Li and Xiaolan Xu

Institute of Medicinal Plant Development, Chinese Academy of Medical Science, Beijing, PR China

1 INTRODUCTION

On the basis of the definition from Wikipedia, "A microarray is a multiplex lab-on-a-chip. It is a 2D array on a solid substrate (usually a glass slide or silicon thin-film cell) that assays large amounts of biological material using high throughput screening methods." *Microarray* is a rather generic name and can be used to refer to any setup designed for high throughput analysis. There are many types of microarrays based on the biological materials immobilized on the solid substrate and the purpose of the microarray. The microarray commonly used in genomic research and drug discovery include: (i) DNA microarrays, such as cDNA microarrays, oligonucleotide microarrays, and single nucleotide polymorphism (SNP) microarrays; (ii) protein and peptide microarrays for detailed analyses or optimization of protein–protein interactions; (iii) cellular microarrays, also called *transfection microarrays*; (iv) tissue microarrays (TMAs); and (v) chemical compound microarray, including carbohydrate arrays (glycoarrays), and so on.

On one end of the spectrum, the advanced microarray technology is built around a device called *lab-on-a-chip* (*LOC*), which integrates one or several laboratory functions on a single chip of only millimeters to a few square centimeters in size. LOCs deal with the handling of extremely small fluid volumes down to less than pico liters. LOC devices are a subset of micro-electromechanical system and often known as *micro total analysis systems* (*μTASs*). LOCs may provide various advantages, which are specific to their application. Typical advantages are: (i) low fluid volumes consumption (less waste, lower reagents costs, and less required sample volumes for diagnostics); (ii) faster analysis and response times due to short diffusion distances, fast heating, high surface to volume ratios, and small heat capacities; (iii) better process control because of a faster response of the system (e.g., thermal control for exothermic chemical reactions); (iv) compactness of the systems due to integration of much functionality and small volumes; (v) massive parallelization due to compactness, which allows high throughput analysis; (vi) lower fabrication costs, allowing cost-effective disposable chips, fabricated in mass production; and (vii) safer platform for chemical, radioactive, or biological studies because of integration of functionality, smaller fluid volumes, and stored energies.

On the other end of the spectrum, a microarray is a simple multiplex of the thousands or even millions of single assay system, which do not utilize any electromechanical systems. Although simple in fabrication, they can offer many advantages similar to those for the "LOC" described above. For example, in tissue array technology, thousands of tissue slice are spotted to a chip to allow automated, parallel, and standard immunohistochemistry and hybridization assays to be conducted. In the following text, we will focus on those microarray technologies that are used in genomic research and drug discovery. We start in Section 2 to briefly overview the principle for various

types of microarray technologies and summarize their potential applications, hoping to provide ideas for readers to explore.

2 PRINCIPLE AND APPLICATIONS OF VARIOUS TYPES OF MICROARRAY TECHNOLOGIES

In this section, we will briefly discuss the principles and applications of various types of microarray technologies. It should be pointed out that many microarray technologies are essentially a target approach in the sense that they need information about the DNA and protein sequences for the organism under study. This might not be the case for many medicinal plants whose genetic information are minimal. In this section, we review the applications that have been applied to all organisms beyond medicinal plants, hoping that they would help researchers on medicinal plants in their particular studies.

2.1 DNA Microarray (DMA)

DNA microarray (also commonly known as *gene chip*, *DNA chip*, or *biochip*) is a collection of microscopic DNA spots attached to a solid surface. The technology is a key element in today's functional genomics research. The power of the method lies in miniaturization, automation, and parallelism, permitting large-scale and genome-wide acquisition of quantitative biological information from multiple samples (Chan *et al*., 2009).

The core principle behind microarrays is hybridization between two DNA strands, the property of complementary nucleic acid sequences to specifically pair with each other by forming hydrogen bonds between complementary nucleotide base pairs. The higher number of complementary base pairs will translate into more hydrogen bonds and tighter non-covalent bonding between the two strands. At appropriate conditions, nonspecific bonding sequences can be washed off and only the strongly paired strands will remain hybridized. As a whole, the signal-resulted binding between fluorescently labeled target sequences and a probe sequence depends on the number of complementary bases, the hybridization conditions, and the washing off conditions. Total strength of the signal, from a spot

(feature), depends on the amount of target sample binding to the probes present on that spot.

DNA microarray technology is the most used microarray technologies in biological research. There are many excellent reviews and books on DNA microarray and its applications in biological research, and the interested users are recommended to read them (Howbrook *et al*., 2003; Shaw *et al*., 2009; Stears, Martinsky, and Schena, 2003). The DNA microarray can be divided into two categories, the cDNA array and oligo array (Figure 1). Below, we summarize the main applications of DMA.

1. *Gene Expression Profiling*. This is the most popular application of DMA. In such an experiment, the expression levels of thousands of genes are simultaneously monitored to study the effects of certain treatments, diseases, and developmental stages on gene expression. For example, the effects of active components of medicinal plants on various human cells can be studied (Adomas *et al*., 2008).
2. *Comparative Genome Hybridization*. Using this method, the genome content in different cells or closely related organisms can be assessed (Pollack *et al*., 1999; Moran *et al*., 2004).
3. *GeneID*. By combining PCR (polymerase chain reaction) and microarray technologies, the identification of particular organisms can be tested (Maskos and Southern, 1992).
4. *Chromatin Immunoprecipitation* (*ChIP*). DNA sequences bound to a particular protein can be isolated by immunoprecipitating that protein (ChIP); these fragments can be then hybridized to a microarray (such as a tiling array), allowing the determination of protein binding site occupancy throughout the genome (Collas, 2010).
5. *DamID* (*DNA adenine methyltransferase identification*). This method is analogous to ChIP. Genomic regions bound by a protein of interest can be isolated and used to probe a microarray to determine binding site occupancy. However, unlike ChIP, DamID does not require antibodies but makes use of adenine methylation near the protein's binding sites to selectively amplify those regions, introduced by expressing minute amounts of protein of interest fused to bacterial DNA adenine methyltransferase (van Steensel and Henikoff, 2000).
6. *SNP Detection*. This is the second most popular applications of DMA. This technology can
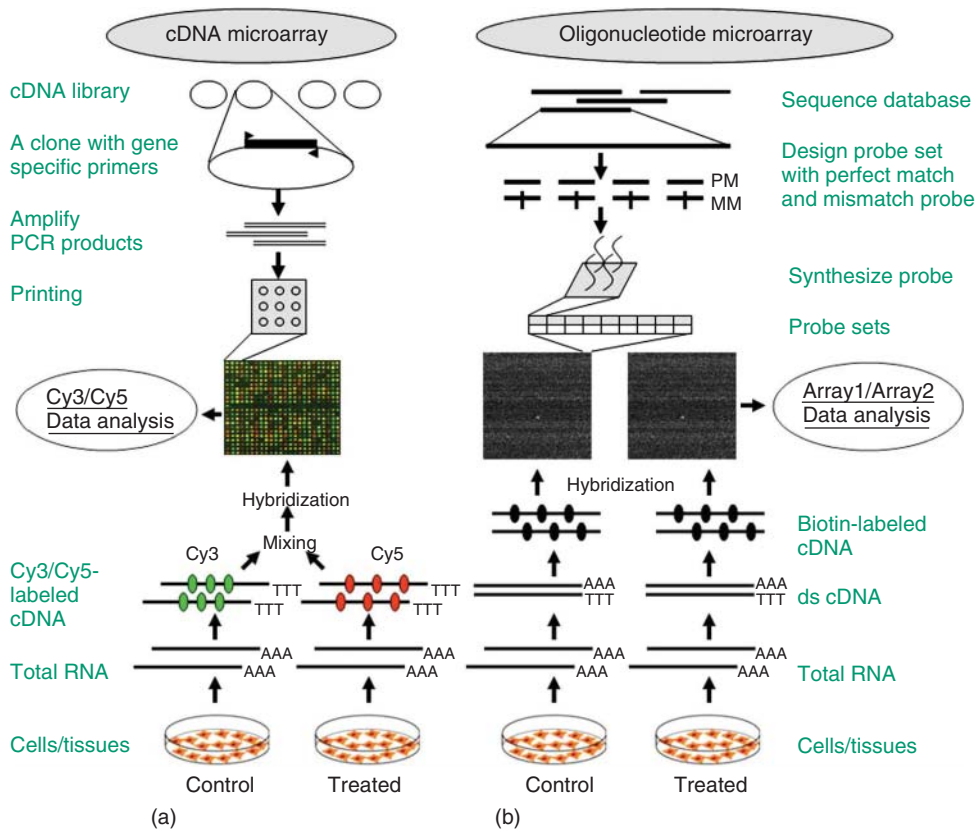
Figure 1 Schematic overview of array and target preparation for (a) cDNA and (b) oligonucleotide microarrays. Detailed descriptions can be found in da Silva and Yücel (2008). Left and right panes represent cDNA and oligonucleotide arrays, respectively. Top and bottom panes represent target and array preparation, respectively. Middle pane displays array images and data analyses. Inlets show spots and features in cDNA and oligonucleotide arrays, respectively.

be used to identify SNP among alleles within or between populations. Several applications of microarrays make use of SNP detection, including genotyping, forensic analysis, measuring predisposition to disease, identifying drug candidates, evaluating germ line mutations in individuals or somatic mutations in cancers, assessing loss of heterozygosity, or genetic linkage analysis (Hacia *et al*., 1999).

7. *Alternative Splicing*. An exon junction array design uses probes specific to the expected or potential splice sites of predicted exons for a gene. Exon arrays have a different design employing probes designed to detect each individual exon for known or predicted genes, and can be used for detecting different splicing isoforms. On the basis of the same principle, a fusion gene microarray can be used to detect fusion transcripts (Johnson

et al., 2003) and the tiling array can be used to identify alternative splicing in a more systematic manner.

8. *MicroRNA(miRNA) Array*. This method is used to detect the expression level of miRNAs. Known miRNA sequences will be spotted on slides or chips as probes to hybridize to hundreds or thousands of miRNA targets; the relative levels of miRNAs can then be determined (Shingara *et al*., 2005).

9. *DNA-Protein Array*. Double-stranded DNA corresponding to the exact binding sequence of proteins of interests can be spotted on an array. The array will then be used to probe samples with their proteins labeled with fluorescent dyes. The fluorescent detection method is compatible with standard microarray scanners; the spots on

the resulting image can be quantified by commonly used microarray quantification software packages.

2.2 Protein Microarray (PMA)

Protein microarrays (PMAs) (also biochip, protein chip) are used to determine the presence and/or amount of proteins in biological samples. They have the potential to be an important tool for proteomics research. Usually, a multitude of different capture agents, most frequently, monoclonal antibodies, DNA, or small molecules are deposited on a chip surface (glass or silicon) in a miniature array. Protein array has several applications:

1. to determine protein–protein interactions (Figure 2A);
2. to determine protein–lipid interactions (Figure 2B);
3. to determine protein–small molecule interactions (Figure 2C);
4. to determine antigen–antibody specificity (Figure 2D); and
5. to conduct enzymatic assays (Figure 2E);

Among these applications, antibody microarray is one of the most popular ones. With this antibody array, a collection of capture antibodies are spotted and fixed on a solid surface such as glass, plastic, or silicon chip, for the purpose of detecting antigens. Antibody microarray is often used for detecting protein expressions from cell lysates in general research and special biomarkers from serum or urine for diagnostic applications (Rivas *et al*., 2008; Chaga, 2008).

As described above, protein arrays are fabricated by spotting proteins on chips. The proteins can be externally synthesized, purified, and attached to the array. Alternatively, they can be synthesized *in situ* and directly attached to the array by cell-free DNA expression or chemical synthesis. With cell-free DNA expression, proteins are attached to the support right after their production. For chemical synthesis, peptides can be chemically produced by solid phase peptide synthesis methods and the resulted array is also called *peptide microarray*, commonly known as *peptide chip* or *peptide epitope microarray*. The
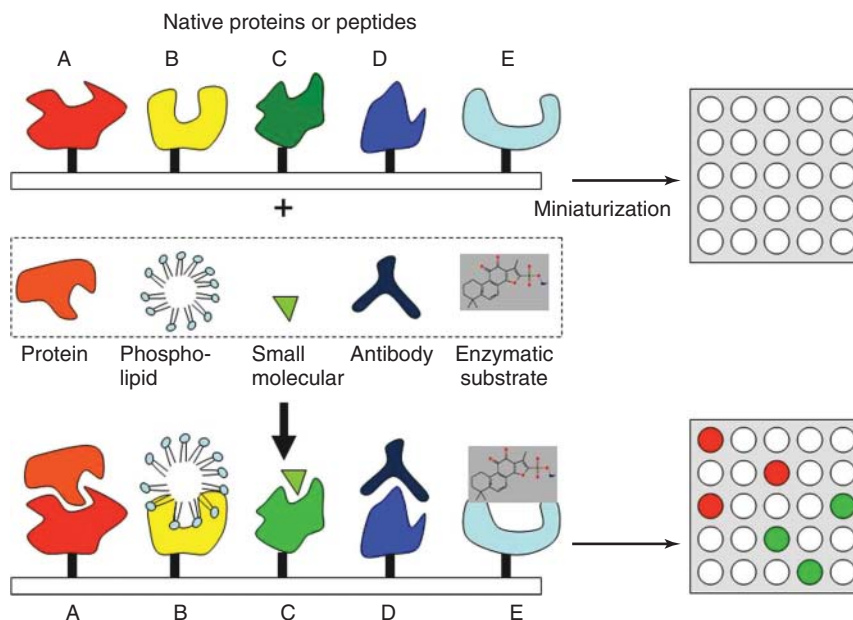
Figure 2 Functional protein microarray. Detailed descriptions can be found in Smith *et al*. (2005) and Phizicky *et al*. (2003). Native proteins or peptides are individually purified or synthesized and arrayed onto a suitable surface to form the functional protein microarrays. Spotted protein arrays can be used to study many interactions, including protein–protein, protein–lipid, and protein–small molecule. Consequently, this class of chips is particularly useful in drug and drug-target identification and in building biological networks.

assay principle of peptide microarrays is similar to an ELISA (enzyme-linked immunosorbent assay) protocol. The peptides (up to thousands in several copies) are linked to the surface of a glass chip. This peptide chip can directly be incubated with a variety of different biological samples such as purified enzymes or antibodies, patient or animal sera, cell lysates, and so on. After several washing steps, a secondary antibody with the needed specificity (e.g., anti-IgG human/mouse or anti-phosphotyrosine or anti-myc) is applied. Usually, the secondary antibody is tagged by a fluorescence label that can be detected by a fluorescence scanner (Panse *et al.*, 2004). Other detection methods are chemiluminescence or autoradiography.

Peptide microarrays can be used to study all kinds of protein–protein interactions. Most publications can be found in the context of immune monitoring and enzyme profiling. For immune monitoring, it has been used for mapping of immunodominant regions in antigens or whole proteomes (Lin *et al.*, 2009; Linnebacher *et al.*, 2012), seromarker discovery (Callaway, 2011), monitoring of clinical trials (Garren *et al.*, 2008), profiling of antibody signatures (Gaseitsiwe *et al.*, 2010), and finding neutralizing antibodies (Tomaras *et al.*, 2011). For enzyme profiling, it has been used for identification of substrates for orphan enzymes (Kindrachuk *et al.*, 2012); optimization of known enzyme substrates (Lizcano *et al.*, 2002); elucidation of signal transduction pathways (Delgado *et al.*, 2007); and detection of contaminating enzyme activities, consensus sequence, and key residues determination (Thiele *et al.*, 2011).

2.3 Small Molecule Microarray (SMM)

Small molecule microarrays or SMMs (Foong *et al.*, 2012) are a subclass of microarrays, which encompass molecules of low molecular weight, such as peptides, carbohydrates, chemical compounds, and natural products. The last decade has introduced extensive immobilization methods for the capture of diverse molecular classes onto arrays (Moran *et al.*, 2004; van Steensel and Henikoff, van Steensel and Henikoff, 2000). Libraries may also be fabricated on the surface directly, through iterative couplings or reactions performed *in situ*. As a result, large libraries with several thousands of molecules have been synthesized and screened on SMMs for many different applications. The library designs have matured,

through the active incorporation of encoding strategies, click chemistry, diversity-oriented synthesis, and fragment-based strategies that have transformed the diversity and efficiency of libraries synthesized (Moran *et al.*, 2004). Detection methods for binding on arrays have also extended beyond fluorescent labels, with the adoption of label-free optical impedance measurements (Hacia *et al.*, 1999), surface plasmon resonance (Johnson *et al.*, 2003), and mass spectrometry (MS)-based techniques. The readers are encouraged to consult some of the recent excellent review paper on this subject.

SMMs were first invented in 1999 by MacBeath and Schreiber who immobilized thiolated compounds onto maleimide-derivatized slides through a thioether bond and applied them to screening for binding against model proteins (Collas, 2010). By now, SMMs have many applications, extending from (i) drug and biomarker discovery, (ii) disease detection, and (iii) profiling.

2.4 Cell Microarray (CMA)

Cell microarray (CMA) really evolved from the need to miniaturize cell-based high throughput screening (HTS). Cell-based assays are the workhorse to evaluate potential drug targets by functionally characterizing their effect in cells, to assess specificity and efficacy of drug leads, and to identify the targets for drugs of unknown mechanism of action. Consequently, there is a continuous need to scale up the throughput level of cell-based assay. Initially, cell assays have been miniaturized by growing cells in 96-, 384-or 1536-well microtiter plates or by mixing cells with one-bead, one-compound chemical libraries. However, as chemical libraries and gene collections grow, one will need to further miniaturize cell assays to increase parallelism of cell analysis. To that end, CMAs provide an attractive solution. Because they can hold at least 5000–6000 spots on one slide, this enables a genome-wide screen on only a few slides, which would increase the throughput significantly (Carles *et al.*, 2005).

CMA can be very useful in two key steps of the drug discovery process: target identification and lead assessment. Many potential applications of CMA have been proposed (Castel *et al.*, 2006):

1. cDNA clones expressing particular drug targets, such as GPCRs (G-protein coupled receptors), and

then arrayed. Transfection reagents and cells can then be added to the chips; a detection assay will be used to identify the cells that have successful expression of the genes encoded by the cDNA clone. The chips can then be used to screen small molecules for their capacity to specifically interact with GPCR targets. This information could be useful as a predictive tool to identify chemical structures with potential to associate with particular drug targets.
2. CMAs could be used to study cellular responses to ligand binding, such as protein phosphorylation, calcium waves, gene transcription, and protein translocation.

For research involving medicinal plants, CMA can be used to study the effects of nature compounds isolated from medicinal plants, lead to the identification of lead compounds.

2.5   Tissue Microarray (TMA)

TMA technology allows rapid and simultaneous morphological assessment of molecular targets in large collectives of tissues. The principle of TMA is similar to other microarray technology in that cylindrical cores of tissues are acquired from morphologically representative areas of paraffin-embedded tissue samples and arrayed tightly into a recipient paraffin block. Typically, the tissue cores have a diameter of 0.6 mm and up to 1000 samples can be assembled within a single recipient block of conventional size. The TMA sections are suitable for all *in situ* analyses, including immunohistochemistry and *in situ* hybridization, as are conventional large sections (Tzankov *et al.*, 2005).
TMA has several advantages compared to the conventional large section studies:

1. It allows simultaneous analysis of a large number of cases.
2. It allows an exceptional level of laboratory procedure standardization, thereby reducing intra- and interlaboratory variations.
3. It allows a high level of standardization of morphological interpretation.
4. It facilitates automation of signal detection, minimizing machinery selection bias, as the areas of interest have already been achieved during the selection of the regions to be arrayed.

5. It will save time, material, technical assistance, and money.

Various TMA types can be constructed for the corresponding applications such as:

1. to determine the organotypic and histotypic distribution and expression pattern of candidate diagnostic and therapeutic targets using normal tissue TMA;
2. to determine whether the specific qualitative and quantitative target distribution within different neoplastic diseases is screened using multitumor TMA;
3. to survey molecular changes and identify potential diagnostic and therapeutic targets associated with progression using progression TMA that contains samples of different stages or precursor lesions of one particular disease or tumor entity;
4. to identify prognostically different disease subtypes using prognosis TMA, which contains samples with available clinical follow-up data; and
5. to conduct immunohistochemistry and hybridization study using TMA constructed from cell culture and frozen tissues.

3   RECENT APPLICATIONS OF DNA MICROARRAY ANALYSIS IN MEDICINAL PLANT RESEARCH

In Section 2, the principles of applications of various microarray technologies have been introduced. Among them, DNA microarray technology is best developed and, as a result, it is the microarray technology that has been used most extensively in medicinal plant research. The research can be divided into two categories. One category of application actually studies the medicinal plant themselves, whereas the second category of application studies the changes observed in the cells treated with active components derived from medicinal plants. The following sections will discuss the practical usage of DNA microarray technology on gene expression profiling, identification of toxic medicinal plants, functional genomics studying of the medicinal plant, and miRNA analysis.

3.1 Identification of Molecular Mechanisms and Novel Targets Involved in Natural Compounds of Medicinal Plants' Action on Different Human Diseases by Gene Expression Profiling

Most microarray studies implicating medicinal plants are using gene expression profiling to identify molecular mechanisms and novel targets involved in natural compounds of medicinal plants' action on different human diseases.

For example, we have recently analyzed Tanshinone IIA's (Tan IIA) apoptotic effects on leukemia cells by genome-wide expression profiling. Tan IIA is a diterpene quinone extracted from the root of *Salvia miltiorrhiza*, a Chinese traditional herb. Our results (Liu *et al.*, 2012) revealed that Tan IIA had different cytotoxic activities on five types of leukemia cells, with the highest toxicity on U-937 cells. Annexin V and Caspase-3 assays showed that Tan IIA induced apoptosis in U-937 cells. We employed Affymetrix GeneChip to identify the genes that are differentially expressed among leukemia cell lines exhibiting various Tan IIA sensitivities and significantly regulated after Tan IIA treatment as a way to elucidate the molecular mechanisms of Tan IIA in a systematic manner. 366 significantly differentially expressed genes (SDEGs) in Tan IIA-treated U-937 cells were identified. From the annotation, many of these genes were found belonging to cell-cycle control, apoptosis pathways. Among the genes related to the apoptosis pathway, chemokine ligand 2 (CCL2, also named monocyte chemoattractant protein-1, MCP-1) is the most significantly downregulated gene, whose expression has been downregulated to -3.06- and -3.64-fold in the log scale after 12- and 24-h treatment, respectively. This corresponded to ~10-fold downregulation. To validate GeneChip results, we selected 30 genes that are involved in apoptosis-related pathways for RT-qPCR (reverse transcription quantitative polymerase chain reaction) analysis. Three genes are upregulated >1 in the log2 (fold change) scale after Tan IIA treatment. And eight genes including CCL2, BIRC5 (Survivin), BCL2, IL8, HSPA1A, MYC, PRKCZ, and SERPINB2 are downregulated <-1 in the log2 (fold change) scale. In total, we found that the directions of expression change for 80% of genes are consistent between GeneChip data and real-time PCR data. Confirmation of CCL2 expression regulations was also carried out using real-time quantitative PCR and ELISA.

Our results suggested that Tan IIA-induced apoptosis might result from the activation of PXR (pregnane X receptor), which suppresses the activity of NF-κB (nuclear factor kappa B) and leads to the downregulation of CCL2 expression.

Artemisinin, a sesquiterpene isolated from *Artemisia annua* L., is used in traditional Chinese medicine for the treatment of fever and chills. Artesunate (ART) is a semisynthetic derivative of artemisinin. In addition to the well-known antimalarial activity of ART, the cytotoxic action of ART against different tumor cell lines has also been identified. Youns *et al.* (2009) developed a gene expression profiling to identify novel key players involved in the cytotoxic effect of ART on pancreatic cancer cells. cDNA microarray chips containing some 7000 genes representing apoptotic, angiogenic, growth factors, anti-apoptotic, and metastasis-associated genes were used. Results showed that ART mediated growth inhibitory effects and induced apoptosis in pancreatic cancer cells through modulation of multiple signaling pathways. Several molecular targets involved in the intrinsic and extrinsic apoptotic pathways were affected after treatment with ART. Among those molecular targets are APAF1, BAX, BAK, CASP 2, CASP 3, CASP 4, CASP 5, CASP 6, CASP 8, CASP 9, and CASP 10. The finding of this study is that ART is a novel topoisomerase IIα inhibitor. Moreover, the cytotoxic effect of ART on pancreatic cancer cells could be mediated, in part, through upregulation of DDIT3 and NAG-1 genes and downregulation of PCNA and RRM2 genes.

Mastic oil from *Pistacia lentiscus* variation chia, a blend of bioactive terpenes with recognized medicinal properties, has been recently shown to exert antitumor growth activity through inhibition of cancer cell proliferation, survival, angiogenesis, and inflammatory response. To investigate molecular mechanisms triggered by mastic oil at genome-wide gene expression level, Moulos *et al.* (2009) treated Lewis lung carcinoma cells with mastic oil or DMSO (dimethyl sulfoxide) and collected RNA at five distinct time points (3–48 h) for microarray expression profiling. The results demonstrated that exposure of Lewis lung carcinomas to mastic oil caused a time-dependent alteration in the expression of 925 genes. GO (gene ontology) analysis associated expression profiles with several biological processes and functions. Among them, modifications on cell cycle/proliferation, survival, and NF-κB cascade in conjunction with concomitant

regulation of genes encoding for PTEN, E2F7, HMOX1 (upregulation), and NOD1 (downregulation) indicated some important mechanistic links underlying the anti-proliferative, proapoptotic, and anti-inflammatory effects of mastic oil. The expression profiles of Hmox1, Pten, and E2f7 genes were similarly altered by mastic oil in the majority of test cancer cell lines. Inhibition of PTEN partially reversed mastic oil effects on tumor cell growth, indicating a multi-target mechanism of action. Finally, k-means clustering organized the significant gene list in eight clusters demonstrating a similar expression profile. These results provide novel evidence on the molecular basis of tumor growth inhibition mediated by mastic oil and set a rational basis for application of genomics and bioinformatic methodologies in the screening of natural compounds with potential cancer chemopreventive activities.

The natural compound *n*-butylidenephthalide (BP), which is isolated from the chloroform extract of Angelica sinensis, has been investigated for its antitumoral effects on glioblastoma multiform (GBM) brain tumors both *in vitro* and *in vivo*. To determine the mechanism of BP-induced growth arrest and apoptosis, Lin *et al*. (2008) examined BP-induced changes in gene expression by microarray screening using human GBM brain tumor cells. Several BP-inducible genes were identified, including the nuclear receptors NOR-1, Nurr1, and Nur77. Among these genes, Nur77 is particularly interesting because it plays an important role in the apoptotic processes in various tumor cell lines. BP was able to increase Nur77 mRNA and protein expression in a time-dependent manner. After BP treatment in GBM 8401 cells, Nur77 translocated from the nucleus to the cytoplasm, the cytochrome c was released from the mitochondria, and caspase 3 became activated. Furthermore, using Nur77 promoter-luciferase assay, BP-increased Nur77 was AP1 related. Inhibition of BP-induced Nur77 expression by Nur77 short interfering RNA blocked BP-induced apoptosis in GBM 8401 cells, suggesting that the induction of Nur77 negatively affected GBM 8401 cell survival. Their results suggest that upregulation of Nur77 may explain the antitumoral activity of BP in brain tumor cells.

The same strategy has also been used in the identification of targets for many other natural products (Sertel *et al*., 2011; Yin *et al*., 2010; Tseng-Crank *et al*., 2010; Lee *et al*., 2007; Lee, Kang, and Yoon, 2010; Im *et al*., 2009; Nones *et al*., 2009; Zhang *et al*., 2004, 2005, 2009; Sun *et al*., 2008; Wang *et al*., 2008; Skommer, Wlodkowic, and Pelkonen, 2007; Bemis *et al*., 2006; Ramachandran *et al*., 2005, 2006; Zhou *et al*., 2005; Kang *et al*., 2005; Roy *et al*., 2005; Efferth, 2005; Van Erk *et al*., 2004; Chen *et al*., 2004; Iizuka *et al*., 2003; Yang *et al*., 2003; Bonham *et al*., 2002).

3.2 Identification of Medicinal Plants

One challenge to ensure the effectiveness and safety of traditional Chinese medicine is to ensure that the appropriate medicinal materials have been used in traditional Chinese medicine (TCM) products. There are several herbal and animal parts that are toxic that need to be correctly identified. Carles *et al*.(2005) designed a silicon-based DNA microarray for the identification of toxic traditional Chinese medicinal plants. Species-specific oligonucleotide probes were derived from the 5S ribosomal RNA gene of *Aconitum carmichaeli*, *Aconitum kusnezoffi*, *Alocasia macrorrhiza*, *Croton tiglium*, *Datura inoxia*, *Datura metel*, *Datura tatula*, *Dysosma pleiantha*, *Dysosma versipellis*, *Euphorbia kansui*, *Hyoscyamus niger*, *Pinellia cordata*, *Pinellia pedatisecta*, *Pinellia ternata*, *Rhododendron molle*, *Strychnos nux-vomica*, *Typhonium divaricatum*, and *Typhonium giganteum* and the leucine transfer RNA gene of *Aconitum pendulum* and *Stellera chamaejasme*. The probes were immobilized via dithiol linkage on a silicon chip. Genomic target sequences were amplified and fluorescently labeled by asymmetric polymerase chain reaction. Multiple toxic plant species were identified by parallel genotyping.

Zhang *et al*. (2003) used DNA microarray for identification of medicinal *Dendrobium* species (HerbaDendrobii) from Chinese medicinal formulations. A DNA microarray was constructed by incorporating the ITS1-5.8S-ITS2 sequences of 16 *Dendrobium* species on a glass slide. Using fluorescence-labeled ITS2 sequences as probes, distinctive signals were obtained for the five medicinal *Dendrobium* species listed in the Chinese Pharmacopoeia. The established microarray was also able to detect the presence of *Dendrobium nobile* in a Chinese medicinal formulation containing nine herbal components.

Zhu, Fushimi, and Komatsu (2008) developed a DNA microarray (PNX array) for authentication of ginseng drugs based on 18S rRNA gene

sequence. The nucleotide sequences of the nuclear 18S rRNA gene of 13 Panax taxa were determined as the specific polymorphic nucleotides. Thirty-five kinds of specific oligonucleotide were designed and synthesized as probes spotting on a decorated glass slide, which included 33 probes corresponding to the species-specific nucleotide substitutions and two probes as positive and negative controls. The species-specific probes were of 23–26 bp in length, in which the substitution nucleotide was located at the central part. Triplicate probes were spotted to warrant accuracy by correcting variation in fluorescent intensity. Partial 18S rRNA gene sequences amplified from *Panax* plants and drugs as well as their derived health foods were fluorescently labeled as targets to hybridize to the PNX array. After hybridization under optimal condition, specific fluorescent patterns were detected for each *Panax* species, and the analyzed results could be indicated as barcode patterns for quick distinction. The developed PNX array provided an objective and reliable method for the authentication of *Panax* plants and drugs as well as their derived health foods.

Asterids is one of the major plant clades comprising of many commercially important medicinal species. One of the major concerns in medicinal plant industry is adulteration/contamination resulting from misidentification of herbal plants. Mantri *et al*. (2012) constructed and validated a microarray capable of fingerprinting medicinally important species from the asterids clade. Pooled genomic DNA of 104 non-asterid angiosperm and non-angiosperm species was subtracted from pooled genomic DNA of 67 asterid species. Subsequently, 283 subtracted DNA fragments were used to construct an asterid-specific array. The validation of asterid-specific array revealed a high (99.5%) subtraction efficiency. Twenty-five asterid species (mostly medicinal) representing 20 families and nine orders within the clade were hybridized onto the array to reveal its level of species discrimination. All these species could be successfully differentiated using their hybridization patterns. A number of species-specific probes were identified for commercially important species such as tea, coffee, dandelion, yarrow, motherwort, Japanese honeysuckle, valerian, wild celery, and yerba mate. Thirty-seven polymorphic probes were characterized by sequencing. A large number of probes were novel species-specific probes whereas some of them were from chloroplast region including genes such as atpB, rpoB, and ndh that have extensively been used for fingerprinting and phylogenetic analysis of plants.

In summary, chip-based authentication of medicinal plants may be useful as an inexpensive and rapid tool for quality control and safety monitoring of herbal pharmaceuticals and neutraceuticals.

3.3 Analysis of Functional Genomics of Medicinal Plants

cDNA microarray could be a powerful tool for studying functional genomics of the medicinal plants. In order to study functional genomics of *S. miltiorrhiza*, Cui *et al*. (2007) established the first cDNA microarray for traditional Chinese herbs. A total of 4354 genes were singled out from the first 8736 PCR product and used for cDNA microarray manufacturing. Accordingly, 30-days hairy root was chosen as a reference, which was hybridized with 45- and 60-days hairy root separately. In total, 203 different expressed genes were obtained and several important functional genes related to second metabolite synthesis were cloned such as P450 and copalyl diphosphate synthase genes. Northern blot analysis showed that the result was identical with the microarray result. After sequencing, there were 172 genes clustered into 114 clusters (Unigenes). Among them, 62 unigenes had known functions, 34 unigenes were hypothetical proteins, nine unigenes were homologs with no similarity, and nine unigenes were unidentified proteins with low similarity. In all, 67 genes were classified into cellular component ontology, molecular function ontology, and biological process ontology based on GO analysis. A total of 26 genes, which represented 29 metabolic-related enzymes, were located in metabolic maps based on KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway classification.

Falara *et al*. (2008) used DNA microarrays and transcriptome analysis approaches to isolate trichome-specific genes from the medicinal plant *Cistus creticus* ssp. creticus. *C. creticus* is a plant of intrinsic scientific interest owing to the distinctive pharmaceutical properties of its resin. Labdane-type diterpenes, the main constituents of the resin, exhibit considerable antibacterial and cytotoxic activities. Analysis of trichomes isolated from different developmental stages revealed that young leaves of 1–2-cm length displayed the highest content of

labdane-type diterpenes ($80\,\text{mg}\,\text{g}^{-1}$ fresh weight), whereas trichomes from older leaves (2–3 or 3–4 cm) exhibited gradual decreased concentrations. A cDNA library was constructed to enrich transcripts from trichomes isolated from young leaves, which are characterized by high levels of labdane-type diterpenes. Custom DNA microarrays constructed with 1248 individual clones from the cDNA library facilitated transcriptome comparisons between trichomes and trichome-free tissues. In addition, gene expression studies in various Cistus tissues and organs for one of the genes highlighted as the most differentially expressed by the microarray experiments revealed a putative sesquiterpene synthase with a trichome-specific expression pattern. Full length cDNA isolation and heterologous expression in *Escherichia coli* followed by biochemical analysis led to the characterization of the produced protein as germacrene B synthase.

3.4 Analysis of MicroRNA

miRNAs are small noncoding RNAs that act as posttranscriptional gene modulators. It is reported that large numbers of miRNAs play important roles in plant development. *P. pedatisecta* is an important aroid medicinal plant that possesses the only pedate leaf blades and the largest tubers and inflorescences among all *Pinellia* species. Used as an *in situ* synthesized custom miRNA microarray, a total of 99 miRNAs belonging to 22 miRNA families were identified from *P. pedatisecta* and showed different expression levels of miRNAs in different tissues. Following the verification for the presence of the miRNAs through reverse transcription polymerase chain reaction (RT-PCR) and the quantitative reverse transcription polymerase chain reaction (qRT-PCR), 14 miRNAs were detected to validate the microarray results (Wang *et al*., 2012).

Curcumin (diferuloylmethane), a naturally occurring flavinoid derived from the rhizome of *Curcuma longa*, has shown its inhibitory anticancer, antioxidant, anti-inflammatory, anti-proliferative, and proapoptotic activities *in vitro* and *in vivo* preclinical studies. Yet, the mechanisms underlying the antitumor activity of curcumin have not, however, been completely delineated. Sun *et al*. (2008) developed an oligonucleotide microarray chip to profile miRNA expressions in human pancreatic cells treated with

curcumin. Their results indicated that curcumin alters miRNA expression in human pancreatic cells, upregulating miRNA-22 and downregulating miRNA-199a*, as confirmed by TaqMan real-time PCR. Upregulation of miRNA-22 expression by curcumin or by transfection with miRNA-22 mimetics in the PxBC-3 pancreatic cancer cell line suppressed expression of its target genes SP1 transcription factor (SP1) and estrogen receptor 1 (ESR1), whereas inhibiting miRNA-22 with antisense enhanced SP1 and ESR1 expression.

Tang, Zhang, and Du (2010) used an oligonucleotide microarray chip to profile miRNA expressions in A549/DDP cells treated with and without curcumin. Their aim is to illustrate whether curcumin could promote the apoptosis of A549/DDP cells through regulating the expression of miR-186*. Curcumin, a natural compound, is derived from the rthizom of *C. longa*. *In vitro* and *in vivo* preclinical studies have shown its anti-inflammatory, antioxidant, anticancer activities, and so on. miR-186*, which was found by microarray technology, was highly expressed in lung carcinoma cells A549/DDP. An oligonucleotide microarray chip was used to profile miRNA expressions in A549/DDP cells treated with and without curcumin. The significant differentially expressed miRNA, which was selected from microarray chip, was validated by quantitative real-time PCR. Ultimately, the remarkably expressed miRNA modulated the apoptosis assaying by flow cytometry experiments and the survival rate was measured by MTT method. The microarray chip results demonstrated the following: curcumin altered the expression level of miRNAs compared with untreated control in A549/DDP cell line; miR-186* was significantly downregulated after curcumin treatment, which was confirmed by quantitative real-time PCR. Downregulation of miR-186* expression by curcumin elevated the apoptosis, and the survival rate of A549/DDP cells decreased; however, upregulation of miR-186* expression by transfection its mimics restrained the apoptosis, the survival rate of A549/DDP cells increased, which were assayed by flow cytometry experiments and MTT method.

Ginsenoside-Rg1, one of the active components of ginseng, has been confirmed as an angiogenesis inducer. Using miRNA microarray analysis, a total of 17 (including miR-214) and five miRNAs were found to be downregulated or upregulated by Rg1 in human umbilical vein endothelial cells (HUVECs), respectively (Chan *et al*., 2009).

4 EXPERIMENTAL DESIGN AND PROTOCOLS OF USING DNA MICROARRAY TECHNOLOGIES

4.1 Experimental Design

Like all the research projects, researchers have to be clear about the goal of the experiments. Only when the goal is clear, one can design appropriate experiments. Owing to the biological complexity of gene expression, only when an experiment is appropriately designed, statistically and biologically valid conclusions can be drawn. This is particularly important when the expression levels of thousands of genes are measured simultaneously. There are three main elements to consider when designing a microarray experiment. First, replication of the biological samples is essential for drawing conclusions from the experiment. The biological replicates can be RNA extracted from different biological samples. Second, technical replicates (two RNA samples obtained from each experimental unit) help ensure precision and allow for testing differences within treatment groups. Usually, technical differences at two different levels should be considered. For example, the difference from two different batches of RNA extraction from the same biological samples can help determine variations between different RNA extraction runs. Furthermore, for a novel platform, the variations between different spots or on different chips also need to be monitored. Even for commercial platforms, the intra-chip variations have to be tested to be minimal. And this is certainly the reason why commercial platforms are much more in favor compared with those made in individual laboratory. Researchers are often recommended to consult with their in-house statistician before initiating any microarray experiments (Churchill, 2002).

4.2 Standardization

As described below, there is a long list of providers for technologies related to microarray, from providing particular equipments to providing complete solutions. One of the difficulties is to compare results obtained from various platform fabrication, assay protocols, and analysis methods. This presents an interoperability problem for bioinformatics. During the past years, various data standards have been proposed in an attempt to ease the exchange and analysis of data produced with non-proprietary chips.

The earliest effort is the "Minimum Information About a Microarray Experiment" (MIAME) checklist (http://www.mged.org/Workgroups/MIAME/miame.html). It helps define the level of detail that should exist in any microarray experiments and is being adopted by many journals as a requirement for the submission of papers incorporating microarray results. However, MIAME does not describe the format for the information, so while many formats can support the MIAME requirements, no format permits verification of complete semantic compliance.

The "MicroArray Quality Control (MAQC) Project" (http://www.fda.gov/ScienceResearch/BioinformaticsTools/MicroarrayQualityControlProject/default.htm) is being conducted by the US Food and Drug Administration (FDA) to develop standards and quality control metrics that will eventually allow the use of MicroArray data in drug discovery, clinical practice, and regulatory decision-making. Results obtained from their studies have been widely used in the selection of microarray platforms.

The MGED (Microarray Gene Expression Data) Society (http://www.mged.org/), which has recently changed their name to Functional Genomics Data (FGED), has developed many standards for biological research data quality, annotation, and exchange. Furthermore, they facilitate the creation and use of software tools that build on these standards. In this way, it supports the integration of multiple sources of microarray data and potentially leads to novel discovery.

4.3 Experimental Procedures

Here, we would like to describe in detail the procedures to conduct microarray experiments using a homemade cDNA microarray platform (Figure 1a). This is most applicable to medicinal plant research because no commercial chips are available in most cases. Then, we will briefly describe several commercially available platforms. The users are recommended to consult the detailed technical manuals for each of the platforms. Figure 1b explains the principle of oligo array using Affymetrix platform as an example.

4.3.1 Homemade cDNA Microarray

This set of protocols is intended to serve as a basic introduction to make and use cDNA microarrays for those embarking on this path. There are three fundamental steps.

1. *cDNA Amplification and Printing*. This step makes the cDNA microarray. One needs to obtain an inventory of cDNA bacterial clones that represent the genes to study. Plasmid templates are made from these clones and used as PCR substrates to produce DNA representations of the EST (expressed-sequence tag) inserts. The PCR products are then purified and spotted onto poly-L-lysine-coated microscope slides.
2. *RNA Extraction and Labeling*. This step obtains the samples to test. RNA is extracted from the cell samples to be examined, purified, and used as the substrate for reverse transcription in the presence of fluor-derivatized nucleotides. This procedure provides the tagged representations of the mRNA pools of the samples that will be used in the next step.
3. *Hybridization and Data Extraction*. This process covers the steps in which fluor-labeled cDNAs hybridize to their complements on the microarray, and the resulting localized concentrations of fluorescent molecules are quantified.

The three steps are described in more detail below. Owing to the space limitation, we omitted the details for steps that are not unique to cDNA microarray experiments such as RNA extraction and labeling, as there are many effective protocols to achieve the same goal.

4.3.1.1 Fabrication of cDNA Microarray

This protocol describes the steps required to produce a cDNA microarray. Gene-specific DNA is produced by PCR amplification of purified template plasmid DNAs from cloned ESTs. The PCR product is purified by ethanol precipitation, thoroughly resuspended in 3× SSC, and printed onto a poly-L-lysine-coated slide. This step can be further divided into the following 10 steps. Detailed procedures are given for slide coating, blocking, and printing, which are the unique steps.

1. EST clone growth.
2. Isolate plasmid templates.
3. Amplify EST inserts.

4. Check PCR products by agarose gel electrophoresis of ESTs.
5. Purify PCR products.
6. Resuspend the PCR products.
7. Check PCR resuspension for yield by fluorometric determination of DNA concentration.
8. Slide coating.

Slides coated with poly-L-lysine have a surface that is both hydrophobic and positively charged. The hydrophobic character of the surface minimizes spreading of the printed spots, and the charge appears to help position the DNA on the surface in a way that makes cross-linking more efficient.

(i) Place slides into 50-slide racks and place racks in glass tanks with 500 mL of cleaning solution. Gold Seal slides are highly recommended, as they have been found to consistently have low levels of autofluorescence.
(ii) Place tanks on platform shaker for 2 h at 60 rpm.
(iii) Pour out cleaning solution and wash in H_2O for 3 min. Repeat wash four times.
(iv) Transfer slides to 30-slide plastic racks and place into small plastic boxes for coating.
(v) Submerge slides in 200-mL poly-L-lysine solution per box.
(vi) Place slide boxes on platform shaker for 1 h at 60 rpm.
(vii) Rinse slides three times with H_2O.
(viii) Submerge slides in H_2O for 1 min.
(ix) Spin slides in centrifuge for 2 min at 400 × g and dry slide boxes used for coating.
(x) Place slides back into slide box used for coating and let stand overnight before transferring to new slide box for storage.
(xi) Allow slides to age for two weeks on the bench, in a new slide box, before printing on them. The coating dries slowly, becoming more hydrophobic with time.

9. Slide blocking

(i) Place slides, printed face up, in casserole dish and cover with cling wrap. Expose slides to a 450-mJ dose of ultraviolet irradiation in the Stratalinker.
(ii) Transfer slides to a 30-slide stainless steel rack and place rack into a small glass tank.
(iii) Dissolve 6.0-g succinic anhydride in 325-mL 1-methyl-2-pyrrolidinone in a glass beaker by stirring with a stir bar.

(iv) Add 25-mL 1 M sodium borate buffer (pH 8.0) to the beaker. Allow the solution to mix for a few seconds, then pour rapidly into glass tank with slides.

(v) Place the glass tank on a platform shaker in a fume hood for 20 min.

(vi) While the slides are incubating on the shaker, prepare a boiling H_2O bath to denature the DNA on the slides.

(vii) After the slides have incubated for 20 min, transfer them into the boiling H_2O bath. Immediately turn off the heating element after submerging the slides in the bath. Allow slides to stand in the H_2O bath for 2 min.

(viii) Transfer the slides into a glass tank filled with 100% ethanol and incubate for 4 min.

(ix) Remove the slides and centrifuge at 400 rpm for 3 min in a horizontal microtiter plate rotor to dry the slides.

(x) Transfer slides to a clean, dust-free slide box and let stand overnight before hybridizing.

10. Printing

(i) Pre-clean the print pens according to the manufacturer's specification.

(ii) Load the printer slide deck with poly-L-lysine-coated slides.

(iii) Thaw the plates containing the purified EST PCR products and centrifuge briefly, 2 min, at 1000 rpm in a horizontal microtiter plate rotor to remove condensation and droplets from the seals before opening.

(iv) Transfer 5–10 μL of the purified EST PCR products to a plate that will serve as the source of solution for the printer.

(v) Run a repetitive test print on the first slide. In this operation, the pens are loaded with the DNA solution, and then the pens serially deposit this solution on the first slide in the spotting pattern specified for the print.

(vi) If one or more of the pens is not performing at the desired level, reclean or substitute another pen and test again.

(vii) If all pens are performing, carry out the full print.

4.3.1.2   RNA Extraction

This protocol details the methods used to extract RNA from cells, purify the RNA by a combination of phase extraction and chromatography, and prepare a labeled cDNA copy of the message fraction of the purified RNA. The protocol also describes the process of making fluorescent cDNA representations of the message pools within the isolated total RNA pools. This is accomplished by using the pure total RNA as a substrate for reverse transcription in the presence of nucleotides derivatized with either a Cy3 or a Cy5 fluorescent tag.

4.3.1.3   Hybridize Fluorescent cDNA to Slide

1. Determine the volume of hybridization solution required. The rule of thumb is to use 0.033 μL for each square millimeter of slide surface area covered by the coverslip used to cover the array.

2. For a 40-μL hybridization reaction as example, pool the Cy3- and Cy5-labeled cDNAs into a single 0.2-mL thin wall PCR tube and adjust the volume to 30 μL by either adding DEPC H_2O or removing water in a SpeedVac. If using a vacuum device to remove water, do not use high heat or heat lamps to accelerate evaporation. The fluorescent dyes could be degraded.

3. Set up the hybridization components.

4. Mix the components well by pipetting, heat at 98°C for 2 min in a PCR cycler, cool quickly to 25°C, and add 0.6 μL of 10% SDS (sodium dodecyl sulfate).

5. Centrifuge for 5 min at 14,000 × g. The fluor-labeled cDNAs have a tendency to form small, very fluorescent, aggregates that result in bright, punctuate background on the array slide. Hard centrifugation will pellet these aggregates, allowing you to avoid introducing them to the array.

6. Apply the labeled cDNA to a 24 mm × 50 mm glass coverslip and then touch with the inverted microarray.

7. Place the slide in a microarray hybridization chamber; add 5 μL of 3× SSC in the reservoir, if the chamber provides one, or at the scribed end of the slide; and seal the chamber. Submerge the chamber in a 65°C water bath and allow the slide to hybridize for 16–20 h.

8. Remove the hybridization chamber from the water bath, cool, and carefully dry off. Unseal the chamber and remove the slide.

9. Place the slide, with the coverslip still affixed, into a Coplin jar filled with 0.5× SSC/0.01% SDS wash buffer. Allow the coverslip to fall from the slide and then remove the coverslip from the

jar with a forceps. Allow the slide to wash for 2–5 min.

10. Transfer the slide to a fresh Coplin jar filled with 0.06× SSC. Allow the slide to wash for 2–5 min.

11. Transfer the slide to a slide rack and centrifuge at low rpm (700–1000) for 3 min in a clinical centrifuge equipped with a horizontal rotor for microtiter plates.

4.3.1.4   Detection of Hybridization Signal

Fluorescent hybridization signals are scanned and detected using laser confocal devices. Data are filtered and processed.

4.3.2   Commercial DNA Microarray Platforms

Owing to the difficulty in making "in-house" microarray, many commercial microarray analysis platforms have been developed, for example, Affymetrix, Agilent, Applied Biosystems (ABI), Amersham (now GE Healthcare), Compugen (now Sigma-Genosys), Mergen, and MWG BioTech (now Ocimum Biosolutions). However, to our knowledge, these platforms do not provide any premade chips for medicinal plants. For each platform, there are series of products that meet the needs for model organisms to be studied and there are detailed technical manuals. The general principle is the same as those for the cDNA microarray. As a result, they will not be described in more detail here. Instead, the users are advised to review a recent paper that has carried out a systematic comparison of these platforms (Liu *et al.*, 2012). Furthermore, many web resources such as Microarray Station (http://www.microarraystation.com/) provide extensive information on microarray-related vendors, software tools, and protocols.

4.4   Statistical Analysis

The size of data produced by various microarray platforms is usually very large, and analytical precision is influenced by a number of variables. Statistical challenges include effects of background noise and appropriate normalization of the data. Normalization methods may be suited to particular platforms. In the case of some commercial platforms, the analysis methods may be proprietary and not available for open inspection. Various statistical methods can be applied at the steps specified below:

1. *Image Analysis*. Various methods are needed to determine the grid and recognize the spots of the scanned image and to remove or mark poor-quality and low intensity features.
2. *Data Processing*. Various methods are need to for background subtraction (based on global or local background), determination of spot intensities and intensity ratios, visualization of data (e.g., see MA plot), and log-transformation of ratios, global or local normalization of intensity ratios, and segmentation into different copy number regions using step detection algorithms (Little and Jones, 2011).
3. *Identification of Statistically Significant Changes*. Commonly used statistical algorithms include *t*-test, ANOVA, Bayesian method (Ben-Gal *et al.*, 2005), and Mann–Whitney test methods tailored to microarray data sets, which take into account multiple comparisons (Leung and Cavalieri, 2003) or cluster analysis (Priness, Maimon, and Ben-Gal, 2007). These methods assess statistical power based on the variations present in the data and the number of experimental replicates, and can help minimize type I and type II errors in the analyses (Wei, Li, and Bumgarner, 2004).
4. *Network-Based Methods*. After the identification of statistically significant changes, other statistical methods can be used to reveal the underlying structure of gene networks, representing either associative or causative interactions or dependencies among gene products (Dehmer and Emmert-Streib, 2008).

5   CONCLUSIONS AND FUTURE DIRECTION

DNA microarrays are a powerful and easy-to-use genomic tool. The genomic-wide data provided after microarray applications can provide potential information that helps find the causes of disease, the mechanism of drug action, and the discovery of gene products that are targets for therapy in various diseases. With the use of this approach, novel molecular targets and new therapeutic options can be identified. In this chapter, we have reviewed recent experiments using microarray technology involving medicinal plants derived from TCM as examples.

Furthermore, we showed that expression analysis was effective in identifying novel pathways and molecular targets mediating their effects.

However, DNA microarray technology has its own limitations. DNA microarrays only measure expression levels for those genes for which a probe, either a clone or a sequence, is available and, therefore, should be considered a targeted approach to expression profiling. At present stage, this poses a problem, as the gene sequences for most medicinal plants are unknown, and consequently, DNA microarray technology is not suitable to profile their gene expression. With the rapid progress made in Next Generation DNA Sequencing (NGS) technology, many techniques based on NGS have been developed. One of them, RNA-SEQ (ref) has posed to replace DNA microarray technology in many of its applications. So, the readers are advised to fully explore the various options before embarking on any microarray-based expression profiling studies.

6 RELATED ARTICLES

Transcriptome Analysis of Medicinal Plants with Next-Generation Sequencing Technologies; Multivariate data analysis.

REFERENCES

Adomas, A., Heller, G., Olson, A., *et al.* (2008) *Tree Physiol.*, **28**(6), 885–897.

Bemis, D. L., Capodice, J. L., Gorroochurn, P., *et al.* (2006) *Int. J. Oncol.*, **29**(5), 1065–1073.

Ben-Gal, I., Shani, A., Gohr, A., *et al.* (2005) *Bioinformatics*, **21**(11), 2657–2666.

Bonham, M., Arnold, H., Montgomery, B., *et al.* (2002) *Cancer Res.*, **62**(14), 3920–3924.

Callaway, E. (2011) *Nature* [Internet]. 2011 Sep 16 [cited 2011 Nov 17].

Carles, M., Cheung, M. K., Moganti, S., *et al.* (2005) *Planta Med.*, **71**(6), 580–584.

Castel, D., Pitaval, A., Debily, M. A., *et al.* (2006) *Drug Discov. Today*, **11**(13–14), 616–622.

Chaga, G. S. (2008) *Methods Mol. Biol.*, **441**, 129–151.

Chan, L. S., Yue, P. Y., Mak, N. K., *et al.* (2009) *Eur. J. Pharm. Sci.*, **38**(4), 370–377.

Chen, H. W., Yu, S. L., Chen, J. J., *et al.* (2004) *Mol. Pharmacol.*, **65**(1), 99–110.

Churchill, G. A. (2002) *Nat. Genet.*, **32**(Suppl), 490–495.

Collas, P. (2010) *Mol. Biotechnol.*, **45**(1), 87–100.

Cui, G. H., Mao, Y., Huang, L. Q., *et al.* (2007) *Zhongguo Zhong Yao Za Zhi*, **32**(14), 1393–1395.

Dehmer, M., and Emmert-Streib, F. (2008) *Analysis of Microarray Data: A Network-Based Approach*, Vch Verlagsgesellschaft Mbh, Hoboken, NJ.

Delgado, JY., Coba, M., Anderson, CN., *et al.* (2007) *J. Neurosci.*, **27**(48), 13210–13221.

Efferth, T. (2005) *Oncol. Rep.*, **13**(3), 459–463.

Falara, V., Fotopoulos, V., Margaritis, T., *et al.* (2008) *Plant Mol. Biol.*, **68**(6), 633–651.

Foong, YM., Fu, J., Yao, SQ., *et al.* (2012) *Curr. Opin. Chem. Biol.*, **16**(1–2), 234–242.

Garren, H., Robinson, WH., Krasulova, E., *et al.* (2008) *Ann. Neurol.*, **63**(5), 611–620.

Gaseitsiwe, S., Valentini, D., Mahdavifar, S., *et al.* (2010) *Clin. Vaccine Immunol.*, **17**(1), 168–175.

Hacia, JG., Fan, JB., Ryder, O., *et al.* (1999) *Nat. Genet.*, **22**(2), 164–167.

Howbrook, DN., van der Valk, AM., O'Shaughnessy, MC., *et al.* (2003) *Drug Discov. Today*, **8**(14), 642–651.

Iizuka, N., Oka, M., Yamamoto, K., *et al.* (2003) *Int. J. Cancer*, **107**(4), 666–672.

Im, R., Mano, H., Matsuura, T., *et al.* (2009) *J. Ethnopharmacol.*, **121**(2), 234–240.

Johnson, JM., Castle, J., Garrett-Engele, P., *et al.* (2003) *Science*, **302**(5653), 2141–2144.

Kang, JX., Liu, J., Wang, J., *et al.* (2005) *Carcinogenesis*, **26**(11), 1934–1939.

Kindrachuk, J., Arsenault, R., Kusalik, A., *et al.* (2012) *Mol. Cell Proteomics*, **11**(6), DOI: 10.1074/mcp.M111.015701.

Lee, S. B., Cha, K. H., Selenge, D., *et al.* (2007) *Biol. Pharm. Bull.*, **30**(6), 1074–1079.

Lee, H., Kang, R., Yoon, Y., (2010) *J. Ethnopharmacol.*, **127**(3), 709–717.

Leung, Y. F. and Cavalieri, D. (2003) *Trends Genet.*, **19**(11), 649–659.

Lin, P. C., Chen, Y. L., Chiu, SC., *et al.* (2008) *J. Neurochem.*, **106**(3), 1017–1026.

Lin, J., Bardina, L., Shreffler, W. G., *et al.* (2009) *J. Allergy Clin. Immunol.*, **124**(2), 315–322, 322, e311–313.

Linnebacher, M., Lorenz, P., Koy, C., *et al.* (2012) *Anal. Bioanal. Chem.*, **403**(1), 227–238.

Little, M. A., and Jones, N. S. (2011) *Proc. Math. Phys. Eng. Sci.*, **467**(2135), 3115–3140.

Liu, C., Li, J., Wang, L., *et al.* (2012) *BMC Complement. Altern. Med.*, **12**, 5.

Lizcano, J. M., Deak, M., Morrice, N., *et al.* (2002) *J. Biol. Chem.*, **277**(31), 27839–27849.

Mantri, N., Olarte, A., Li, C. G., *et al.* (2012) *PLoS One*, **7**(4), e34873.

Maskos, U., and Southern, E. M. (1992) *Nucleic Acids Res.*, **20**(7), 1679–1684.

Moran, G., Stokes, C., Thewes, S., *et al.* (2004) *Microbiology*, **150**(Pt 10), 3363–3382.

Moulos, P., Papadodima, O., Chatziioannou, A., *et al.* (2009) *BMC Med. Genomics*, **2**, 68.

Nones, K., Dommels, Y. E., Martell, S., *et al.* (2009) *Br. J. Nutr.*, **101**(2), 169–181.

Panse, S., Dong, L., Burian, A., *et al*. (2004) *Mol. Divers.*, **8**(3), 291–299.

Phizicky, E., Bastiaens, P. I., Zhu, H., *et al*. (2003) *Nature*, **422**(6928), 208–215.

Pollack, J. R., Perou, C. M., Alizadeh, A. A., *et al*. (1999) *Nat. Genet.*, **23**(1), 41–46.

Priness, I., Maimon, O., and Ben-Gal, I. (2007) *BMC Bioinformatics*, **8**, 111.

Ramachandran, C., Rodriguez, S., Ramachandran, R., *et al*. (2005) *Anticancer Res*, **25**(5), 3293–3302.

Ramachandran, C., Nair, P. K., Alamo, A., *et al*. (2006) *Int. J. Cancer*, **119**(10), 2443–2454.

Rivas, L. A., Garcia-Villadangos, M., Moreno-Paz, M., *et al*. (2008) *Anal. Chem.*, **80**(21), 7970–7979.

Roy, S., Khanna, S., Shah, H., *et al*. (2005) *DNA Cell Biol.*, **24**(4), 244–255.

Sertel, S., Eichhorn, T., Bauer, J., *et al*. (2011) *J. Nutr. Biochem.*, **23**(8), 875–884.

Shaw, P-C, Wong, K-L, Chan, A. W., *et al*. (2009) *Chin. Med.*, **4**(1), 21.

Shingara, J., Keiger, K., Shelton, J., *et al*. (2005) *RNA*, **11**(9), 1461–1470.

da Silva, J. A. T. and Yücel, M. (2008) *Genes Genomes Genomics*, **2**(1), 14–48.

Skommer, J., Wlodkowic, D., and Pelkonen, J. (2007) *Exp. Hematol.*, **35**(1), 84–95.

Smith, M. G., Jona, G., Ptacek, J., *et al*. (2005) *Mech. Ageing Dev.*, **126**(1), 171–175.

Stears, R. L., Martinsky, T., and Schena, M. (2003) *Nat. Med.*, **9**(1), 140–145.

van Steensel, B., and Henikoff, S. (2000) *Nat. Biotechnol.*, **18**(4), 424–428.

Sun, M., Estrov, Z., Ji, Y., *et al*. (2008) *Mol. Cancer Ther.*, **7**(3), 464–473.

Tang, N., Zhang, J., and Du, Y. (2010) *Zhongguo Fei Ai Za Zhi*, **13**(4), 301–306.

Thiele, A., Krentzlin, K., Erdmann, F., *et al*. (2011) *J. Mol. Biol.*, **411**(4), 896–909.

Tomaras, G. D., Binley, J. M., Gray, E. S., *et al*. (2011) *J. Virol.*, **85**(21), 11502–11519.

Tseng-Crank, J., Sung, S., Jia, Q., *et al*. (2010) *J. Diet. Suppl.*, **7**(3), 253–272.

Tzankov, A., Went, P., Zimpfer, A., *et al*. (2005) *Exp. Gerontol.*, **40**(8–9), 737–744.

Van Erk, M. J., Teuling, E., Staal, Y. C., *et al*. (2004) *J. Carcinog.*, **3**(1), 8.

Wang, C. Y., Staniforth, V., Chiao, M. T., *et al*. (2008) *BMC Genomics*, **9**, 479.

Wang, B., Dong, M., Chen, W., *et al*. (2012) *Gene*, **498**(1), 36–40.

Wei, C., Li, J., and Bumgarner, R. E. (2004) *BMC Genomics*, **5**, 87.

Yang, S. H., Kim, J. S., Oh, T. J., *et al*. (2003) *Int. J. Oncol.*, **22**(4), 741–750.

Yin, S. Y., Wang, W. H., Wang, B. X., *et al*. (2010) *BMC Genomics*, **11**, 612.

Youns, M., Efferth, T., Reichling, J., *et al*. (2009) *Biochem. Pharmacol.*, **78**(3), 273–283.

Zhang, Y. B., Wang, J., Wang, Z. T., *et al*. (2003) *Planta Med.*, **69**(12), 1172–1174.

Zhang, J. P., Ying, K., Xiao, Z. Y., *et al*. (2004) *Int. J. Cancer*, **108**(2), 212–218.

Zhang, P., Mao, Y. C., Sun, B., *et al*. (2005) *Ai Zheng*, **24**(4), 454–460.

Zhang, J., Zuo, G., Bai, Q., *et al*. (2009) *Planta Med.*, **75**(4), 396–398.

Zhou, Z., Tan, H. L., Xu, B. X., *et al*. (2005) *Pharmacol. Rep.*, **57**(6), 818–823.

Zhu, S., Fushimi, H., and Komatsu, K. (2008) *J. Agric. Food Chem.*, **56**(11), 3953–3959.

38

Small RNAs

Shanfa Lu

Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, PR China

1 INTRODUCTION

Small RNAs (sRNAs) are short, noncoding RNAs with significant regulatory roles in many organisms, including animals, plants, and viruses (Kozomara and Griffiths-Jones, 2011). In plants, sRNAs not only regulate plant development but also are involved in plant response to biotic and abiotic stresses through direct cleavage of transcripts, translational inhibition, or epigenetic modification (Aukerman and Sakai, 2003; Brodersen *et al*., 2008; Chen, 2004; Jones-Rhoades, Bartel, and Bartel, 2006; Llave *et al*., 2002; Mallory and Vaucheret, 2006; Rhoades *et al*., 2002; Wu *et al*., 2010). They are abundant in cells and can be classified into microRNAs (miRNAs) and endogenous small interfering RNAs (siRNAs) based on the distinct origins and biogenesis pathways.

Plant miRNAs are usually about 21 nucleotides (nt) in length. They are processed mostly from primary miRNAs (pri-miRNAs), which are transcribed from miRNA loci by polymerase II and can fold into internal hairpin structures (Voinnet, 2009). Cleavage of pri-miRNAs to produce miRNA precursors (pre-miRNAs) is carried out by Dicer-like 1 (DCL1) in interaction with several other proteins (Fang and Spector, 2007; Kurihara, Takashi, and Watanabe, 2006; Park *et al*., 2002). Further processes of pre-miRNAs include cleavage by DCL1 to produce ∼21-bp RNA duplexes bearing 2 nt overhangs at both 3′ ends (Park *et al*., 2005) and unwinding of the duplexes to release single-stranded mature miRNAs.

Plant endogenous siRNAs are a large class of sRNAs and can be classified into several subclasses, such as trans-acting siRNAs (ta-siRNAs), natural-antisense-transcript-derived siRNAs (NAT-siRNAs), and heterochromatic-siRNAs (hc-siRNAs) (Vazquez, Legrand, and Windels, 2010). ta-siRNAs are phased 21-nt RNA molecules originated from nonprotein-coding ta-siRNA genes (*TASs*) (Rajagopalan *et al*., 2006; Vazquez *et al*., 2004b). The 21-nt phase for *TAS* processing is set by various miRNAs, such as miR173, miR390, and miR828 (Allen *et al*., 2005; Rajagopalan *et al*., 2006). NAT-siRNAs are derived from the overlapping regions of two RNAs that are partially complementary to each other and can be classified into *cis*-NAT-siRNAs and *trans*-NAT-siRNAs based on the origin of two overlapping RNAs (Borsani *et al*., 2005; Katiyar-Agarwal *et al*., 2006, 2007; Wang, Chua, and Wang, 2006; Zhou *et al*., 2009). hc-siRNAs are typically 24-nt sRNAs and are also known as repeat-associated siRNAs (ra-siRNAs) and transposable element-derived siRNAs (TE-siRNAs), as they usually arise from repeats and transposable elements (Hamilton *et al*., 2002; Kasschau *et al*., 2007; Lippman and Martienssen, 2004).

So far, many methods have been developed to isolate and identify sRNAs and to elucidate their functions. In this chapter, we briefly review various techniques currently used for the analysis of plant sRNAs and provide two protocols, one for sRNA cloning and the other for sRNA Northern hybridization.

2   sRNA ISOLATION AND SEQUENCING

Total RNA used for sRNA isolation is usually extracted from plant tissues using the TRIzol reagent. Various methods for sRNA library construction have been developed (Elbashir *et al.*, 2001; Lagos-Quintana *et al.*, 2002; Lau *et al.*, 2001; Lu *et al.*, 2005; Lu, Sun, and Chiang, 2009). sRNAs can be cloned into a TA vector and then sequenced using the conventional Sanger sequencing method, which will produce a limited number of sRNA sequences (Lu *et al.*, 2005, 2007; Lu, Sun, and Chiang, 2008). Alternatively, the sRNA library can be directly sequenced using the high-throughput sequencing platforms, such as the 454 (Roche), Solexa (Illumina), and SoLiD (ABI) platforms, which can produce millions of sequence reads (Lu *et al.*, 2006, 2009; Ma, Coruh, and Axtell, 2010; Norden-Krichmar *et al.*, 2011; Szittya *et al.*, 2008).

3   COMPUTATIONAL CHARACTERIZATION OF sRNAS

In order to characterize sRNAs, adapters will be removed and low-quantity sequences will be filtered from raw sRNA sequence data using the PHRED and CROSS MATCH programs (Ewing and Green, 1998). sRNAs with sizes between 18 and 30 nucleotides can be annotated locally by aligning with the whole genome sequences, EST data sets, or known miRNA precursors using PatScan or SOAP2 (Dsouza, Larsen, and Overbeek, 1997; Li *et al.*, 2009). sRNAs can also be analyzed by using some web tools, such as the UEA sRNA toolkit (http://srna-tools.cmp.uea.ac.uk/plant/cgi-bin/srna-tools.cgi), or by BLAST analysis against the GenBank (http://www.ncbi.nlm.nih.gov/BLAST). Known mature miRNAs and miRNA precursors can be downloaded from miRBase (Kozomara and Griffiths-Jones, 2011). Secondary structures of sRNA-matched genomic sequences or ESTs and their flanking sequences can be predicted by the mfold program (Zuker, 2003). The other useful software is MIREAP, which can be used to predict novel miRNAs from the sRNA set (http://sourceforge.net/projects/mireap/).

Because the fundamental defining feature of plant miRNAs is the precise excision of a miRNA/miRNA* duplex from the stem of a stem-loop precursor, it is very important to apply correct parameters to miRNA annotation (Meyers *et al.*, 2008). In recent studies, we successfully predicted several novel miRNAs from a set of *Panax ginseng* sRNAs using the MIREAP software with following parameters: minimal miRNA length (18), maximal miRNA length (26), minimal miRNA (reference) length (20), maximal miRNA reference length (24), uniqueness of miRNA (280), maximal energy (-30), minimal space (5), maximal space (450), minimal mature pair (16), maximal mature bulge (3), maximal duplex asymmetry (4), and flank sequence length (20) (Wu *et al.*, 2012).

To reduce false-positive prediction of plant miRNAs, one may perform a manual check of the stem-loop structures predicted from the mfold program and the MIREAP software. The criteria applied to manual miRNA annotation may include the following: (1) the free energy (dG) value of a single-stranded, hairpin precursor is less than -30, (2) there are four or fewer mismatched miRNA bases, (3) there are no more than two asymmetric bulges within the miRNA/miRNA* duplex, and (4) the minimal folding free energy index (MFEI) calculated as described by Zhang *et al.* (2006) is at least 0.85.

4   COMPUTATIONAL PREDICTION OF sRNA TARGETS

Many sRNAs are involved in gene silencing through direct cleavage of their targets at the region with perfect or near-perfect complementarity to the corresponding sRNA (Carrington and Ambros, 2003; Llave *et al.*, 2002; Reinhart *et al.*, 2002; Rhoades *et al.*, 2002; Tang *et al.*, 2003). On the basis of the complementarity between sRNAs and their targets, various computational programs have been developed to predict the targets of sRNA (Allen *et al.*, 2005; Dai and Zhao, 2011; Lu *et al.*, 2005; Sun *et al.*, 2011; Zhang, 2005). These programs apply the penalty scoring scheme developed by Jones-Rhoades and Bartel (2004) to identify sRNA targets. The total score for mismatched patterns in the sRNA : target duplexes is calculated based on a 20-base sequence window, with 0 points being assigned to each complementary pair, 0.5 points to each G : U wobble, 1 point to each non-G : U wobble mismatch, and 2 points to each bulged nucleotide in either RNA strand. In the

programs developed by Allen *et al.* (2005) and Sun *et al.* (2011), the penalty score is doubled for all mismatch types appeared between the 2nd and the 13th nucleotides of the sRNA, whereas in the programs developed by Zhang (2005) and Dai and Zhao (2011), any mismatch other than G : U wobble in positions 2–7 nucleotides of the sRNA is further penalized 0.5 points in the score. In each case, a cutoff score of 3.0 can predict targets with high confidence. Using a lower cutoff score may result in a lower false-positive prediction, whereas the prediction coverage will be decreased. On the contrary, using a higher cutoff score may result in a higher false-positive prediction, whereas the prediction coverage will be increased.

5 MAPPING OF sRNA-GUIDED CLEAVAGE SITES

Computationally predicted cleavage of miRNA targets can be experimentally validated by the modified 5′RNA ligase-mediated RACE (RLM-RACE) method (Lu *et al.*, 2005; Vazquez *et al.*, 2004a). In this method, total RNAs are extracted from plant tissues and then poly(A) RNAs are enriched. The 5′-RACE was conducted using the GeneRacer kit (Invitrogen). Two-round PCRs are performed. The first round of PCR uses the GeneRacer 5′primer and the nesting gene-specific primers. The second round uses the GeneRacer 5′ nested primer and the gene-specific nested primers. The PCR-amplified products are gel-purified, cloned, and sequenced. Cleavage sites are determined by mapping 5′-ends of the PCR products to the targets. This method has been widely used for sRNA characterization. Additionally, an approach called parallel analysis of RNA ends (PARE) for high-throughput identification of the targets of sRNAs has been developed (Addo-Quaye *et al.*, 2009; German *et al.*, 2009). This approach comprises a modified 5′-rapid amplification of cDNA ends, deep sequencing of 3′ cleavage products of mRNA, and bioinformatic analysis (German *et al.*, 2009). The protocol of this approach has been published in Nature Protocol (German *et al.*, 2009).

6 ANALYSIS OF sRNA EXPRESSION

Analysis of the spatiotemporal expression patterns of sRNAs is very important for understanding the regulatory roles of sRNAs in plant growth and development. Several sRNA-specific techniques have been developed for determining sRNA expression levels. It includes sRNA Northern hybridization, sRNA-specific quantitative real-time PCR (qRT-PCR), and sRNA microarrays.

Northern hybridization is a semi-quantification method and can be used for analyzing a few RNA samples (Hutvágner *et al.*, 2001; Lu *et al.*, 2005). It separates sRNAs from total RNAs by gel electrophoresis. sRNAs are then transferred onto a nylon membrane by an electrophoretic transfer cell and hybridized with DNA oligonucleotide probes labeled by [γ-^{32}P]ATP. The sequence of probes is complementary to the sRNA analyzed. Hybridization signals are visualized by autoradiography on an X-ray film and can be quantified using a Molecular Imager FX system (Bio-Rad).

So far, several sRNA-specific qRT-PCR methods have been developed. It includes the stem-loop RT-PCR method (Chen *et al.*, 2005), the primer extension RT-PCR method (Raymond *et al.*, 2005), and the poly(T) adaptor RT-PCR method (Shi and Chiang, 2005). In the stem-loop RT-PCR method, a stem-loop RT primer is designed. It binds to the 3′ portion of sRNA molecules and is reverse transcribed with reverse transcriptase. The RT products are quantified using conventional TaqMan PCR that includes sRNA-specific forward primer, reverse primer, and a dye-labeled TaqMan probes (Chen *et al.*, 2005). The primer extension RT-PCR method relies on primer extension conversion of RNA to cDNA by reverse transcription followed by quantitative real-time PCR (Raymond *et al.*, 2005). The poly(T) adaptor RT-PCR method extends sRNA sequences by adding a poly(A) tail to the 3′ terminus, and then converts the poly(A) tailed sRNAs into cDNAs in a reverse transcription reaction primed by a poly(T) adaptor. PCR amplification is carried out using a sRNA-specific forward primer and a universal reverse primer complementary to the poly(T) adaptor. PCR products are monitored by a real-time signal detection system (Shi and Chiang, 2005). Compared with sRNA Northern hybridization, the sRNA-specific qRT-PCR methods are more sensitive and specific and less laborious. However, these methods require the expensive SYBR green mix and a real-time signal detection system.

sRNA microarray is the other effective method for analyzing the level of sRNAs (Axtell and Bartel,

2005; Barad *et al.*, 2004; Baskerville and Bartel, 2005; Lim *et al.*, 2005; Lu *et al.*, 2008; Nelson *et al.*, 2004, 2006; Thomson *et al.*, 2004). Like conventional microarrays, sRNA microarray can be used for genome-wide profiling of sRNAs. Additionally, the expression level of sRNAs can be estimated by high-throughput sequencing-based technology (Hoen *et al.*, 2008).

7   A PROTOCOL FOR sRNA CLONING

This protocol is modified from the methods described by Elbashir, Lendeckel, and Tuschl (2001), Lau *et al.* (2001), and Lagos-Quintana *et al.* (2002) and has been applied to the isolation of sRNAs from *Populus trichocarpa* and *Pinus taeda* (Lu *et al.*, 2005, 2007, 2008). In this protocol, total RNA will be separated on a denaturing 12% polyacrylamide gel. The 15–30-mer sRNAs will be recovered from gel and then dephosphorylated by alkaline phophatase. A 5′-phosphorylated 3′ adaptor will be ligated to the dephosphorylated RNA. The ligation products are separated from the nonligated sRNAs on a denaturing 10% polyacrylamide gel. The ligated sRNAs are then recovered from the gel. After 5′-phosphorylation, the 5′ adaptors containing hydroxyl groups at both 5′ and 3′ ends will be ligated to the 5′ ends of sRNAs. After gel purification, sRNAs will be converted to small DNA fragments by reverse transcription. Then, PCR amplification will be performed. The PCR products are digested with *Ban* I and concatamerized using T4 DNA ligase. Concatamers with sizes 400–700-bp are collected. After the unpaired ends are filled by incubation with Taq polymerase, DNA fragments will be cloned into a TA vector and then sequenced.

7.1   Primers

1. 3-adaptor: 5′-pCTGTAGGCACCATTCATCACx-3′ (p, phosphate; x, 3′-amino-modifier C-7).
2. 5-adaptor: 5′-ATGTCGTGaggcacctgaaa-3′ (RNA/DNA version, lowercase RNA).
3. 3-RT: 5′-GATGAATGGTGCCTAC-3′.
4. 5-PCR: 5′-GTCGTGAGGCACCTGAAA-3′.
5. 3-PCR: 5′-GATGAATGGTGCCTACAG-3′.

7.2   Preparation of A 12% PAGE Gel

1. Wash the tank of Protean II apparatus (Bio-Rad, Hercules, CA) with ddH$_2$O and clean glass plates with ddH$_2$O and ethanol.
2. Assemble glass plates for pouring gel.
3. In a bottle for gel pouring, add 14.33-mL sterile ddH$_2$O, 21-g urea, 15.6-mL acrylamide 40% stock (acrylamide : bis-acrylamide = 19 : 1), 5-mL 10× TBE buffer.
4. Dissolve urea in a water bath at 37°C with shaking.
5. Vacuum for 5–10 min to remove bubbles.
6. Add 300 μL of Ap and 21 μL of TEMED. Mix well.
7. Pour gel. Be aware not to introduce bubbles.
8. Apply a 1.5-mm comb with 20 wells.
9. Stand at room temperature for 1 h.

7.3   Pre-Running of PAGE Gel

1. Take out the comb.
2. Assemble the gel running cassette.
3. Add small amount of 1× TBE buffer to the upper tank to make sure that the buffer is not leaking and then add total of about 1.1 L buffer to the upper and lower tanks.
4. Use a syringe to rinse gel wells with running buffer.
5. Connect the Protean II apparatus with a pre-warmed water bath (60°C).
6. Turn on power supply.
7. Pre-run at 250–300 V for about 30 min till gel temperature reaches 60°C.
8. Use a syringe to rinse gel well with running buffer.

7.4   Preparation of RNA Samples and Gel Running

1. Isolate total RNA and resuspend in 50% formamide.
2. Add 1/10 vol. of 10× loading dye to the RNA samples, and then heat for 15 min at 70°C.
3. Load denatured RNA samples to gel wells.
4. Run RNA samples at 250–300 V for about 1 h.
5. Use a syringe to rinse gel wells with running buffer to remove large RNA molecules remained in the wells.

6. Stain the gel in a buffer containing 0.5× TBE and ethidium bromide (EtBr).

7.5 Purification of 15–30 Mer RNAs

1. Cut out the gel slice containing 15–30 mer sRNAs using a clean knife.
2. Put the gel slice into a 2-mL tube.
3. Add 2× gel vol (about 0.4 mL) of 0.3 M NaCl.
4. Elute sRNAs by incubating at 4°C for overnight with gentle agitation.
5. Spin briefly to pellet the polyacrylamide.
6. Transfer the supernatant to a 1.5-mL tube.
7. Rinse the pellet with 0.5× gel vol. (about 0.1 mL) of 0.3 M NaCl. Spin briefly, and then combine the supernatant.
8. Precipitate sRNAs with 2× vol. of cold ethanol at −80°C for about 2 h under the help of DNA/RNA carrier (glycogen) at a final concentration of $1 \, \mu g \, mL^{-1}$.
9. Recover sRNAs by centrifuging at full speed (15,000 rpm) for 30 min at 4°C.
10. Rinse with 70% cold ethanol and air-dry briefly.
11. Resuspend in 35-μL DEPC-treated H_2O.

7.6 Dephosphorylation of Purified RNAs

1. Add 5-μL 10× buffer and 10-μL (10 U) alkaline phosphatase (Fisher) to the 35-μL purified RNAs.
2. Incubate at 50°C for at least 30 min.
3. Add 150-μL DEPC-treated H_2O and 200-μL phenol/chloroform (1 : 1). Extract once.
4. Precipitate sRNAs with 400-μL cold ethanol, 12 μL of 5 M NaCl and carrier (glycogen) at a final concentration of $1 \, \mu g \, mL^{-1}$ at −80°C for about 2 h.
5. Recover RNA samples by centrifuging at full speed for 30 min at 4°C.
6. Rinse RNAs with 70% ethanol and air-dry briefly.
7. Resuspend in 12-μL DEPC-treated H_2O.

7.7 Ligation of 3′-Adaptor and Purification

1. Prepare the ligation mixture containing 12-μL dephosphorylated sRNAs, 4 μL of 100-μM 3′-adaptor (20 μM final), 2 μL of 10× buffer, and 2-μL (40 U) T4 RNA ligase.

2. Incubate at 37°C for at least 30 min.
3. Pour a 10%, 1.5-mm thick denaturing polyacrylamide gel.
4. Pre-run at 60°C for 30 min to warm up gel.
5. Add 25-μL formamide and 5 μL of 10× loading dye to the ligation mixture, and then heat for 10 min at 70°C.
6. Load RNA samples to gel wells.
7. Run RNA samples at 250–300 V for about 1 h.
8. Use a syringe to rinse gel wells with running buffer to remove large RNA molecules remained in the wells.
9. Stain the gel in a buffer containing 0.5× TBE and ethidium bromide (EtBr).
10. Cut out the gel slice containing 35–50 mer RNA.
11. Put the gel slice into a 2-mL tube.
12. Add 2× gel vol (about 0.4 mL) of 0.3 M NaCl.
13. Elute sRNAs by incubating at 4°C for overnight with gentle agitation.
14. Spin briefly to pellet the polyacrylamide.
15. Transfer the supernatant to a clean tube.
16. Rinse the pellet with 0.5× gel vol. (about 0.1 mL) of 0.3 M NaCl. Spin briefly, and then combine the supernatant.
17. Precipitate sRNAs with 2× vol. of cold ethanol at −80°C for about 2 h under the help of DNA/RNA carrier (glycogen) at a final concentration of $1 \, \mu g \, mL^{-1}$.
18. Recover sRNAs by centrifuging at full speed (15,000 rpm) for 30 min at 4°C.
19. Rinse with 70% cold ethanol and dry briefly.
20. Resuspend in 43-μL DEPC-treated H_2O.

7.8 Phosphorylation of Purified RNA

1. Add 5 μL of 10× T4 RNA ligase buffer or T4 DNA ligase buffer (Fisher) that contains ATP and 2-μL (20 U) T4 polynucleotide kinase to the 43-μL sRNAs.
2. Incubate at 37°C for at least 30 min.
3. Add 150-μL DEPC-treated H_2O and 200-μL phenol/chloroform (1 : 1). Extract once.
4. Precipitate sRNAs with 0.4-mL cold ethanol, 12 μL of 5 M NaCl and carrier (glycogen) at a final concentration of $1 \, \mu g \, mL^{-1}$ (1 μL of 1 mg mL^{-1} glycogen) at −80°C for at least 2 h.
5. Recover sRNAs by centrifuging at full speed for 30 min at 4°C.
6. Rinse the pellet with 70% ethanol and dry briefly.
7. Resuspend in 12-μL DEPC-treated H_2O.

7.9 Ligation of 5′-Adaptor and Purification

1. Prepare the ligation mixture containing 12-μL sRNAs ligated with the 3′-adaptors, 4 μL of 100-μM 5′-adaptor (20 μM final), 2 μL of 10× buffer, and 2-μL (40 U) T4 RNA ligase.
2. Incubate at 37°C for at least 30 min.
3. Pour a 10%, 1.5-mm thick denaturing polyacrylamide gel.
4. Pre-run at 60°C for 30 min to warm up gel.
5. Add 25-μL formamide and 5 μL of 10× loading dye to the ligation mixture, and then heat for 10 min at 70°C.
6. Load RNA samples to gel wells.
7. Run RNA samples at 250–300 V for about 1 h.
8. Use a syringe to rinse gel wells with running buffer to remove large RNA molecules remained in the wells.
9. Stain the gel in a buffer containing 0.5× TBE and ethidium bromide (EtBr).
10. Cut out the gel slice containing 55–70 mer RNA.
11. Put the gel slice into a 2-mL tube.
12. Add 2× gel vol (about 0.4 mL) of 0.3 M NaCl.
13. Elute sRNAs by incubating at 4°C for overnight with gentle agitation.
14. Spin briefly to pellet the polyacrylamide.
15. Transfer the supernatant to a clean tube.
16. Rinse the pellet with 0.5× gel vol. (about 0.1 mL) of 0.3 M NaCl. Spin briefly, and then combine the supernatant.
17. Precipitate sRNAs with 2× vol. of cold ethanol at −80°C for about 2 h under the help of DNA/RNA carrier (glycogen) at a final concentration of 1 μg mL^{-1}.
18. Recover sRNAs by centrifuging at full speed (15,000 rpm) for 30 min at 4°C.
19. Rinse with 70% cold ethanol and dry briefly.
20. Resuspend in 18.5-μL DEPC-treated H$_2$O.

7.10 RT-PCR Amplification of sRNAs Ligated with Adaptors

1. Add 3 μL of 100-μM RT primer to the 18.5 μL of sRNAs ligated with adaptors.
2. Heat the mixture at 80°C for 2 min.
3. Cool down on ice and then briefly spin down.
4. Add 3.0 μL of 10× M-Mulv buffer, 1.5 μL of 10-mM dNTP, 1.0 μL of RNase inhibitor (40 U), and 3.0 μL of M-Mulv (200 U μL^{-1}).

5. Incubate at 42°C for 60 min to reverse-transcribe the sRNAs.
6. Kill the enzyme by heating at 70°C for 15 min.
7. Add 1-μL RNase H, and then incubate at 37°C for 30 min.
8. Set up 12 PCR reaction mixtures (50 μL/reaction). Per 100 μL contains 5-μL reverse-transcribed sRNAs, 10 μL of 10× pfx50 DNA polymerase buffer, 10 μL of 1.0-mM dNTP, 1 μL of 100-μM 3-PCR primer, 1 μL of 100-μM 5-PCR primer, 1-μL pfx50 DNA polymerase, and 72-μL sterile water.
9. Carry out PCRs under the following conditions: 94°C for 2 min; 94°C for 15 s, 50°C for 30 s, 68°C for 30 s, 30 cycles; 68°C for 5 min; 4°C, hold.
10. Check RT-PCR products by running 20-μL products on a 12% denaturing polyacrylamide gel.
11. Extract the RT-PCR products by phenol (pH 8.0, twice) and chloroform (twice).
12. Add 1-mL cold ethanol, 30 μL of 5 M NaCl, and 2 μL of 1 mg mL^{-1} glycogen.
13. Precipitate at −80°C for about 2 h.
14. Spin at full speed for 30 min at 4°C to pellet PCR products.
15. Rinse pellet with 70% of ethanol and air-dry briefly.
16. Resuspend in a total of 300-μL sterile water.

7.11 Digestion and Concatermerization

1. Set up a *Ban* I digestion of PCR products: 40-μL PCR products, 30-μL NEBuffer 4, 5-μL *Ban* I (20 U μL^{-1}), and 225-μL sterile water.
2. Incubate overnight at 37°C.
3. Extract the digestion mixture by phenol (pH 8.0, once) and chloroform (once).
4. Add 600-mL cold ethanol, 18 μL of 5 M NaCl, and 1 μL of 1 mg mL^{-1} glycogen.
5. Precipitate DNA at −80°C for about 2 h.
6. Spin at full speed at 4°C for 30 min.
7. Rinse pellet with 70% ethanol and then spin at full speed at 4°C for 2 min.
8. Remove ethanol by a pipette and air-dry the pellet briefly.
9. Resuspend DNA in 43-μL sterile water.
10. Remove small DNA fragments and purify digested DNA by centri.spin-20 (Princeton Separations).

11. Add 5.0 μL of 10× T4 DNA ligase buffer and 2.0 μL of T4 DNA ligase to the purified DNA solution.

12. Incubate at room temperature and check the size of ligated DNA fragments frequently by running 1–2-μL ligation mixture on a 2.0% agarose gel.

13. When the majority of DNA fragments have sizes between 400 and 700 bp, run all of the remaining ligation mixture on a 2.0% agarose gel.

14. Excise the gel slice containing 400–700-bp DNA fragments

15. Recover DNA fragments from gel using a gel purification kit, such as the QIAquick Gel Extraction Kit (Qiagen).

16. Elute DNA fragments in 48-μL sterile water.

7.12 Cloning and Sequencing

1. Set up a Taq fill-in reaction mixture: 48-μL DNA fragments, 6 μL of 10× Taq PCR buffer, 4 μL of 1.0-mM dNTP, and 1-μL Taq DNA polymerase.

2. Take 4-μL ligation mixtures and ligate the DNA fragment to a TA vector, such as pCR2.1-TOPO vector (Invitrogen), as per kit manual.

3. Transfer plasmids into *Escherichia coli* competent cells.

4. Pick up colonies and sequence by a regular sequencing method.

8   A PROTOCOL FOR sRNA NORTHERN HYBRIDIZATION

This protocol is modified from the method described by Hutvágner *et al*. (2001) and has been used for the analysis of sRNAs from *P. trichocarpa* and *P. taeda* (Lu *et al*., 2005, 2007, 2008). In this protocol, total RNA (30 μg) will be denatured and then separated on a denaturing 17% polyacrylamide gel using the Protean II apparatus (Bio-Rad). sRNAs will be electroblotted onto a nylon membrane by the Trans-Blot SD Semi-Dry Electrophoretic Transfer Cell (Bio-Rad). The membrane is UV cross-linked, air-dried, and then prehybridized. Hybridization is carried out in the ULTRAhyb-Oligo hybridization buffer (Ambion). [γ-^{32}P]ATP-labeled DNA oligonucleotides complementary to sRNA sequences are used as probes. Signals are visualized by autoradiography on an X-ray film at −80°C.

8.1   Preparation of RNA Samples

1. Isolate total RNA and resuspend in 50% formamide.

2. Add 1/10 vol. of 10× loading dye to the RNA samples.

3. Prepare 10 μL (1 μg) of 10 bp ladder in 50% formamide and 1× loading dye.

4. Denature RNA samples and the 10-bp ladder at 70°C for 15 min.

5. Stain the gel for 5–10 min in 0.5× TBE with EtBr at room temperature in dark. And then rinse with sterile ddH$_2$O.

8.2   Preparation of A 17% PAGE Gel

1. Wash the tank of Protean II apparatus (Bio-Rad, Hercules, CA) with ddH$_2$O and clean glass plates with ddH$_2$O and ethanol.

2. Assemble glass plates for pouring gel.

3. In a bottle for gel pouring, add 8.68-mL sterile ddH$_2$O, 21-g urea, 21.25-mL acrylamide 40% stock (acrylamide : bis-acrylamide = 19 : 1), 5-mL 10× TBE buffer.

4. Dissolve urea in a water bath at 37°C with shaking.

5. Vacuum for 5–10 min to remove bubbles.

6. Add 300 μL of Ap and 21 μL of TEMED. Mix well.

7. Pour gel. Be aware not to introduce bubbles.

8. Apply a 1.5-mm comb with 20 wells.

9. Stand at room temperature for 1 h.

8.3   Pre-Running of PAGE Gel

1. Take out the comb.

2. Assemble the gel running cassette.

3. Add small amount of 1× TBE buffer to the upper tank to make sure that the buffer is not leaking and then add total of about 1.1 L buffer to the upper and lower tanks.

4. Use a syringe to rinse gel wells with running buffer.

5. Connect the Protean II apparatus with a pre-warmed water bath (60°C).

6. Turn on power supply.

7. Pre-run at 250–300 V for about 30 min till gel temperature reaches 60°C.

8. Use a syringe to rinse gel well with running buffer.

8.4   Running of PAGE Gel

1. Load denatured RNA samples and the 10-bp ladder to gel wells.
2. Run RNA samples at 250–300 V for about 1 h.
3. Use a syringe to rinse gel wells with running buffer to remove large RNA molecules remained in the wells.
4. Stain the gel in a buffer containing 0.5× TBE and ethidium bromide (EtBr).
5. Check RNA bands under a UV light.
6. Equilibrate the gel with prechill 0.5× TBE for 15 min.

8.5   Transfer RNA to Hybond-N$^+$ Nylon Membrane

1. Prepare a Trans-Blot SD semi-dry electrophoretic transfer cell (Bio-Rad, Cat# 170-3940).
2. Pre-wet extra thick filter paper (two papers per gel) and Hybond-N$^+$ nylon membrane in prechill 0.5× TBE buffer for about 15 min.
3. Put a sheet of pre-soaked extra thick filter paper onto the platinum anode. Roll a test tube over the surface of the filter paper to exclude air bubbles.
4. Place the pre-wetted nylon membrane on the filter paper. Remove air bubbles between the nylon membrane and the filter paper.
5. Carefully place the equilibrated gel on the center of the nylon membrane. Remove air bubbles between the gel and the membrane.
6. Put the other sheet of pre-soaked filter paper on the gel. Remove air bubbles between the filter paper and the gel.
7. Carefully place the cathode onto the stack and the safety cover on the unit.
8. Turn power on. Set the voltage limit to be 25 V and the current limit to be 0.4 A. RNA samples will be transferred onto the nylon membrane in about 1 h.
9. Rinse the membrane briefly in 2× SSC buffer.
10. Briefly dry the membrane on a filter paper.
11. Cross-link the RNA samples onto the membrane by a UV cross-linker.
12. Completely dry the membrane on a filter paper at 65–80°C for 1 h.

8.6   Probe Labeling and Purification

DNA oligonucleotides complementary to sRNA sequences are end labeled with [γ-^{32}P]ATP using T4 polynucleotide kinase (New England Biolabs).

1. Prepare a 20-μL reaction mixture containing 2 μL of 2-μM DNA oligos, 2 μL of 10× T4 polynucleotide kinase buffer, 1 μL of 10 U μL^{-1} T4 polynucleotide kinase, and 2 μL of 2 pmol μL^{-1} [γ-^{32}P]ATP.
2. Incubate at 37°C for 0.5–1 h.

Unincorporated [γ-^{32}P] ATP is removed by Mini Quick Spin Oligo columns (Roche, Cat. No. 1814397).

1. Resuspend column matrix.
2. Remove the top cap, snap off the bottom tip, and then place the column into a sterile microcentrifuge tube.
3. Spin column at 1000×g for 1 min to pack the column and to remove residual buffer.
4. Carefully apply the labeled oligonucleotide solution to the center of the column bed.
5. Spin column at 1000×g for 4 min.
6. Recover eluate containing the end-labeled oligonucleotides (probes).

8.7   Hybridization

1. Preheat ULTRAhyb-Oligo (hybridization buffer for oligonucleotides, Ambion) to 68°C to dissolve any precipitated materials.
2. Prehybridize the membrane by 10–20-mL hybridization buffer at 37°C for 30 min. Make sure that the buffer is enough to keep the membrane uniformly wet.
3. Remove extra hybridization buffer. Leave about 5 mL in the tube.
4. Add 10^6 cpm mL^{-1} of the end-labeled oligonucleotides (20 μL).
5. Hybridize overnight at 37°C.
6. Discard hybridization buffer, and then immediately pour 25-mL washing buffer consisting of 2× SSC and 0.5% SDS onto the blot and incubate at 37°C for 30 min with gentle agitation.
7. Replace with fresh washing buffer and repeat the washing of blot at 37°C for 20 min.

8. Expose the blot to an X-ray film at −80°C with one or two intensifying screens for 1–7 days. The actual length of exposure time depends on the radiation intensity.

8.8 Stripping

The membrane can be used to hybridize with the other probes after stripping. The stripping procedures are showed below.

1. Boil some 0.1% SDS solution.
2. Place the membrane in a hybridization tube and pour the SDS solution into the tube.
3. Incubate the tube at 85°C for 0.5 h in a hybridization oven with rotation, and then cool down to room temperature.
4. Repeat once.
5. Rinse the membrane with 2× SSC buffer. Now the membrane is ready for prehybridization.

9   CONCLUSION

Recently, studies on sRNAs have been very hot in biology since the first discovery of miRNAs from *Caenorhabditis elegans* in 1993 (Lee, Feinbaum, and Ambros, 1993; Wightman, Ha, and Ruvkun, 1993) and the first discovery of plant miRNAs from *Arabidopsis* in 2002 (Llave *et al*., 2002; Park *et al*., 2002; Reinhart *et al*., 2002; Rhoades *et al*., 2002). Many sRNAs have been experimentally cloned or computationally predicted from various plant species, particularly those model plants, such as *Arabidopsis*, rice, and *P. trichocarpa* (Llave *et al*., 2002; Reinhart *et al*., 2002; Wang *et al*., 2004; Lu *et al*., 2005). These sRNAs can be classified into miRNAs and siRNAs. The class of plant endogenous siRNAs includes several subclasses, such as ta-siRNAs, NAT-siRNAs, and hc-siRNAs (Allen *et al*., 2005; Borsani *et al*., 2005; Hamilton *et al*., 2002; Rajagopalan *et al*., 2006; Vazquez *et al*., 2004b, 2010). The functions of some sRNAs have been characterized using various sRNA-specific techniques, such as Northern hybridization, computational target prediction, and experimental validation. The results showed that sRNAs regulated plant development and defense response through direct cleavage of transcripts, translational inhibition, or epigenetic modification. However, the regulatory roles of many other sRNAs are still unknown, and novel sRNAs need to be identified from many non-model plants. Application of the techniques currently used will be certainly helpful in sRNA identification and characterization, whereas novel high-throughput methods are particularly expected to speed up the studies on sRNAs.

REFERENCES

Addo-Quaye, C., Snyder, J. A., Park, Y. B., *et al*. (2009) *RNA*, **15**, 2112–2121.

Allen, E., Xie, Z., Gustafson, A. M., *et al*. (2005) *Cell*, **121**, 207–221.

Aukerman, M. J. and Sakai, H. (2003) *Plant Cell*, **15**, 2730–2741.

Axtell, M. J. and Bartel, D. P. (2005) *Plant Cell*, **17**, 1658–1673.

Barad, O., Meiri, E., Avniel, A., *et al*. (2004) *Genome Res.*, **14**, 2486–2494.

Baskerville, S. and Bartel, D. P. (2005) *RNA*, **11**, 241–247.

Borsani, O., Zhu, J., Verslues, P. E., *et al*. (2005) *Cell*, **123**, 1279–1291.

Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., *et al*. (2008) *Science*, **320**, 1185–1190.

Carrington, J. C. and Ambros, V. (2003) *Science*, **301**, 336–338.

Chen, C., Ridzon, D. A., Broomer, A. J., *et al*. (2005) *Nucleic Acids Res.*, **33**, e179.

Chen, X. (2004) *Science*, **303**, 2022–2025.

Dai, X. and Zhao, P. X. (2011) *Nucleic Acids Res.*, **39**, W155–W159.

Dsouza, M., Larsen, N. and Overbeek, R. (1997) *Trends Genet.*, **13**, 497–498.

Elbashir, S. M., Lendeckel, W. and Tuschl, T. (2001) *Genes Dev.*, **15**, 188–200.

Ewing, B. and Green, P. (1998) *Genome Res.*, **8**, 186–194.

Fang, Y. and Spector, D. L. (2007) *Curr. Biol.*, **17**, 818–823.

German, M. A., Luo, S., Schroth, G., *et al.* (2009) *Nat. Protoc.*, **4**, 356–362.

Hamilton, A., Voinnet, O., Chappell, L., *et al.* (2002) *EMBO J.*, **21**, 4671–4679.

Hoen, P. A., Ariyurek, Y., Thygesen, H. H., *et al.* (2008) *Nucleic Acids Res.*, **36**, e141.

Hutvágner, G., McLachlan, J., Pasquinelli, A. D., *et al.* (2001) *Science*, **293**, 834–838.

Jones-Rhoades, M. W. and Bartel, D. P. (2004) *Mol. Cell*, **14**, 787–799.

Jones-Rhoades, M. W., Bartel, D. P. and Bartel, B. (2006) *Annu. Rev. Plant Biol.*, **57**, 19–53.

Kasschau, K. D., Fahlgren, N., Chapman, E. J., *et al.* (2007) *PLoS Biol.*, **5**, e57.

Katiyar-Agarwal, S., Gao, S., Vivian-Smith, A., *et al.* (2007) *Genes Dev.*, **21**, 3123–3134.

Katiyar-Agarwal, S., Morgan, R., Dahlbeck, D., *et al.* (2006) *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 18002–18007.

Kozomara, A. and Griffiths-Jones, S. (2011) *Nucleic Acids Res.*, **39**, D152–D157.

Kurihara, Y., Takashi, Y. and Watanabe, Y. (2006) *RNA*, **12**, 206–212.

Lagos-Quintana, M., Rauhut, R., Yalcin, A., *et al.* (2002) *Curr. Biol.*, **12**, 735–739.

Lau, N. C., Lim, L. P., Weinstein, E. G., *et al.* (2001) *Science*, **294**, 858–862.

Lee, R. C., Feinbaum, R. L. and Ambros, V. (1993) *Cell*, **75**, 843–854.

Li, R., Yu, C., Li, Y., *et al.* (2009) *Bioinformatics*, **25**, 1966–1967.

Lim, L. P., Lau, N. C., Garrett-Engele, P., *et al.* (2005) *Nature*, **433**, 769–773.

Lippman, Z. and Martienssen, R. (2004) *Nature*, **431**, 364–370.

Llave, C., Xie, Z., Kasschau, K. D., *et al.* (2002) *Science*, **297**, 2053–2056.

Lu, C., Kulkarni, K., Souret, F. F., *et al.* (2006) *Genome Res.*, **16**, 1276–1288.

Lu, S., Sun, Y.-H. and Chiang, V. L. (2008) *Plant J.*, **55**, 131–151.

Lu, S., Sun, Y.-H. and Chiang, V. L. (2009) *Nucleic Acids Res.*, **37**, 1878–1885.

Lu, S., Sun, Y.-H., Amerson, H., *et al.* (2007) *Plant J.*, **51**, 1077–1098.

Lu, S., Sun, Y.-H., Shi, R., *et al.* (2005) *Plant Cell*, **17**, 2186–2203.

Ma, Z. R., Coruh, C. and Axtell, M. J. (2010) *Plant Cell*, **22**, 1090–1103.

Mallory, A. C. and Vaucheret, H. (2006) *Nat. Genet.*, **38**, S31–S36.

Meyers, B. C., Axtell, M. J., Bartel, B., *et al.* (2008) *Plant Cell*, **20**, 3186–3190.

Nelson, P. T., Baldwin, D. A., Kloosterman, W. P., *et al.* (2006) *RNA*, **12**, 187–191.

Nelson, P. T., Baldwin, D. A., Scearce, L. M., *et al.* (2004) *Nat. Meth.*, **1**, 155–161.

Norden-Krichmar, T. M., Allen, A. E., Gaasterland, T., *et al.* (2011) *PLoS One*, **6**, e22870.

Park, M. Y., Wu, G., Gonzalez-Sulser, A., *et al.* (2005) *Proc. Natl. Acad. Sci. U. S. A.*, **102**, 3691–3696.

Park, W., Li, J., Song, R., *et al.* (2002) *Curr. Biol.*, **12**, 1484–1495.

Rajagopalan, R., Vaucheret, H., Trejo, J., *et al.* (2006) *Genes Dev.*, **20**, 3407–3425.

Raymond, C. K., Roberts, B. S., Garrett-Engele, P., *et al.* (2005) *RNA*, **11**, 1737–1744.

Reinhart, B. J., Weinstein, E. G., Rhoades, M. W., *et al.* (2002) *Genes Dev.*, **16**, 1616–1626.

Rhoades, M. W., Reinhart, B. J., Lim, L. P., *et al.* (2002) *Cell*, **110**, 513–520.

Shi, R. and Chiang, V. L. (2005) *Biotechniques*, **39**, 519–525.

Sun Y-H, Lu S, Shi R and Chiang VL. 2011. Computational prediction of plant miRNA target, in Methods in Molecular Biology Series, vol 744: RNAi and Plant Gene Function Analysis. Kodama H and Komamine A (Eds.). Humana Press: New York; 175–186.

Szittya, G., Moxon, S., Santos, D. M., *et al.* (2008) *BMC Genomics*, **9**, 593.

Tang, G., Reinhart, B. J., Bartel, D. P., *et al.* (2003) *Genes Dev.*, **17**, 49–63.

Thomson, J. M., Parker, J., Perou, C. M., *et al.* (2004) *Nat. Meth.*, **1**, 47–53.

Vazquez, F., Gasciolli, V., Crété, P., *et al.* (2004a) *Curr. Biol.*, **14**, 346–351.

Vazquez, F., Legrand, S. and Windels, D. (2010) *Trends Plant Biol.*, **15**, 337–345.

Vazquez, F., Vaucheret, H., Rajagopalan, R., *et al.* (2004b) *Mol. Cell*, **16**, 69–79.

Voinnet, O. (2009) *Cell*, **136**, 669–687.

Wang, H., Chua, N. H. and Wang, X. J. (2006) *Genome Biol.*, **7**, R92.

Wang, J.-F., Zhou, H., Chen, Y.-Q., *et al.* (2004) *Nucleic Acids Res.*, **32**, 1688–1695.

Wightman, B., Ha, I. and Ruvkun, G. (1993) *Cell*, **75**, 855–862.

Wu, B., Wang, M., Ma, Y., *et al.* (2012) *PLoS ONE*, **7**, e44385.

Wu, L., Zhou, H., Zhang, Q., *et al.* (2010) *Mol. Cell*, **38**, 465–475.

Zhang, B. H., Pan, X. P., Cox, S. B., *et al.* (2006) *Cell. Mol. Life Sci.*, **63**, 246–254.

Zhang, Y. (2005) *Nucleic Acids Res.*, **33**, W701–W704.

Zhou, X., Sunkar, R., Jin, H., *et al.* (2009) *Genome Res.*, **19**, 70–78.

Zuker, M. (2003) *Nucleic Acids Res.*, **31**, 3406–3415.

39

Metabolomics

Jan Schripsema and Denise Dagnino

Grupo Metabolômica, Universidade Estadual do Norte Fluminense, Campos dos Goytacazes, Brazil

1 INTRODUCTION

Among genomics, proteomics, and metabolomics, metabolomics is the most recent evolved subject and still remains behind the former two subjects in the number of publications (Schripsema, 2010). However, the range of applications is enormous, and the trend indicates that in about 10 years, metabolomics might pass the other two subjects in importance.

The focus of this chapter is on the practical and efficient application of metabolomics. As a handbook, it provides orientations on the following different aspects:

- how to design the experiment;
- what is the analytical technique of choice;
- how to extract and process the material;
- how to obtain the maximum of information from the experiments; and
- how to interpret and verify the data.

The techniques used in metabolomics are still in rapid development, and new techniques are introduced or methods are modified. Metabolomics involves a series of different aspects, involving experimental planning, sample treatment and extraction, analysis with different techniques, data processing, and interpretation. The multidisciplinary nature and the lack of routine techniques and procedures turn it into a rather complicated field of research.

For outsiders, it might seem that results are mainly based on software-based interpretation of the spectra or chromatograms obtained from the measurement of large series of samples, leading to some black-box idea. Maybe as a result, the number of reviews is high in relation to the number of original papers, but most of those reviews focused on specific applications or parts of the metabolomic analysis. In this chapter, a basic introduction is presented and guidelines are provided for the efficient use through citing most recent key publications on specific aspects.

2 METABOLOMICS, ITS OBJECTIVES, AND DEFINITION

In this chapter, a general overview of metabolomics and the state of the art is presented, focusing on promising developments. Some years ago, we already gave a review about the use of nuclear magnetic resonance (NMR) in metabolomics (Schripsema, 2010).

Many definitions have been launched for metabolomics, but the area of research is difficult to catch in a single definition. The name metabolomics has been introduced by Oliver *et al.* in 1998, based on the analogy with the other omics, especially genomics, proteomics, and transcriptomics.

The metabolome can be understood as the complete set of metabolites in an organism or system. Consequently, the term *metabolomics* can be understood as the identification and quantification of all metabolites (Dettmer, Aronov, and Hammock, 2007; Schripsema, 2010).

However, in most metabolomic research, only a limited part of the metabolome is investigated and most of the metabolome remains unexplored, mainly because of the experimental procedures that do not cover the whole set of metabolites or because of some factors such as concentration (below detection limit) or stability (some metabolites being also volatile). Now, the methods used for the analysis of metabolome are especially based on mass spectrometry (MS) and NMR. In any plant species, the presence of thousands of compounds is expected. In 2004, it was roughly estimated that in *Arabidopsis*, about 5000 different compounds, including primary and secondary metabolites, are present (Bino *et al.*, 2004). Recently, it was reported that a little more than 1500 metabolites have been routinely detected in *Arabidopsis*, but only 730 of those were chemically defined (Quanbeck *et al.*, 2012).

Geier *et al.* (2011) described metabolomics as the untargeted profiling of metabolites. However, in metabolomics, there is normally a clear objective to find differences in metabolic fingerprints; therefore, a good general definition would be the following:

Metabolomics is the area of research that strives to obtain complete metabolic fingerprints, to detect differences between them and to provide hypothesis to explain those differences.

In Figure 1, this view of metabolomics is depicted.

In recent years, metabolomic research continued increasing strongly, in contrast to proteomics and genomics, which passed already through their major growth phase (Schripsema, 2010). Together with the other omics, metabolomics also contributes to systems biology, the field of research, which uses a holistic view of the biological system, looking for the interactions operating within it. Furthermore,
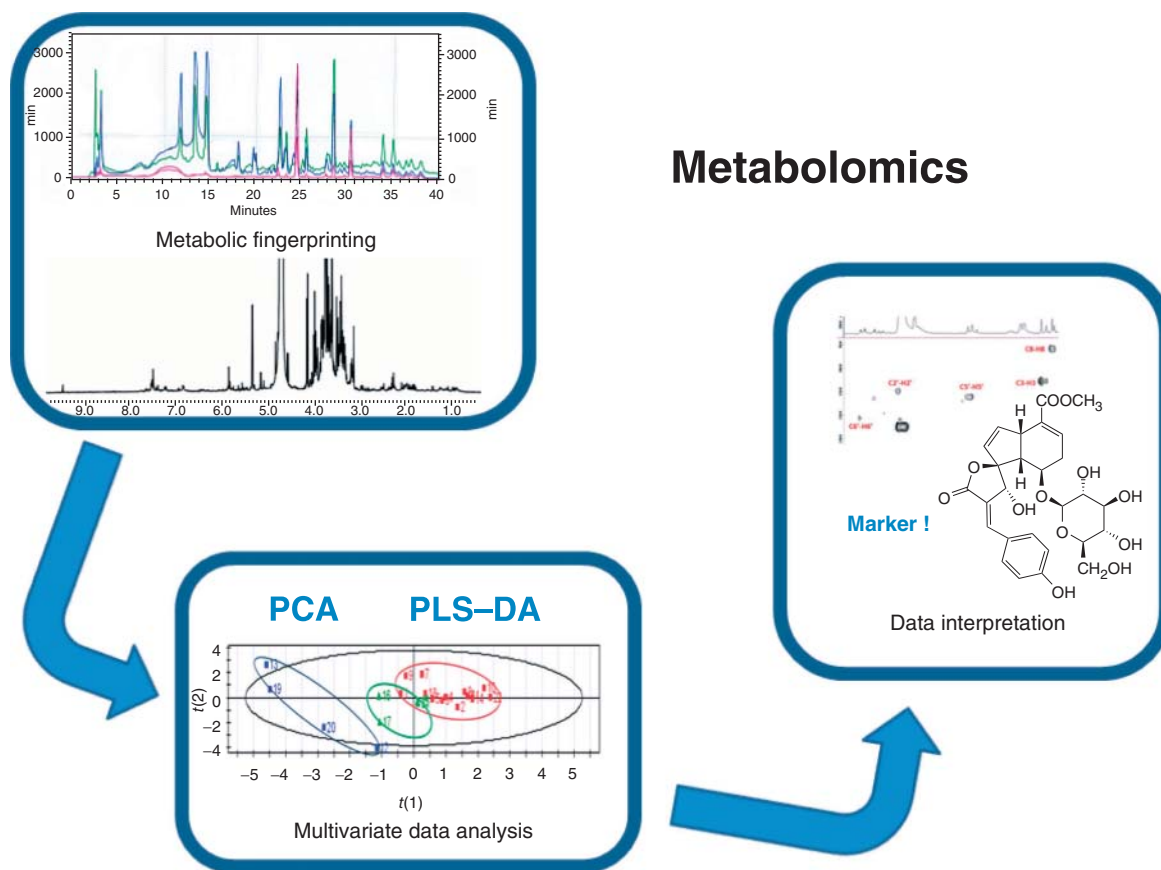
Figure 1 Metabolomics is the area of research that strives to obtain complete metabolic fingerprints, to detect differences between them and to provide hypothesis to explain those differences.

metabolomics can be considered an important tool for functional genomics (Bino *et al*., 2004).

3 DIFFERENT STEPS IN A METABOLOMIC STUDY

In a metabolomic study, different stages can be discerned as follows.

1. Proposing the questions to be answered and the objectives of the study.
2. Experimental design.
3. Sampling of the material to be analyzed.
4. Measurement of spectra and/or chromatograms.
5. Data pretreatment.
6. Multivariate analysis.
7. Interpretation of factors.

The procedures of a metabolomic experiment are depicted in Figure 2. In the following sections, these different stages of the metabolomic experiment are discussed.

3.1 Proposing the Questions to Be Answered and the Objectives of the Study

The most important and often sub-estimated part of metabolomic experiments is the initial planning and experimental design. Errors in this phase generally lead to deficient experiments and lack of data. Moreover, these errors are generally impossible to be repaired in the next phases of the experiments. Imprecisions also turn the data difficult to interpret, independent from and despite the extensive use of multivariate tools.

The objective common to metabolomic studies is to identify compounds characteristic of specific metabolic states of organisms, tissues, or cells in an attempt to find molecular markers. This type of study is of relevance in a variety of fields, with applications in, for instance, medicine and agriculture. Thus, advancement in metabolomics should be monitored in a wide literature range.

Some of the specific aspects investigated by plant metabolomics are the metabolic differences between genotypes (DiLeo *et al*., 2011; Houshyani *et al*., 2011), phenotypes (Sardans, Penuelas, and Rivas-Ubach, 2011) or diseased or stressed (Obata
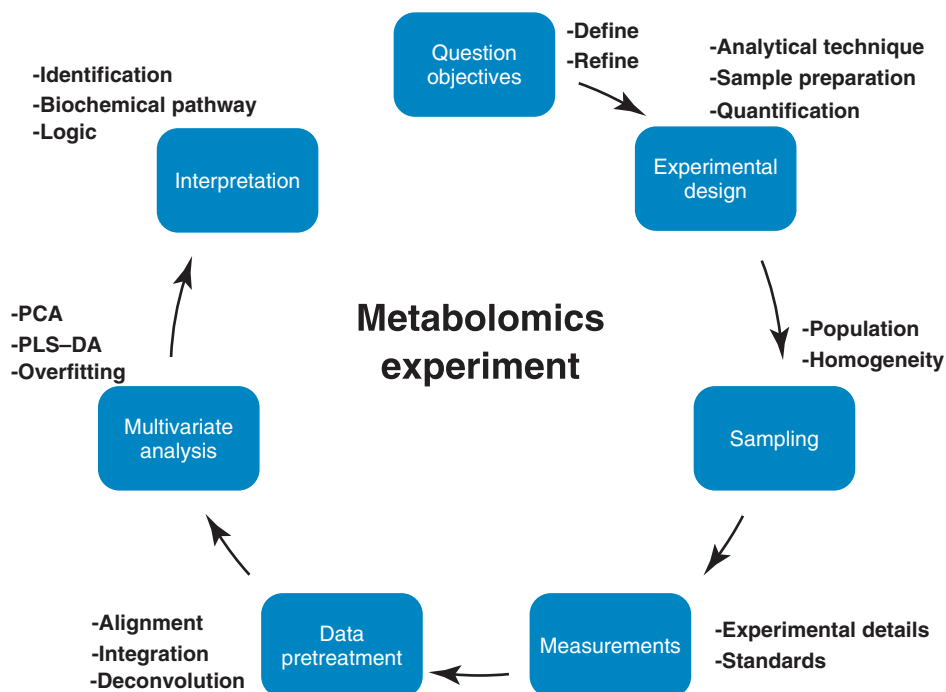
Figure 2 Schematic representation of a typical metabolomics experiment. Details of the individual steps can be found in the text.

and Fernie, 2012) plants. Besides the pure scientific relevance of these questions, these types of studies have implications regarding health and economical issues such as the evaluation of the suitability of genetically modified plants for consumption (Davies, 2010), to assess nutraceuticals, different cultivars, and varieties in agriculture and to find clues to combat metabolic disorders (Schripsema *et al*., 2010) or diseases.

The aim of metabolomics is very ambitious, and is still hampered by technical issues. Thus, in practice, experiments monitor a subset of the metabolites present. Proposing the questions to be answered will guide the "where to measure, what and how?" and the introduction of all the necessary precautions to generate conclusive results. Metabolomics has been presented as a hypothesis-generating functional genomics tool (Quanbeck *et al*., 2012). However, despite generating hypothesis, metabolomic experiments need an initial question or hypothesis to be able to come to an effective experimental setup. A question might be to find the differences between phenotypes of a certain species. It seems simple; just take samples from the two for comparing. However, what if one phenotype has a retarded growth, should they be compared in the same developmental stage, or should the samples be taken on the same day? Maybe it is necessary to make a time course for each. This example shows that the initial question should be well thought over and eventually refined.

3.2 Experimental Design

3.2.1 *Analytical Techniques*

The analytical technique to be used for analyzing the samples is the next point to be thought over. A number of very powerful techniques that can combine spectrometric and chromatographic methods are now available but it is necessary to keep always in mind that none of them is able to analyze the whole range of compounds present in the metabolome.

In principle, any technique can be used, which provides information about the composition of metabolites in a sample. Nevertheless, factors such as the amount of compound available, the preparation of the sample, the variety of compounds investigated, and the possibility of identifying these compounds should all be considered when choosing the technique.

Table 1 The strong and weak points of the major analytical platforms, which are used in metabolomics.

	GC–MS	LC–MS	NMR
Absolute quantitation	++	+	+++
Precision	++	++	+++
Reproducibility	++	+	+++
Quantity of compounds	+++	+++	+
Sensitivity	+++	+++	+
Limit of detection	+++	+++	+
Sample preparation	+	++	+++
Duration of analysis	+	++	+++

In metabolomics, gas chromatography–mass spectrometry (GC–MS), liquid chromatography–mass spectrometry (LC–MS), and ^{1}H NMR are the most commonly used analytical techniques. Each of these techniques has its specific advantages and disadvantages, which are given in Table 1.

NMR is an outstanding technique for the identification and quantification of compounds. In metabolomics, ^{1}H NMR is used owing to the abundance of the isotope and because essentially all of the metabolites contain it and thus can be detected. It has been used for decades to aid the identification of natural products and thus a huge library of spectra is available to aid the identification of compounds in samples. Quantification by NMR is also straight forward as the integral of each signal corresponds to the absolute quantity of hydrogen atoms responsible for the signal. Therefore, relative concentrations of compounds in mixtures can be determined straight away.

The ^{1}H NMR spectra of samples in metabolomic experiments are usually complex because they contain the signals of all compounds of the mixture, and the total spectrum shows a superposition of the spectra of the individual compounds. Furthermore, the sensitivity of ^{1}H NMR is relatively low (at least when compared to MS). Therefore, with ^{1}H NMR, only dozens of compounds can be detected in a single analysis.

However, NMR has some great advantages in identification and quantitation of the individual compounds, as discussed later in this chapter.

MS is the other spectrometric method commonly used in metabolomics. The sensitivity of MS is much higher than that of ^{1}H NMR, at least a factor of 1000 higher, and consequently, many more compounds can be detected. Direct introduction mass spectrometry (DI-MS), generally combined

with high resolution mass spectrometers, such as Fourier transfer ion cyclotron resonance (FT-ICR) or linear trap quadrupole (LTQ)-Orbitrap instruments (Junot *et al.*, 2010), has been successfully used, but disadvantages include matrix effects and the formation of adduct and product ions. MS is thus usually coupled to a chromatographic system. In metabolomics, the combinations GC–MS and LC–MS are most frequently used. Both platforms provide the possibility for the detection of hundreds of compounds in a single analysis.

Capillary GC–MS instruments are used routinely for decades for relatively volatile or derivatized compounds. Nowadays, a further increase of chromatographic separation can be achieved with two-dimensional gas chromatography (GC) (Marriott and Shellie, 2002; Adahchour *et al.*, 2006). The MS spectra generated by electron impact have a good reproducibility and extensive libraries exist to aid the identification of compounds.

LC–MS instruments generally work with electrospray ionization (ESI). Little fragmentation occurs in this type of ionization, but with tandem MS, fragmentation can be induced. In general, there is a certain difficulty in comparing this type of MS spectra because of their variability in relation to the experimental conditions.

A general disadvantage of MS in relation to NMR is that the resulting spectra contain much less information about the chemical structure. Nevertheless, in GC–MS or LC–MS, combining the information of the retention in the chromatographic system with spectral information is generally sufficient to confirm the presence of known substances, but identification of new substances is nearly impossible.

A second general serious limitation of MS regards quantification. Owing to differences in ionization characteristics, each compound gives a specific response, with much variation between individual compounds. This means that for quantification, a calibration is needed for each component. Often isotope-labeled analogs are used for this purpose (Wu *et al.*, 2005).

A metabolomic experiment can be done with any of the mentioned techniques. In choosing the technique, limits are posed generally by the availability of the equipment and/or expertise. Buscher *et al.* (2009) made a systematic cross-platform comparison of different separation and detection methods for quantitative metabolomics. The techniques GC–MS,

LC–MS, and capillary electrophoresis–mass spectrometry (CE–MS) were compared using a defined mixture of 91 metabolites, mainly abundant primary metabolites. LC–MS provided the best combination of versatility and robustness. In this study, the influence of the matrix on the quantification was also investigated. Most compounds were found to be influenced by matrix effects and some compounds became immeasurable.

When available, multiplatform approaches are advisable in cases a broader spectrum of compounds is of interest. Beltran *et al.* (2012) emphasize the complementarity of NMR and MS in investigating the chemically diverse metabolites of the metabolome and show that it is possible to use both platforms using the same sample preparation procedures. Using multiple platforms allows the expansion of the coverage of the metabolites investigated, exploring the advantages of each technique.

With the technique to be used chosen, the setup of the experiment should be decided. In this setup, one needs to have some previous information about the quantities of the compounds that are expected to be present and are intended to be determined. A sample should be obtained with suitable concentrations to have an accurate detection. It also should be decided how the sample needs to be prepared. Generally, an extraction step is necessary. What solvent should be used in this extraction? Is it possible to use a single solvent?

3.2.2 Sample Preparation

A series of methods are available in the literature for the preparation of samples. Nevertheless, tests should be performed to guarantee the adequate implementation of a suitable routine. As a rule, the sample preparation routine should be as simple as possible. This will decrease both costs and errors. Any additional step in a procedure will cost time and introduces a possibility of error.

The usual procedure used in metabolomics is the extraction of the metabolites by solvents or mixtures of solvents from the matrix whose constituents vary according to the type of material. For example, the extracellular matrix in plant material consists mainly of cellulose and glycoproteins such as pectin and extensin, whereas in animals, two main classes of

macromolecules make up the matrix, glycosamino-glycans, usually covalently bound to proteins, and fibrous proteins, such as collagen and elastin.

As organisms share many metabolic pathways (primary metabolism), metabolites from different sources can be extracted under similar conditions.

To speed up the extraction of solid biological material, it might be necessary to perform a previous tissue disruption, if necessary under cooling, before the extraction. Tissue disruption can be achieved by several methods. Geier *et al.* (2011) did a comparison of *Caenorhabditis elegans* tissue extraction strategies to obtain a comprehensive metabolome coverage. In their study, beadbeating with 80% aqueous methanol provided the best overall results on different platforms (GC–MS, LC–MS, and NMR). For the extraction of plant material, drying and milling or freezing and lyophilization are generally sufficient to obtain complete extraction of the material. Anyway, for every specific study, it is recommended to do some tests to verify the extraction efficiency.

The solvent used has a great influence on the compounds extracted. Every solvent will provide a different qualitative profile and a quantitative profile. A series of papers have been published on this question of solvent selection (Beltran *et al.*, 2012; Kim and Verpoorte, 2010). Generally, it is verified what solvent or solvent mixture extracts most compounds or provides more signals in a chromatographic system. However, other factors also need consideration. The specific characteristics of the solvent should be considered, regarding its physical properties, the toxicity, and possibilities of artifact formation (Kim and Verpoorte, 2010). For instance, a low boiling point can be useful if samples need to be dried, but might be problematic regarding quantification because of evaporation. The toxicity of the solvent is a further issue, and the choice for less toxic ones should be preferred.

It should also be verified what is the influence of the solvent in the subsequent analysis. If derivatization is required in a next step, the solvent should be compatible. If NMR is used, it can be useful to do the extraction directly with deuterated solvents. If, in NMR experiments, the identification of the components is important, the solvent should be compatible with the spectral libraries that can be used. Traditionally, most NMR spectra have been recorded in single solvents, $CDCl_3$ for apolar compounds and D_2O for polar compounds. Therefore, we routinely use

in our laboratory an extraction method employing a two-phase system of D_2O and $CDCl_3$ (Schripsema, 2008). This yields two completely complementary fractions, and the NMR data of unknown components can be directly compared with previous published NMR data (Figure 3).

In many studies, mixtures of solvents are used with the intention to extract as much as possible compounds in a single extract. The advantage is that only one single extract needs to be measured, but the quantity of compounds that can be observed in a single extract is generally less than what can be observed in two complementary extracts. Furthermore, a specific database has to be made to be able to identify individual components in the mixture of solvents. Beltran *et al.* (2012) compared 12 extraction protocols for the liver tissue. The extraction efficiency of 30 primary metabolites was tested with four solvent mixtures and three temperatures. Most efficient extraction was found with the mixture of methanol–chloroform–water (7 : 2 : 1).

For metabolite fingerprinting of *Arabidopsis thaliana*, the extraction was performed with a mixture of D_2O and CD_3OD (80 : 20) (Ward *et al.*, 2003). In this study, different ecotypes could be distinguished by their metabolite profile using ^{1}H NMR. In an evaluation of extraction protocols for characterizing the metabolome of the hepatobiliary fluke *Fasciola hepatica*, using ultra performance liquid chromatography (UPLC) and CE, both coupled with mass spectroscopy, it was concluded that the most suitable extraction solvent was a mixture of methanol–water–chloroform in the ratio 59 : 26 : 15. The evaluated analytical platforms were UPLC–MS with two types of columns, C18 or hydrophilic interaction liquid chromatography (HILIC), and CE–MS. The performance of these systems was evaluated in both positive and negative ion modes in MS (Saric *et al.*, 2012). From the 142 identified metabolites, only nine were identified across all analytical platforms, and it was concluded that these platforms were largely complementary.

3.2.3　Analysis without Extraction

The extraction step is a source of variation and the compounds detected are dependent on the solvent used. If the extraction step could be omitted, it would
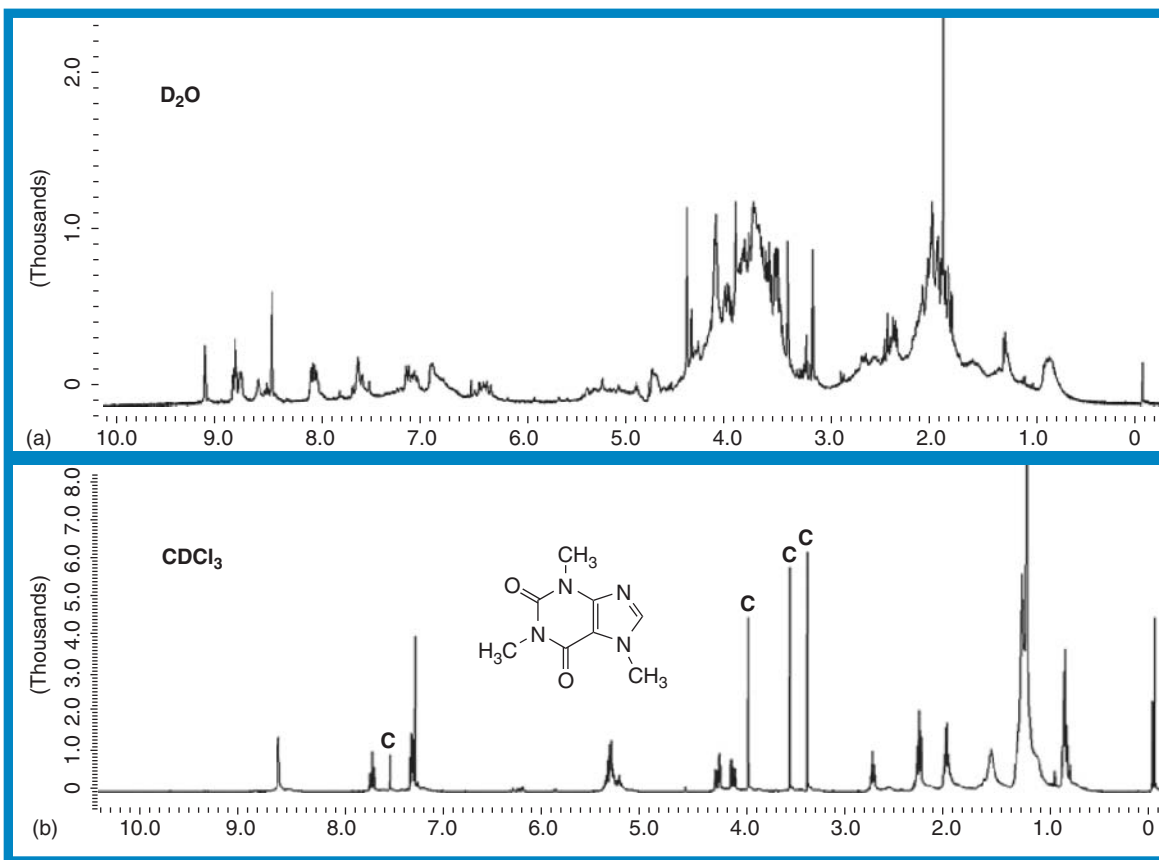
Figure 3 ^{1}H NMR spectra from the (a) D_2O and (b) $CDCl_3$ extracts of coffee powder. A two-phase extraction was performed, and after centrifugation, NMR spectra were obtained from both phases. In the $CDCl_3$ phase, signals from caffeine (marked with C) and triglycerides are visible, whereas the D_2O phase shows mainly caffeic acid esters and carbohydrates.

theoretically simplify the procedure and consequently there would be less potential sources of error in the experiment. High resolution magic angle spinning nuclear magnetic resonance (HR-MAS NMR), desorption electrospray ionization mass spectrometry (DESI–MS), and direct analysis in real time mass spectrometry (DART-MS) are techniques that are able to measure spectra from crude nonprocessed materials.

3.2.3.1 HR-MAS NMR (High Resolution Magic Angle Spinning NMR)

With NMR, there is the option of solid-state NMR and with the new developments over the past decades, it has become feasible to obtain high resolution spectra from unprocessed samples. However, in the technique of HR-MAS NMR, the sample needs to be introduced in a special holder, and in some way, the tissue should be stabilized, for example, for the liver and the kidney samples snap-freezing was applied (Waters *et al*., 2000). By selecting the pulse sequence in these samples, signals from macromolecules can be suppressed, yielding spectra comparable to liquid NMR. [1]H-HRMAS was also used for the metabolic profiling of *C. elegans* (Blaise *et al*., 2009). The robustness of the method was shown in the distinction of genetically modified strains.

If the sample is dry or lyophilized, the spectra will be dominated by macromolecules and small molecules are difficult to detect (Marszalek *et al*., 2010).

Again, it depends on the initial question of the study, if HR-MAS NMR offers advantages or not.

3.2.3.2 DESI–MS (Desorption Electrospray Ionization Mass Spectrometry)

In DESI–MS, the sample is treated with a stream of electrically charged droplets. This leads to the generation of ions from sample constituents, which are drawn inside the mass spectrometer for analysis (Wiseman and Laughlin, 2007). Different applications of the technique have been reported, including for analysis of pharmaceutical products, drug detection, and mass spectral imaging of tissues. It is a very powerful technique, and application in metabolomics is possible.

3.2.3.3 DART-MS (Direct Analysis in Real Time Mass Spectrometry)

Similar to DESI–MS, DART-MS is an ambient desorption ionization technique. The ionization takes place outside the mass spectrometer at ambient conditions. In DART-MS, the ionization is achieved by helium metastable ions at high temperature. The DART-MS technique is more suitable for less polar compounds, but is less suitable for high molecular weight compounds, such as proteins. DART-MS is reported to be more suitable for quantitative analysis than DESI–MS and good fingerprints can be obtained from, for example, plant materials. This turns the technique quite suitable for use in metabolomics. However, in relation to sensitivity, selectivity and detection limits LC–MS or GC–MS still have advantages. A review of applications of DART-MS in food quality and safety analysis is given by Hajslova, Cajka, and Vaclavik (2011). In a study from Zhao *et al*. (2008), it was shown that the quantification of many small molecules in plasma is possible with DART-MS without sample preparation or chromatographic separation. The sensitivity is lower than that using LC–MS, and in some cases, problems occurred with matrix effects.

Independent of the method of choice are the standard precautions to avoid degradation of the metabolites during and after sample preparation (storage until analysis) and the standardization of the whole procedure.

3.2.4 Practical Aspects of Quantification

It should be kept in mind that in metabolomic experiments, the quantitative data are essential, because the effects observed between series of samples are quantitative. No matter how accurate the analytical step is, its accuracy will not replace a lack of care in the previous steps of sampling, sample preparation, and storage. To ensure reproducibility and thus conclusive results, maximum care should be taken in all steps of the analytical procedure. Some standard precautions have been mentioned earlier and here are listed some additional points for attention, which are usually taken for granted during the sample preparation procedure.

An important aspect in quantitative analysis is the way of transferring liquid or sample. Most commonly, liquids are transferred by automatic pipetting. It is, however, important to keep in mind that volumetric pipettes are designed for use with water or diluted solutions in this solvent, and if organic solvents are transferred, deviations can occur. Furthermore, the ability of the person performing the liquid transfers in operating the pipette is essential for accurate results. An error in volume transfer is not detected if no control mechanisms are in place. To maximize reproducibility in our laboratory, we generally check all volume transfers by registering the transferred mass. Weighing is also more accurate than pipetting, if, of course, a balance with sufficient precision is used.

In standard quantitative measurements, errors in volume transfer can be corrected using adequate internal standards. If an internal standard is directly added to the material to be extracted, all further quantifications can be done in relation to the added standard. However, the internal standard should behave in a similar way as other compounds to be analyzed, and as extracts contain a large variety of compounds, its use attributes limited confidence to the resulting data.

The procedure of drying and resolubilization of the extracts should, whenever possible, be avoided because it is rarely possible to redissolve a crude extract completely and in a reproducible manner. Drying frequently exposes the samples to unnecessary heat that can lead to degradation. For this, a sufficient amount of sample should be extracted with a minimum amount of solvent, both considering the analytical method chosen.

3.3 Sampling of the Material to Be Analyzed

In the experimental design, sampling of the material to be analyzed is crucial. If plant material is collected

and the objective of the study is to investigate the influence on the metabolome of a single factor, for example, what is the influence of a certain parasite on the metabolism of a plant, or why certain plants are resistant to disease, one should take a lot of care when sampling the two populations of plants, because many other factors might have an influence. Interfering factors to be considered are, for example, type of soil, solar illumination, and proximity of other species.

It should also be considered that the population to be sampled is not homogeneous, within a plant variation exists between leafs, between twigs and leafs, in different parts of the leaf, and even within the individual cells.

If leaves are collected, what is the position of the leaves, are they young or old? It all has influence. When the samples are analyzed, generally small parts of the total sample are taken for processing. This can be as little as a few milligrams. If the total sample has small particle sizes and is well homogenized, the variation between individual subsamples might be small; however, if the particle size is bigger, a single subsample for analysis might show much variation compared with other subsamples.

It means that sufficient subsamples should be analyzed to determine the variation within the sample and within the population. Only then, a reasonable comparison is possible with samples from another population.

Adequate sampling will provide samples that are, as much as possible, comparable and preferably only differ in the factor under investigation. This will greatly facilitate the drawing of adequate conclusions from a metabolomic study.

As explained, for any existing species, organ, or cell type, there will be a metabolome plasticity. Plants and microorganisms impose an additional challenge owing to their ability to produce secondary metabolites in often cryptic metabolic routes (e.g., for production of phytoalexins). Thus, in the experimental design, adequate sampling of the material to be analyzed is crucial in order to obtain conclusive results. The variability found within the reference samples should be measured and be smaller than the ones found between the reference and treated groups in order to obtain conclusive results.

For instance, Blaise *et al.* (2009) investigated differences in the metabolome of wild-type and mutant nematodes. For this, the experimental design ensured that the variability in the collected data linked to genetic differences was larger than any variability connected with any other possible confounding effects of technical origin such as sample culturing, freezing, and NMR acquisition. Confounding effects from biological origin such as intersample and interindividual variability were minimized using samples consisting of hundreds of individuals. This strategy can be used with relative ease with small organisms or microorganisms (often clonal cultures).

Plasticity within a single species was systematically verified by Houshyani *et al.* (2011), who investigated metabolic distances of *A. thaliana* populations sown from seeds of different origins in noninduced metabolic states. The authors used the independent replicates of pooled material collected from several plants of the different accessions investigated as samples and thus were able to assess intersample variability. Metabolic differences were found between the nine accessions investigated and four were selected as representing the metabolome variation across the group. In some cases, primary metabolites were identified among the most discriminating compounds and suggested that this was due to fundamental differences in the central C metabolism. There was only a weak correlation between the genetic and the metabolic distances, and thus it could be concluded that genetic diversity does not always lead to metabolic differences.

3.4 Measurement of Spectra and/or Chromatograms

After designing the best experimental setup and collecting the samples, one arrives at the next stage: the measurement of samples, obtaining spectra, and/or chromatograms.

In the measurements, care should be taken to obtain the best possible results from the equipment. This requires adequate procedures.

3.4.1 NMR Spectroscopy

For obtaining the best NMR spectra, the equipment should be optimally tuned and shimmed. This will provide maximum sensitivity on the specific equipment and narrow signal lines.

The increased resolution of signals will enable to detect the maximum number of compounds, with the best signal-to-noise ratio.

As mentioned earlier, sensitivity and resolution are most important in NMR. The following formula expresses the relation of some common factors with NMR sensitivity.

$$S/N = c \cdot n \cdot \gamma_e \sqrt{(\gamma_d{}^3 \cdot B_o{}^3 \cdot t)}$$

In this formula,

c = constant dependent on the equipment and experimental conditions (see below);
n = number of spins observed;
γ_e = gyromagnetic ratio of the excited spin;
γ_d = gyromagnetic ratio of the detected spin;
t = experiment acquisition time;
B_o = magnetic field strength.

Using this formula, the effects of these factors can be evaluated.

- Concentration: It corresponds with n in the formula. With a fixed volume, n doubles if the concentration doubles. This means also that the S/N ratio increases with a factor of 2. There is thus a direct relation between the concentration and the S/N ratio. It is important here to point out that if one has a specific mass of sample, the most straightforward way to increase the S/N ratio is to maintain the highest concentration possible. In an NMR tube of 5 mm, it is recommended to use a sample volume of 0.5 mL. If the sample is dissolved in 1 mL, only half the S/N ratio is obtained.
- Number of scans: In the NMR experiment, the time for a single pulse consisting of the pulse (of a few microseconds), acquisition, and relaxation delay generally takes about 5 s. The pulses can be repeated many times and the resulting NMR FIDs (free induction decay) can be summed, yielding an improved S/N ratio. The number of scans corresponds in the formula with the time t. With more time, more scans are obtained. The noise in the scans also sums up; however, because it is random, it only increases with the square root of the number of scans, whereas the signal intensity increases directly with the number of scans. Consequently, an experimental time four times as long is required to obtain an increase in the S/N ratio of a factor of 2. Therefore, if a sample is dissolved in twice the volume, the experimental time to obtain the same S/N ratio increases with a factor of 4.
- Field strength of the magnet, B_o in the formula, influences the NMR sensitivity through different effects: first of all because of higher energy differences between the spin orientations (the Larmor frequency is directly related to B_o). Furthermore, the Boltzman distribution is influenced by the higher energy difference, increasing the population difference. Owing to the influence of the field strength, one gains about a factor of 2 in S/N if an apparatus of 600 MHz is used, instead of an apparatus of 400 MHz.

Besides the factors mentioned earlier, also the following factors have a direct effect on NMR sensitivity.

- *Homogeneity of the Magnetic Field*. If the magnetic field is not homogeneous within the region of the sample, broadened or distorted signals are obtained. This gives loss in the S/N ratio and resolution.
- *Relaxation Time*. After a pulse, the magnetization should return to equilibrium to obtain the full signal intensity with the next pulse. Especially for slowly relaxing nuclei (for instance, tetramethylsilane (TMS) with a relaxation delay in ^{1}H NMR of about 6 s), a long delay is required (at least 5× the relaxation delay).
- *Acquisition Time*. The acquisition time is the time during which the FID is recorded. If chosen too long, only noise is recorded and the S/N ratio decreases. Preferentially, a sufficiently long acquisition time is taken, and the digital resolution of the spectrum is improved by zero-filling.
- *Cryoprobes*. These probes were developed some decades ago. Because the transmission and receiver coils are maintained at low temperatures, the noise contributions from random thermal motion of electrons have been reduced; furthermore, the decrease of resistance provides a much more sensitive detection. These probes provide a direct gain in the S/N ratio of about a factor of 5.

One of the major advantages of NMR for its use in metabolomics is that the NMR signal integrals are directly related to the number of nuclei of the signal.

However, to obtain very accurate quantitations, the NMR spectra should be measured carefully, adjusting a large number of parameters (Pauli, Jaki, and Lankin, 2007; Bharti and Roy, 2012). This use of NMR for quantitative purposes is known as *quantitative nuclear magnetic resonance* (*qNMR*). Recent applications are reviewed by Pauli *et al.* (2012). It is true that for very accurate absolute quantifications, the parameters should be carefully selected, but measurement of spectra in this way increases the experimental time with at least a factor of 2. At the other side for normal metabolomic experiments, the absolute and very accurate quantification is not essential. It is essential that between individual spectra of metabolomic experiments, no variation exists, and that the relative quantification of any component in relation to other spectra is guaranteed. Deviations in absolute quantification might be as high as several tens of percents; however, with accurate registration of experimental parameters, posterior calibration is possible.

In the 1D ^{1}H NMR spectra of mixtures, superposition of signals is common and this limits the number of compounds that can be detected. LC–NMR and especially LC–solid phase extraction (SPE)–NMR are interesting developments (Sturm and Seger, 2012) but routine application in metabolomics does not offer any advantages, compared to, for example, 2D-NMR.

3.4.2 LC–MS and GC–MS

In both LC–MS and GC–MS analyses, a chromatographic separation precedes the MS.

For metabolomics, the reproducibility of both the chromatographic separation and the MS is important.

3.4.2.1 Liquid Chromatography

For LC, octadecyl columns are used generally in combination with a mobile phase gradient with increasing percentages of acetonitrile or methanol in an aqueous acid solution. An alternative more focused on polar metabolites is HILIC, where polar columns are used, with an opposite gradient compared to reversed phase chromatography (Gika *et al.*, 2012).

With the same column, good reproducibility is possible, but the temperature should be controlled. However, with the aging of a column, its efficiency decreases with changes in retention times and peak shapes. When other columns are used, especially if they are of another brand, large differences in retention times are possible and even inversion of peak order can occur. Important for metabolomic experiments is therefore the measurement of all samples of an experiment sequentially on the same column. For verification of the good function of the column and system, it is therefore recommended to run standard samples between other samples.

Tistaert, Dejaeger, and Vander Heyden (2011) published an extensive review about chromatographic separation techniques and data handling methods for herbal fingerprints.

There are some general problems in chromatographic analysis. Very polar compounds might end up in the injection peak, whereas more apolar compounds might not be eluted and stay on the column.

3.4.2.2 Gas Chromatography

In GC, the same problems exist in relation to reproducibility of chromatograms. The main factor determining the retention times is the volatility of the compounds. Compounds without volatility are not eluted. In addition, compounds that are not stable at higher temperatures are not analyzed. Over the years, the range of compounds that can be analyzed by GC has increased tremendously. Originally, only apolar compounds with considerable volatility, such as essential oils, could be analyzed. With the introduction of capillary columns, the range increased; for example, in 1991, a method was reported for the analysis of monoterpenoid indole alkaloids (Dagnino *et al.*, 1991). New derivatization reagents gave another increase.

In metabolomic studies, it is most common to perform a two-step derivatization. Firstly, a methoximation to inhibit the ring formation of reducing sugars, protecting also all other aldehydes and ketones. Secondly, a trimethylsilylation for derivatization of acidic protons (Kind *et al.*, 2009).

A new development in GC is 2D-GC (Marriott and Shellie, 2002). This provides a considerably enhanced peak capacity (Kiefl *et al.*, 2012). However, automatic cross-sample analysis does get more complicated. This problem was addressed by Kiefl *et al.* (2012) in the analysis of the comprehensive two-dimensional GC–MS data obtained by fingerprinting raw and roasted hazelnuts.

3.4.2.3   Mass Spectrometry

The mass spectrometer coupled to GC is generally based on electron impact. The resulting mass spectra are very characteristic for the individual molecules, and because the technique is already older, large databases exist of electron impact spectra (e.g., the most widely used National Institute of Standards and Technology (NIST) mass spectral database). Specific for metabolomics and for samples that went through a specific derivatization process, with methoximation and trimethylsilylation, a database (FiehnLib) was constructed, which contains at the moment 5598 individual spectra, together with the retention index for the individual compounds (Kind *et al.*, 2009). This library was reported to contain much more compounds and spectra than the public Golm Metabolome Database (Kopka *et al.*, 2005).

Electrospray mass spectrometers are generally coupled to LC and the resulting mass spectra are much less informative and no extensive databases exist.

Common to mass spectrometric detection is the difficulty in exact quantification of individual compounds. To get accurate quantification, a calibration should be performed for each individual compound. It is also not a general analysis method, because the peak intensity depends mainly on the ionization efficiency, and some compounds cannot be ionized at all, turning them invisible. Recently, Xiao, Zhou, and Resom (2011) published a review about metabolite identification and quantification in LC–MS-based metabolomics.

3.5   Data Pretreatment

To be able to analyze the data from a metabolomic experiment, the data from each individual analysis need to be converted to a format that permits the subsequent analysis with multivariate analysis.

In NMR spectra, variation can occur in peak width or shape, generally because of shimming defects. In addition, the position of signals can vary because of differences in the pH of the sample or ionic strength. By preprocessing of the data, corrections can be made. For NMR data, three common methods of data pretreatment can be distinguished.

1. Binning or bucketing. In this process, the spectrum is divided into discrete regions, which are individually integrated.

2. Peak picking, registering the peaks and their intensities or integrals. However, to be able to compare different spectra, a spectral alignment should be performed.
3. Spectral deconvolution. In this approach, the spectrum is deconvoluted in the individual spectra of the mixture components.

For being the simplest method, binning or bucketing is most commonly used. However, it leads to a tremendous loss of resolution and the information of minor peaks close to larger peaks is lost. The preferred method would be spectral deconvolution in which the individual components are identified and quantified. Some automatic procedures have been reported. Recently, MacKinnon *et al.* (2013) presented a graphical user interface package, MetaboID, for the assignment of ^{1}H NMR spectra of body fluids and tissues. The interface works with a library of about 400 compounds. In the research involving plant material, databases will help in analyzing the primary metabolites, but the wide occurrence of secondary metabolites, which are often specific to certain plant families or even species, presents an additional challenge.

In chromatographic data, the peaks are recorded with their integral or intensity. For further analysis, an alignment procedure is essential. Furthermore, the peak identities should be confirmed with the mass spectral data. In the review of Katajamaa and Oresic (2007), the different steps in this procedure are discussed extensively. These steps involve filtering, feature detection, alignment, and normalization. Filtering is aimed at noise reduction and baseline correction. In feature detection, the true signals are identified in the chromatographic trace by analysis of the mass spectral data. Alignment is necessary to adjust the retention times, which always show some variation between runs, owing to concentration effects, pH, or other factors. Normalization is necessary to correct for variation in peak intensities. Usually, internal standards are used for this purpose.

Accurate quantification is one of the major problems in MS. Mass spectral intensities vary according to the design of the equipment, internal voltage settings, and can also be influenced by co-occurring of other compounds. Suitable internal standards are therefore essential, and in fact, for each compound, a specific calibration should be performed. In MS, it is a common practice to use isotope-labeled standards for absolute quantifications of individual compounds;

however, because in metabolomics mixtures are analyzed, many standards are necessary. For the analysis of microbial metabolomes, the use of standardized uniformly ^{13}C-labeled extracts has been proposed (Wu *et al.*, 2005).

Many software packages are available for the data pretreatment and/or subsequent multivariate analysis. Blekherman *et al.* (2011) list a series of software packages that are now available. Both commercial and freely available packages are available. The limitation of most commercial software packages, provided by equipment manufacturers is that they only work with specific data formats (Blekherman *et al.*, 2011). In 2007, Katajamaa and Oresic reviewed data-processing software for MS-based metabolomics.

3.6 Multivariate Analysis

After the adequate organization of the experimental data in spreadsheets, they can be submitted to multivariate analysis. Specific software packages are available, which can perform different modifications.

The simplest way of multivariate analysis is principle component analysis (PCA). In this analysis, the variables are identified, which show most variation within the dataset.

If the data are analyzed in relation to independent variables, one talks about partial least square (PLS) analysis. In this case, the data that show most variation in relation to the independent variable are identified.

Much care should be taken with the results provided by this multivariate analysis, and adequate checks should be performed to verify the validity.

A common problem is overfitting. This is likely to occur when samples have many variables and the number of samples is limited. By pure chance, one of these variables can show a perfect correlation with the independent variable. Statistical strategies have been pointed out to avoid and detect this overfitting (Broadhurst and Kell, 2006). In addition, repeating the experiment will most probably show that the previously detected correlation does not exist. One should always be aware of these false positives. In these cases, it is also very important to interpret the conclusions, by identifying the compounds that are responsible for the distinguishing samples.

3.7 Interpretation of Factors

3.7.1 Identification of Specific Metabolites

The final interpretation of the data should involve the identification of the compounds that cause the differences between datasets. This is generally one of the most complex parts of the metabolomics studies. In this aspect, NMR has large advantages, because the structure of the compounds is clearly reflected in its NMR spectrum. One problem is that in the NMR spectrum from the extract at the most few peaks from a compound are clearly visible and resolved (Figure 4).

For interpretation of the results with NMR spectroscopy, one needs to use spectroscopy of other nuclei, especially ^{13}C and two-dimensional techniques. With those techniques, individual signals in the 1D ^{1}H spectrum can be assigned to specific compounds (Figure 5).

By heteronuclear multiple-quantum correlation (HMQC) or heteronuclear single-quantum correlation (HSQC), one can find out to which carbon the hydrogen is linked. Through correlation spectroscopy (COSY) spectra, one can find out to which other nearby hydrogens the hydrogen is coupled. Moreover, space interactions are analyzed with nuclear Overhauser effect spectroscopy (NOESY) (1D NOE or 2D NOESY spectra).

Three to four bond interactions of hydrogen with other carbons are analyzed with the heteronuclear multiple-bond correlation (HMBC) spectrum.

Using the complete set of spectra, even in crude mixtures, the presence of many compounds can be confirmed (Figures 4 and 5). However, if completely new structures are encountered, the isolation of the compound will generally be necessary in order to adequately identify the new structure and obtain the complete set of spectroscopic data. The process of structure elucidation gets easier with the availability of specialized software. When sufficient spectroscopic data are available, the complete structure can often be elucidated automatically (Elyashberg *et al.*, 2009). However, often owing to limited availability of the compound, only a limited amount of data can be obtained and the structure assignment remains a manual process. Over time, many structures have been reported, and many of them were subsequently revised through new interpretation of data (Schripsema and Dagnino, 1996) or through
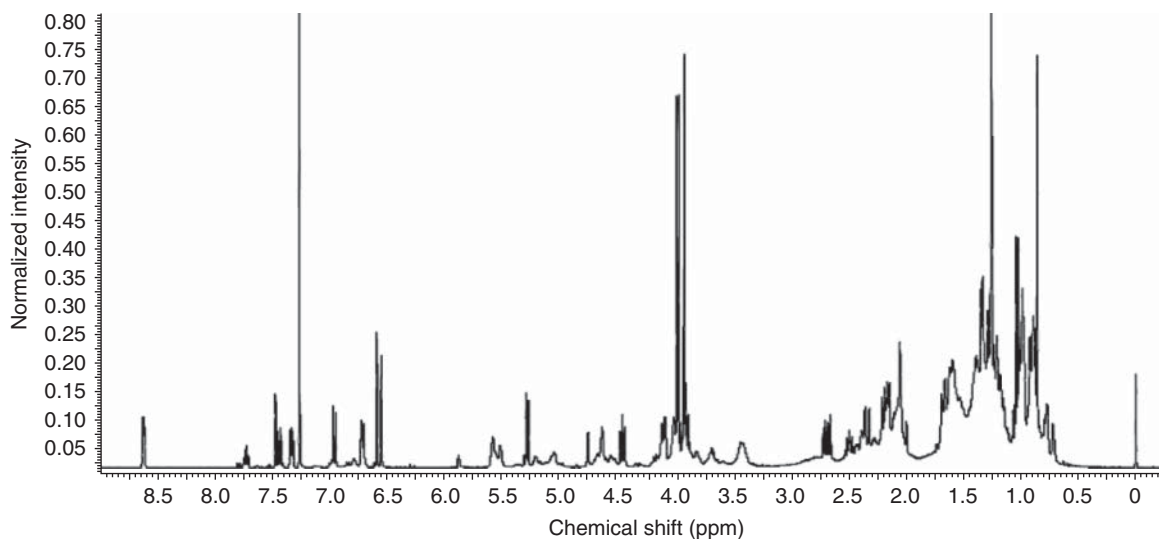
Figure 4 The ^{1}H NMR spectrum of a CDCl$_3$ extract from *Baccharis trimera*.
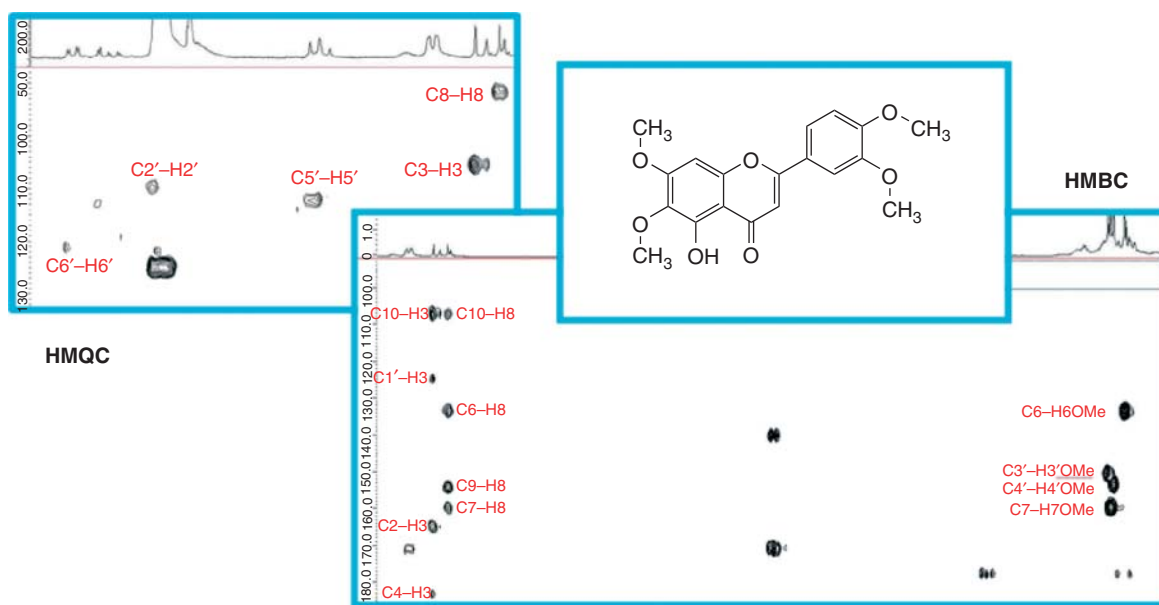
Figure 5 Details from the HSQC and HMBC spectra from a CDCl$_3$ extract from *Baccharis trimera*. Correlations are observed between specific hydrogen atoms with specific carbon atoms. These correlations permit the assignment of most chemical shifts and the deduction of the chemical structure of the flavonoid, present in the crude extract of which the ^{1}H NMR spectrum is shown in Figure 4.

synthesis of the structures (Nicolaou and Snyder, 2005).

About the best choice of specific pulse sequences, some nice papers have been published (Reynolds and Enriquez, 2002).

The identification of compounds by MS is much more complicated. LC–MS provides a retention time and a mass spectrum. These data are very useful for the recognition of known compounds, but if the compound is unknown, NMR spectroscopy

will be necessary to elucidate the structure. A recent paper from Funari *et al*. (2012) describes the dereplication strategy in ultra high performance liquid chromatography–photo diode array-time of flight-mass spectroscopy (UHPLC–PDA-TOF-MS) metabolite profiling of Brazilian *Lippia* species. In this case, the UV (ultraviolet) spectra were also available. Efficient LC peak annotation was done using a database with previously reported compounds, the matching of UV spectra and the prediction of the chromatographic behavior with log *P* calculations.

The MetaboAnalyst integrated web-based platform for comprehensive analysis of quantitative metabolomic data was extensively described by Xia and Wishart (2011). This platform offers next to processing, data normalization and statistical analysis of all types of metabolomics data, the possibility of pathway analysis and peak annotation.

3.7.2 *Interpretation of Metabolic Changes*

With the identification of the metabolites that show variations within the metabolomics investigation, possibilities are obtained to interpret the metabolic changes.

In a recent review by Obata and Fernie (2012), the use of metabolomics to dissect the plant responses to abiotic stress was addressed. The changes in metabolite levels permitted conclusions about the pathways affected by the different types of stress.

Houshyani *et al*. (2011) characterized the variation within nine accessions from *A. thaliana*, grown under various conditions. A statistical method was introduced for estimating the metabolic distance between genotypes or treatments.

Carvalho *et al*. (2010) investigated the co-catabolism of carbon substrates in *Mycobacterium tuberculosis*. After observing the simultaneous catabolism of different carbon sources, further experiments with ^{13}C-labeled substrates revealed that besides the rather unusual co-catabolism, an unusual topologic organization of the central carbon metabolism exists.

4 FUTURE DIRECTIONS OF METABOLOMICS

In metabolomics, samples are analyzed for their metabolic fingerprints. However, how is the homogeneity of the sample? Plants consist of different parts, such as roots, stems, leaves, and flowers. Each of these parts provides a specific and different metabolic fingerprint.

This also applies to different parts of the organs; a leaf has different types of cells and each type of cell provides a specific metabolic fingerprint. Within the cells, different compartments can be distinguished and again differences exist.

It will certainly be one of the great challenges of metabolomics to show and study the differences of metabolic fingerprints at the level of organs, cells, and subcellular compartments. Recently, Rubakhin, Lanni, and Sweedler (2013) reviewed the progress toward single-cell metabolomics, especially based on the application of capillary separation techniques and MS detection. However, the range of metabolites observed is still rather limited. Another review published by Moco, Schneider, and Vervoort (2009) focused on laser-assisted microdissection to obtain samples from specific cell types, which could then be studied by LC–MS and NMR. Studies have been published already in which it was possible to investigate the distribution of specific molecules at the cellular level. Hoelscher *et al*. (2009) reported the localization of different secondary metabolites in *A. thaliana* and *Hypericum* species. With laser desorption/ionization mass spectrometry (LDI-MS), images could be produced showing the distribution of specific characteristic masses within the tissue. In *Hypericum* species, the distribution of flavonoids, biflavonoids, and hypericins could be verified. In *A. thaliana*, the distribution of flavonoids within the flower petals and sepals was studied.

Despite the rather insensitivity of NMR, it has been used to study the content of individual cells. Laser-microdissected trichomes of medicinal *Cannabis sativa* were studied using LC–MS and cryogenic NMR (Happyana *et al*., 2013). Cannabinoids were detected in extracts of 25–143 collected cells of capitate-sessile and capitate-stalked trichomes. NMR permitted the identification of cannabichromenic acid in the extracts.

Another thing to be remembered is that metabolic fingerprints are snapshots of the metabolism. What is the timescale at which the metabolic fingerprints are changing? In microorganisms, changes in the timescale of seconds are reported (de Jonge *et al*., 2012). In plants, the changes are supposed not to be so rapid, but rapid changes are not impossible and might occur after certain stimuli, for example, *Mimosa pudica* reacts within seconds on a mechanical touch.

The study of the changes of metabolic fingerprints over time might reveal important information. The new area of research fluxomics is trying to obtain the information about the actual fluxes of metabolites through the metabolic pathways. A series of metabolic fingerprints through time will provide information about changes in the quantities of compounds, but does not actually reflect the metabolic fluxes, because the quantities of metabolites present are the result of co-occurring reactions of biosynthesis and breakdown. To measure the velocity of these co-occurring reactions, labeled precursors are used. In early work with cell suspension cultures from *Tabernaemontana divaricata* and *Catharanthus roseus*, the biosynthesis and breakdown of indole alkaloids were measured through administration of ^{15}N- or ^{2}H-labeled indole alkaloids (Dagnino, Schripsema, and Verpoorte, 1993; Schripsema *et al*., 1994). In fluxomics research, it is usual to utilize ^{13}C-labeled precursors. In a recent paper by Antoniewicz (2013), the optimal design of isotopic labeling experiments for measuring fluxes by ^{13}C metabolic flux analysis was discussed. A dramatic improvement in the quality of the results can be obtained with parallel labeling experiments and rational design of tracers.

The investigation in all areas of cell biology, genomics, transcriptomics, proteomics, metabolomics, and fluxomics will provide a more complete view on the systems biology. Links between these different areas of research will provide the means for direct applications in medicine, biology, and agriculture.

Genome-wide association studies (GWAS) show the relations between genetic variance and predisposition to disease (Adamski and Suhre, 2013). With metabolomics, the genetic factors can be identified, which influence the metabolite levels. In humans, these studies are already in an advanced state. Recently, a quite complete roadmap of human metabolism was published (Thiele *et al*., 2013). This model of human metabolism contains 5063 metabolites and 7440 reactions. It permits the prediction of metabolic perturbations caused by drugs or inborn errors with a high accuracy.

Many challenges exist. One of the challenges in human metabolomics studies is to reduce the variation between subjects, for example, gut microflora and diet have a big influence on the human metabolome (Johnson and Gonzalez, 2012). Similarly, in plants, the influence of endophytic microorganisms is still largely unknown (Kusari *et al*., 2009).

5 CONCLUSIONS

It should always be remembered that errors made in the sample collection, preparation, and measurement cannot be corrected afterwards and will cause prejudice to the outcome of metabolomic experiments. The difference between two series of samples generally is not the presence or absence of a single marker compound, but a quantitative difference. Any factor negatively influencing the precision of experiments might obscure the quantitative differences between series. Metabolomics has the potential to supply essential information to many areas of research, and can be considered a key area for the further development of functional genomics and systems biology. Poste (2011) stated that many disease-associated biomarkers are awaiting discovery, but what is lacking is a coordinated big science approach. This also applies to many other areas of research, for example, the investigation of agricultural crops, ethnopharmacology (Heinrich, 2010), or food analysis (Mannina et al., 2012).

6 RELATED ARTICLES

Extraction Methodologies: General Introduction; NMR as Analytical Tool for Crude Plant Extracts; NMR of Small Molecules; Multivariate data analysis; On-line and At-line LC-NMR and Related Micro NMR Methods.

REFERENCES

Adahchour, M., Beens, J., Vreuls, R. J. J., *et al*. (2006) *Trends Anal. Chem.*, **25**, 438–454.

Adamski, J., and Suhre, K. (2013) *Curr. Opin. Biotechnol.*, **24**, 39–47.

Antoniewicz, M. R. (2013) *Curr. Opin. Biotechnol.*, **24**, 1116–1121.

Beltran, A., Suarez, M., Rodríguez, M. A., *et al*. (2012) *Anal. Chem.*, **84**, 5838–5844.

Bharti, S. K., and Roy, R. (2012) *Trends Anal. Chem.*, **35**, 5–26.

Bino, R. J., Hall, R. D., Fiehn, O., *et al*. (2004) *Trends Plant Sci.*, **9**, 418–425.

Blaise, B. J., Giacomotto, J., Triba, M. N., *et al*. (2009) *J. Proteome Res.*, **8**, 2542–2550.

Blekherman, G., Laubenbacher, R., Cortes, D. F., *et al*. (2011) *Metabolomics*, **7**, 329–343.

Broadhurst, D. I., and Kell, D. B. (2006) *Metabolomics*, **2**, 171–196.

Buscher, J. M., Czernik, D., Ewald, J. C., *et al*. (2009) *Anal. Chem.*, **81**, 2135–2143.

Carvalho, L. P., Fischer, S. M., Marrero, J., *et al*. (2010) *Chem. Biol.*, **17**, 1122–1131.

Dagnino, D., Schripsema, J., Peltenburg, A., *et al*. (1991) *J. Nat. Prod.*, **54**, 1558–1563.

Dagnino, D., Schripsema, J., and Verpoorte, R. (1993) *Phytochemistry*, **32**, 325–329.

Davies, H. (2010) *Food Control*, **21**, 1601–1610.

Dettmer, K., Aronov, P. A., and Hammock, B. D. (2007) *Mass Spectrom. Rev.*, **26**, 51–78.

DiLeo, M. V., Strahan, G. D., den Bakker, M., *et al*. (2011) *PLoS One*, **6**pp. e26683, 1–10.

Elyashberg, M., Blinov, K., Molodtsov, S., *et al*. (2009) *J. Cheminform.*, **1**, 3.

Funari, C. S., Eugster, P. J., Martel, S., *et al*. (2012) *J. Chromatogr. A*, **1259**, 167–178.

Geier, F. M., Want, E. J., Leroi, A. M., *et al*. (2011) *Anal. Chem.*, **83**, 3730–3736.

Gika, H. G., Theodoridis, G. A., Vrhovsek, U., *et al*. (2012) *J. Chromatogr. A*, **1259**, 121–127.

Hajslova, J., Cajka, T., and Vaclavik, L. (2011) *Trends Anal. Chem.*, **30**, 204–218.

Happyana, N., Agnolet, S., Muntendam, R., *et al*. (2013) *Phytochemistry*, **87**, 51–59.

Heinrich, M. (2010) *Front. Pharmacol.*, **1**, 1–3.

Hoelscher, D., Shroff, R., Knop, K., *et al*. (2009) *Plant J.*, **60**, 907–918.

Houshyani, B., Kabouw, P., Muth, D., *et al*. (2011) *Metabolomics*, **8**, S131–S145.

Johnson, C. H., and Gonzalez, F. J. (2012) *J. Cell. Physiol.*, **227**, 2975–2981.

de Jonge, L. P., Douma, R. D., Heijnen, J. J., *et al*. (2012) *Metabolomics*, **8**, 727–735.

Junot, C., Madalinski, G., Tabet, J. C., *et al*. (2010) *Analyst*, **135**, 2203–2219.

Katajamaa, M., and Oresic, M. (2007) *J. Chromatogr. A*, **1158**, 318–328.

Kiefl, J., Cordero, C., Nicolotti, L., *et al*. (2012) *J. Chromatogr. A*, **1243**, 81–90.

Kim, H. K., and Verpoorte, R. (2010) *Phytochem. Anal.*, **21**, 4–13.

Kind, T., Wohlgemuth, G., Lee, D. Y., *et al*. (2009) *Anal. Chem.*, **81**, 10038–10048.

Kopka, J., Schauer, N., Krueger, S., *et al*. (2005) *Bioinformatics*, **21**, 1635–1638.

Kusari, S., Zuehlke, S., Kosuth, J., *et al*. (2009) *J. Nat. Prod.*, **72**, 1825–1835.

MacKinnon, N., Somashekar, B. S., Tripathi, P., *et al*. (2013) *J. Magn. Reson.*, **226**, 93–99.

Mannina, L., Sobolev, A. P., and Viel, S. (2012) *Prog. Nucl. Magn. Reson. Spectrosc.*, **66**, 1–39.

Marriott, P., and Shellie, R. (2002) *Trends Anal. Chem.*, **21**, 573–583.

Marszalek, R., Pisklak, M., Horsztynski, D., *et al*. (2010) *Solid State Nucl. Magn. Reson.*, **37**, 21–27.

Moco, S., Schneider, B., and Vervoort, J. (2009) *J. Proteome Res.*, **8**, 1694–1703.

Nicolaou, K. C., and Snyder, S. A. (2005) *Angew. Chem. Int. Ed.*, **44**, 1012–1044.

Obata, T., and Fernie, A. R. (2012) *Cell. Mol. Life Sci.*, **69**, 3225–3243.

Oliver, S. G., Winson, M. K., Kell, D. B., *et al*. (1998) *Trends Biotechnol.*, **16**, 373–378.

Pauli, G. F., Jaki, B. U., and Lankin, D. C. (2007) *J. Nat. Prod.*, **70**, 589–595.

Pauli, G. F., Gödecke, T, Jaki, B. U., and Lankin, D. C. (2012) *J. Nat. Prod.*, **75**, 834–851.

Poste, G. (2011) *Nature*, **469**, 156–157.

Quanbeck, S. M., Brachova, L., Campbell, A. A., *et al*. (2012) *Front. Plant Sci.*, **3** (15), 1–12.

Reynolds, W. F., and Enriquez, R. G. (2002) *J. Nat. Prod.*, **65**, 221–244.

Rubakhin, S. S., Lanni, E. J., and Sweedler, J. V. (2013) *Curr. Opin. Biotechnol.*, **24**, 95–104.

Sardans, J., Penuelas, J., and Rivas-Ubach, A. (2011) *Chemoecology*, **21**, 191–225.

Saric, J., Want, E. J., Duthaler, U., *et al*. (2012) *Anal. Chem.*, **84**, 6963–6972.

Schripsema, J. (2008) *J. Agric. Food Chem.*, **56**, 2547–2552.

Schripsema, J. (2010) *Phytochem. Anal.*, **21**, 14–21.

Schripsema, J., and Dagnino, D. (1996) *Phytochemistry*, **42**, 177–184.

Schripsema, J., Dagnino, D., dos Santos, R. I., *et al*. (1994) *Plant Cell Tissue Organ Cult.*, **38**, 299–305.

Schripsema, J., Vianna, M. D., Rodrigues, P. A. B., *et al*. (2010) *Acta Hort.*, **851**, 505–511.

Sturm, S., and Seger, C. (2012) *J. Chromatogr. A*, **1259**, 50–61.

Thiele, I., Swainston, N., Fleming, R. M., *et al*. (2013) *Nat. Biotechnol.*, **31**, 419–425.

Tistaert, C., Dejaegher, B., and Vander Heyden, Y. (2011) *Anal. Chim. Acta*, **690**, 148–161.

Ward, J. L., Harris, C., Lewis, J., *et al*. (2003) *Phytochemistry*, **62**, 949–957.

Waters, N. J., Garrod, S., Farrant, R. D., *et al*. (2000) *Anal. Biochem.*, **282**, 16–23.

Wiseman, J. M., and Laughlin, B. C. (2007) *Curr. Separ. Drug Dev.*, **22**, 11–14.

Wu, L., Mashego, M. R., van Dam, J. C., *et al*. (2005) *Anal. Biochem.*, **336**, 164–171.

Xia, J., and Wishart, D. S. (2011) *Nat. Protoc.*, **6**, 743–760.

Xiao, J. F., Zhou, B., and Ressom, H. W. (2011) *Trends Anal. Chem.*, **32**, 1–14.

Zhao, Y., Lam, M., Wu, D., *et al*. (2008) *Rapid Commun. Mass Spectrom.*, **22**, 3217–3224.

Proteomics and Its Research Techniques in Plants

Liming Yang[1], Haibin Xu[2] and Shilin Chen[3]

[1]*Huaiyin Normal University, Huaian, Jiangsu, PR China* , [2]*The Key Laboratory of Bioactive Substances and Resources Utilization of Chinese Herbal Medicine, Ministry of Education, Institute of Medicinal Plant Development, Beijing, PR China* and [3]*The National Engineering Laboratory for Breeding of Endangered Medicinal Materials, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, PR China*

1 INTRODUCTION

With the implementation and completion of increasing plant genome sequencing projects, the focus of life science research has shifted to gene identification, function distributions, interaction networks, and also extended to systems biology.

Theoretically, genome analysis can help us make predictions, but it cannot reflect spatial and temporal specificities or regulatory network relationships of gene expressions during the life process of an organism. It also cannot reflect the multiple-level regulatory relationships among DNA, mRNA, and protein, including transcription-level, post-transcription-level, translation-level, and post-translation-level relationships. There is no definite correlation between the abundance of mRNA from gene transcription and that of proteins from translation, especially for low abundance proteins. Gene expressions and modes of interaction are affected by everything from post-translation modifications of proteins, subcellular localizations, protein–protein interactions, and protein conformations to higher order structures and protein degradations. Thus, protein-level research is an important component of life science.

2 CONCEPT OF PROTEOMICS

Traditional protein research focuses on one or more proteins. It is not conducive to the understanding of protein functions and interaction networks at the organism level. It also cannot meet the large-scale genome data processing requirements of the post-genome era. For such circumstances, proteomics, which had emerged in the mid-1990s (Wasinger *et al*., 1995), can identify and quantify proteins at the organism level, and carry out protein expressions under specific time or environmental conditions, where localizations, modifications, activities, protein–protein interactions, and their functions may also be studied (Cordwell *et al*., 1997; Fields, 2001).

Proteomics is an important bridge between genomics and phenomics and is gradually becoming a focal field in life science. Furthermore, the research techniques and strategies of proteomics are increasingly popular within plant science research.

3 RESEARCH METHODS AND TECHNIQUES OF PLANT PROTEOMICS

To date, there have been some researches in plants, especially in medicinal plants, using proteomics
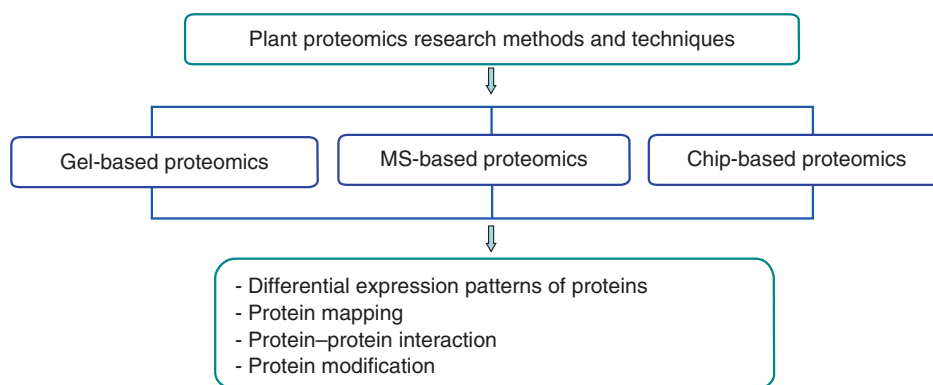
Figure 1   Strategies and techniques of plant proteomics. On the basis of different research strategies, the methods and techniques of plant proteomics mainly include gel-based proteomics, mass spectrometry-based proteomics, and chip-based proteomics. These methods and techniques are used to analyze differential expression responding to different conditions, protein–protein interactions, and protein modifications.

methods and techniques reported (Decker *et al.*, 2000; Jacobs *et al.*, 2001, 2005; Oldham *et al.*, 2010; Hew and Gam, 2011; Debnath, Pandey, and Bisen, 2011; Koay and Gam, 2011; Wu *et al.*, 2012; Yang *et al.*, 2013; Zeng *et al.*, 2013). In accordance with different proteomics research strategies, research techniques of plant proteomics can primarily be divided into gel-based proteomics, mass spectrometry (MS)-based proteomics, and chip-based proteomics (Figure 1).

3.1   Gel-Based Proteomics Analysis

The core of gel-based proteomics analysis is two-dimensional gel electrophoresis. Since O'Farrell (1975) and Klose (1975) invented 2-DE in 1975, a number of techniques have since emerged and developed as a result. Such key developments include protein separation, detection sensitivity, quantitative accuracy, mapping, and mass spectrum identification and analysis of proteins on gel have all been greatly improved (Klose, 2009).

3.1.1   Two-Dimensional Gel Electrophoresis

The basic principles of 2-DE are to conduct a first-dimension separation with an isoelectric focusing, using the difference in charge of different protein molecules and conducting the second-dimension separation with SDS–PAGE based on the difference in protein molecular weights.

In theory, on a $21 \times 25\,\text{cm}^2$ 2-DE gel, approximately 10,000 protein spots can be separated, and experimental molecular weights and isoelectric points of the protein spots can be identified. The resolution and stability of isoelectric focusing can be greatly improved by employing immobilized pH gradient gels. Homologous isomers of proteins can also be identified on the 2-DE gel.

3.1.2   Two-Dimensional Fluorescence-Difference Gel Electrophoresis

Two-dimensional fluorescence-difference gel electrophoresis (2D-DIGE) is a method of quantitatively analyzing proteins developed on the basis of traditional 2-DE (Unlu, Morgan, and Minden, 1997). The *N*-hydroxysuccinyl ester derivative of the fluorescent cyanine dyes (Cy2, Cy3, Cy5, etc.) used in the technique can be bonded to lysine residues in protein molecules. The fluorescence groups for labeling have similar chemical structures and basically the same molecular weights, and all of them are positively charged. This allows the same proteins in all the samples to be transferred to the same site when they react with lysine residues. Using this property, 2D-DIGE labels the protein samples from different sources with different fluorescent cyanine dyes. 2-DE separation is then conducted after mixing the protein samples in equal parts. The proteins labeled with different fluorescent groups are detected with the excitation wavelengths corresponding to specific

fluorescent dyes. The protein expression profile of the samples is subsequently analyzed in combination with fully automated differential protein expression analysis software (DeCyder™). This technique can compare and quantify the protein expression of multiple samples on the same gel, which improves the repeatability and quantitation accuracy of 2-DE (Tonge *et al*., 2001).

3.1.3 Multiplexed Proteomics

Multiplexed proteomics (MP) is a technique that first detects the total protein spectrum of a gel separation and then detects other functional attributes (glycosylation, acetylation, phosphorylation, drug binding capability, etc.) of these samples using fluorescent dyes with different excitation wavelengths and emission spectra. The total protein spectrum and the functional attribute spectrum are integrated, and then they are matched and analyzed to compare the protein expression level and specific functional attributes of each sample or between samples (Patton and Beechem, 2001). This technique enables protein expression analysis and protein post-translational modification research without pretreating the protein samples (Steinberg *et al*., 2003; Zhou *et al*., 2007).

3.1.4 Two-Dimensional Electrophoresis Limitations

2-DE still has some limitations, especially in analyzing protein samples. These limitations include (i) difficulties with protein loading capacity and detection sensitivity to detect low abundance proteins (Gygi *et al*., 2000); (ii) limited detection of insoluble transmembrane proteins, particularly highly hydrophobic proteins and proteins with extreme pIs or molecular weights; (Görg, Weiss, and Dunn, 2004) (iii) co-migration of different proteins and migration of proteins to different sites on the gel because of their different modifications, which both affect accurate quantitation of the protein sample.(Gygi *et al*., 2000; Görg, Weiss, and Dunn, 2004)

3.2 Mass Spectrometry-Based Proteomics

MS-based proteomics is a non-gel technique characterized by high throughput, high resolution, high speed, and full automation (Aebersold and Mann, 2003). It can fulfill most of the requirements for plant biology, and owing to the rapid development of new instruments, computational tools, and analytical strategies, it has already become an essential tool to study molecular and cellular processes in plants under diverse conditions (Sabidó, Selevsek, and Aebersold, 2012). Several new MS-based proteomics approaches have been used in plant biology, mainly differing in their analytical performance in terms of reproducibility, dynamic range, detection limitations, and resolving power (Domon and Aebersold, 2010), including multidimensional protein identification technology (MudPIT), isotope-coded affinity tagging (ICAT), isobaric tags for relative and absolute quantitation (iTRAQ), and stable isotope labeling by amino acids in cell culture (SILAC).

3.2.1 Multidimensional Protein Identification Technology

MudPIT involves conducting one-dimensional strong ion exchange chromatography, two-dimensional reversed-phase chromatographic separation, and multidimensional separation of polypeptide mixtures produced by protein enzymolysis, and finally conducting protein identification by tandem MS (Washburn, Wolters, and Yates, 2001). The protein detection efficacy of MudPIT is much higher than that of 2-DE and it is notable that they can detect different proteins, in which each can detect proteins that the other cannot (Koller *et al*., 2002).

3.2.2 Isotope-Coded Affinity Tagging

Invented by Gygi *et al*. (1999), ICAT involves four steps: (i) labeling the free sulfhydryl of cysteine residues in different protein samples with ICAT light and heavy reagents; (ii) mixing two labeled samples in equal parts and conducting enzymolysis; (iii) conducting biotin affinity chromatography separation on the labeled peptides containing cysteine residues; (iv) analyzing the separated peptides by liquid chromatography – tandem MS. Commercially, available ICAT is found in two types (D0 and D8) that differ by eight mass units and both include three parts: the first terminal is a biotin label with its affinity function, the middle is a link to which eight

hydrogen or deuterium atoms can be introduced, and the third terminal is an iodo-acetylimino structure with a group that specifically reacts with sulfhydryl, and thus can bond with the peptides containing cysteine residues. A pair of peaks that differ by eight mass units in the mass spectrum are considered to be originated from the same peptide. Therefore, according to the intensity or area of two mass spectra of the same peptide, the relative abundance of the protein between samples can be quantitatively analyzed. Compared to MudPIT, ICAT reduces sample complexity. Furthermore, mixed samples can be tested directly and low abundance or hydrophobic protein can be both qualitatively and quantitatively identified with this technique. The molecular weight of the ICAT label, however, is approximately 500 Da and remains on each peptide during mass spectrum analysis. Thus, under collision-induced dissociation, the label is easily fragmented, resulting in complications with tandem mass spectrum analysis of the labeled peptide and necessitating more complex database search algorithms.

3.2.3   Isobaric Tags for Relative and Absolute Quantitation (iTRAQ)

iTRAQ label peptides with an isobaric isotope tag and conduct relative or absolute quantitation of a specific protein on different samples (Ross *et al.*, 2004). iTRAQ consist of three parts: the first terminal is a reporter group, the middle is a balance group, and the third terminal is a peptide reactive group. There are eight types of reporter groups with different mass-to-charge ratios (*m/z*: 113, 114, 115, 116, 117, 118, 119, and 121). The peptide reactive group links the tag to the N-terminal and the lysine-side chain of a peptide. The corresponding balance group ensures that the same peptides labeled with these tags have the same mass-to-charge ratios. The same peptides in different samples have the same mass and show as a single peak in mass spectrum scanning no matter which tag is used for labeling. After the collision-induced dissociation, however, the same peptides in different samples release reporter groups with different masses. The peaks and peak areas of different reporter groups reflect the relative amount of the same protein in different samples. Therefore, the difference in protein expression levels of different samples can be analyzed simultaneously using iTRAQ reagent. Furthermore, a known

standard protein can be used as a reference to conduct absolute quantitation.

3.2.4   Advantages and Limitations of Mass Spectrometry-Based Proteomics

Relative to gel-based techniques, the advantages of mass-based proteomics include wide detection range, high resolution and sensitivity, applicability to structural analysis, and high degree of automation. Its limitations are poor tolerance to salts in samples, complex mixture maps, and sensitivity to and limited by the use of solvents.

Of course, effective proteome analysis is generally considered to depend heavily on the availability of a high quality DNA or protein reference database. As such, proteomics has long been taxonomically restricted, with limited inroads being made into the proteomes of "non-model" organisms. However, next-generation sequencing (NGS), particularly RNA-Seq, now allows deep coverage detection of expressed genes, which in turn potentially facilitates the matching of peptide mass spectra with cognate gene sequence (Lopez-Casado *et al.*, 2012; Ning, Fermin, and Nesvizhskii, 2012). Moreover, RNA-Seq provides a cost-effective and robust platform for protein identification and will be increasingly valuable to the field of proteomics (Lopez-Casado *et al.*, 2012).

3.3   Chip-Based Proteomics

Chip-based proteomics entails forming micro-arrays or chips by immobilizing bioactive molecules such as proteins and polypeptides (antibodies, antigens, ligands, receptors, enzymes, etc.) on the surface of a solid phase medium (glass, polyvinylidene difluoride membrane (PVDF), nitrocellulose membrane, etc.) in a preset arrangement (Uetz *et al.*, 2000; Espejo *et al.*, 2002). Owing to the interactions between antigens and antibodies, between receptors and ligands, and between enzymes and substrates, the chip can specifically capture the target protein from the sample and then analyze it qualitatively and quantitatively with the appropriate method. This technique can be used to study the interactions between proteins, between antibodies and antigens, between proteins and nucleic acids, between proteins and lipids,

between proteins and other micromolecules, and between enzymes and substrates at high capacity (Zhu *et al.*, 2001; Espejo *et al.*, 2002; Michaud *et al.*, 2003; Hall *et al.*, 2004; Kofler, Motzny, and Freund, 2005; Kopf, Shnitzer, and Zharhary, 2005). As proteomics develops, chip-based proteomics – characterized by high capacity, high specificity, and high sensitivity – has been successfully applied in proteomics research to find the molecules interacting with the target protein, analyze differential protein expressions and protein post-translational modifications, analyze protein domain, and detect interactions between protein and DNA or RNA (Templin *et al.*, 2003).

As proteins cannot be amplified by a reaction such as PCR, the preparation of a proteome chip is much more complicated than that of a gene chip. The limitation of this method is that purified protein is necessary, and protein purification is generally difficult; therefore, it is both time consuming and expensive.

4 PROTEIN IDENTIFICATION

Data from MS is used to search protein or DNA/EST (expressed sequence tag) sequence databases to confirm protein identities. The success of protein identification mainly depends on data quality, the conditions under which the search is performed, and the database used. These databases generally include (i) the database for peptide mass fingerprinting data that compares the peptide mass units resulting from MALDI-MS with the theoretical data from the enzymolysis of protein primary structure sequences in the database; (ii) the database for the peptide sequence tag that uses peptide sequences obtained from tandem MS; and (iii) the FASTA procedure search, mainly applicable to organisms where the genome has not been completely sequenced or the database information is incomplete (Mackey, Haystead, and Pearson, 2002).

5 PROCEDURAL STEPS FOR TWO COMMON EXPERIMENTAL TECHNIQUES IN PROTEOMICS RESEARCH

The common two-dimensional electrophoresis of gel-based proteomics research and the iTRAQ operational process of MS-based proteomics research are introduced below.

5.1 Two-Dimensional Electrophoresis Procedure

Two-dimensional electrophoresis includes the following steps: selection of solid phase pH gradient; selection of loading method; IPG strip swelling; isoelectric focusing; IPG strip equilibration; second-dimension SDS–PAGE electrophoresis; gel dyeing; gel scanning; and image analysis (Figure 2).

5.1.1 *Selection of Solid Phase pH Gradient*

Extract the protein using an appropriate method. Select an IPG strip with appropriate pH gradient range and length according to the characteristics of the separated sample and the isoelectric point distribution of the target protein.

5.1.2 *Selection of Loading Method*

Load the sample by adding it to the rehydration fluid or load directly onto the hydrated IPG strip with a sample cup or loading hole.

5.1.3 *IPG Strip Swelling*

To swell an IPG strip without samples, pipette an appropriate volume of swelling solution into the groove of a swelling tray. Carefully remove the protective surface film from the acidic end of the IPG strip. Place the strip gel surface down into the groove of the swelling tray.

To swell an IPG strip with samples, uniformly mix an appropriate amount of protein solution and swelling solution and centrifugally remove the insoluble fraction. Transfer the swelling solution–sample mixture into the groove of the swelling tray. Remove the protective film of the IPG strip, and carefully place the strip gel surface down into the groove of the swelling tray from one side. Pipette 2 mL of mineral oil, and add it slowly into the groove of the swelling tray so that it completely covers the strip and swelling solution. Close the lid of the swelling tray. Set the temperature (17–20 °C) and swell for 12–16 h.
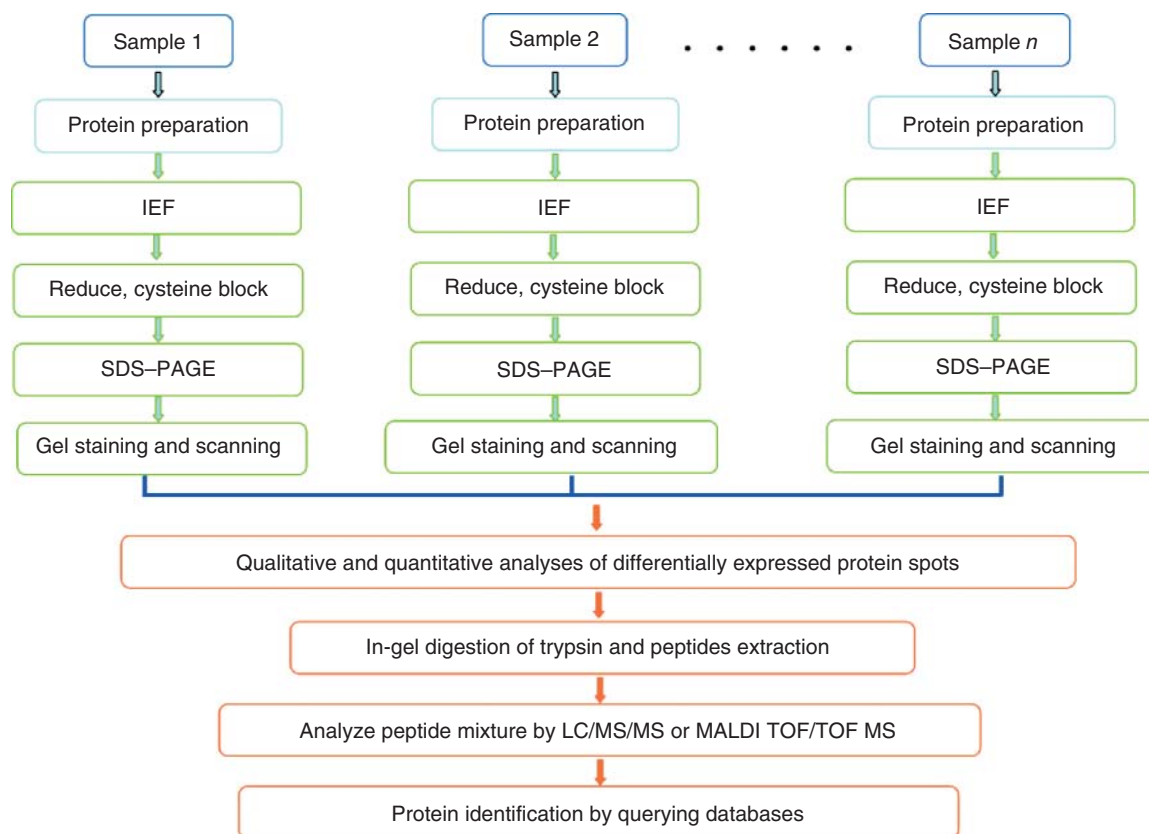
Figure 2 Schematic workflow of two-dimensional electrophoresis. Protein samples are extracted from different plant materials and subsequently separated by IEF and SDS–PAGE. Targeted protein spot is subjected to in-gel digestion, whereas the peptide mixtures are analyzed by MS, and protein is identified by querying protein databases.

5.1.4 *Isoelectric Focusing*

1. Set up the isoelectric focusing procedure based on the pH range and the length of the IPG strip;
2. Close the lid of the electrophoresis tank and conduct isoelectric focusing electrophoresis;
3. After completing the isoelectric focusing electrophoresis, turn off the power and open the lid of the electrophoresis tank, then use tweezers to remove the strip and drain the mineral oil. Place the strip with its gel surface facing upward and place on filter paper moistened with ultrapure water. To absorb excess mineral oil, gently place a piece of moistened filter paper on top of the strip, then balance and transfer the strip.

5.1.5 *IPG Strip Equilibration*

1. Place a thick, dry piece of filter paper on the table and place the focused strip on the paper with its gel surface facing upward. Wet another thick piece of filter paper with MilliQ water, squeeze out the excess water, and then gently place it directly on top of the strip to absorb mineral oil and any excess sample on the strip.
2. Transfer the strip to the swelling tray, add 5 mL of strip equilibration buffer I into the groove with the strip, and shake it slowly on a horizontal shaking table for 15 min;
3. When the first equilibration is complete, thoroughly pour off strip equilibration buffer I, add strip equilibration buffer II, and shake it slowly on the horizontal shaking table for 15 min;

4. When the equilibration is complete, pour off strip equilibration buffer II and absorb excess equilibrium liquid with a piece of filter paper;

5. Remove the IPG strip from the sample hydration tray and use tweezers to move one end of the strip and immerse its gel surface completely in the 1 × electrophoresis buffer. Next, place the strip on a long glass plate of gel with its gel surface upward to begin a second-dimension electrophoresis.

5.1.6 Second-Dimension SDS–PAGE Electrophoresis

1. Prepare acrylamide gel at a specific concentration. Inject the solution into each glass plate layer, leaving a space of 1 cm at the top. Use MilliQ water, ethanol, or water-saturated *n*-butyl alcohol to seal the gel surface and maintain the smoothness;

2. After the gel solidifies, pour off the MilliQ water, ethanol, or water-saturated *n*-butyl alcohol on the separation gel surface and flush the gel surface with MilliQ water;

3. Use filter paper to absorb the excess liquid between the glass plates on the SDS–PAGE gel. Place the treated second-dimension gel on the table with the long glass plate down, the short glass plate up, and the gel top facing the experimenter;

4. When the second equilibration is complete, remove the IPG strip from the equilibration tray and use tweezers to move one end of the strip and immerse its gel surface completely in the 1 × electrophoresis buffer, then place the strip on the long glass plate of gel with its gel surface facing upward;

5. Transfer the SDS–PAGE gel with the strip to a casting apparatus and add low melting-point agarose blocking liquid above the gel to avoid bubbles;

6. Push the strip gently with tweezers or a tongue depressor to ensure it is in full contact with the polyacrylamide gel surface;

7. Hold for 5–20 min to thoroughly solidify the low melting-point agarose blocking liquid, and then transfer the gel into the electrophoresis tank;

8. Turn on the power after adding the electrophoresis buffer into the electrophoresis tank. Conduct the electrophoresis with a low current (2–5 mA/gel) and a constant power (1–5 W/gel) for the first 45 min, and then increase the current or power. Stop the electrophoresis when the bromophenol blue indicator reaches the bottom edge;

9. When the electrophoresis ends, gently pry open the two-layer glass and take the gel out for dyeing. Conduct expression analysis, MS analysis, and identification on the target protein.

5.1.7 2-DE Technique to Analyze Protein Profile in Somatic Embryos of Liriodendron Hybrids

As one of the oldest methods, 2-DE is still widely used and compelling that it is less demanding from technical standpoint (Millioni *et al*., 2012). Currently, 2-DE-based differential proteomics approach has still been widely utilized to eliminate the molecular mechanisms during plant growth and development as well as in plant defense responses to stress conditions (Yang *et al*., 2011a,b, 2013; Bai *et al*., 2011; Chen *et al*., 2012a,b).

In order to understand the molecular mechanism at protein level, proteomic analysis about the development of somatic embryos of *Liriodendron* hybrids was conducted. Protein extraction of somatic embryos of *Liriodendron* hybrids was performed according to the method described by Chen *et al*. (2012a, b) IEF was carried out with 24-cm Immobiline DryStrip IPG strips (pH range 4.0–7.0; Amersham Biosciences, USA). The strips were first hydrated for 16 h in 450 μL of rehydration buffer [7 M urea, 2 M thiourea, 4% (w/v) CHAPS, 1% (v/v) IPG buffer, 50 mM DTT] that contained 400 μg of protein and a trace of bromophenol blue. Strips were then subjected to IEF at 20 °C for 64 kVh in an IPG-phor II system (Amersham Biosciences). After IEF, the strips were equilibrated for 15 min in 6 M urea, 0.375 M Tris, pH = 8.8, 2% (w/v) sodium dodecyl sulfate, 20% (v/v) glycerol, 2% (w/v) DTT, and 2.5% (w/v) IAM. For the second dimension, SDS–PAGE was performed with 12.5% (w/v) polyacrylamide gels at 12 W/gel for 4.5 h. The gel was stained with silver staining. After staining and destaining, the gel was scanned using ImageScanner II and analyzed using Image Master 2-D Version 5.0 software (Amersham Biosciences). The average number of electrophoretically separated, silver-stained spots from somatic embryos of *Liriodendron* hybrids was ~1150 (Figure 3).
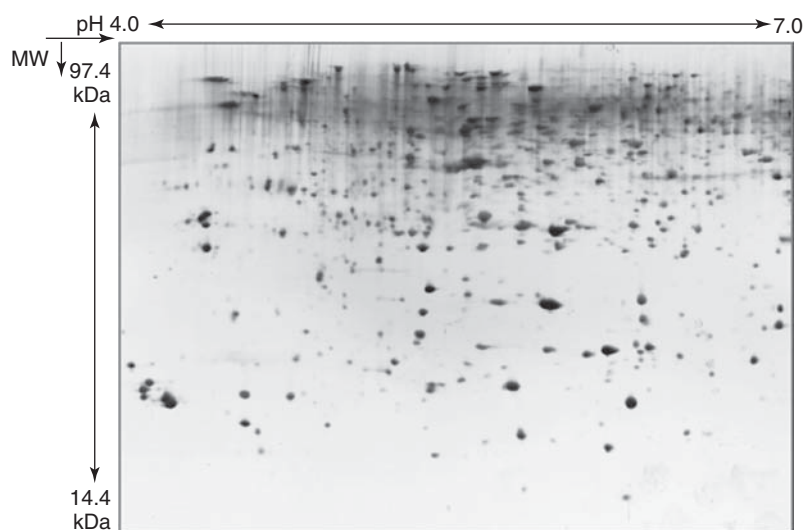
Figure 3   2-DE profiles of somatic embryos of *Liriodendron* hybrids. Total protein was extracted from somatic embryos of *Liriodendron* hybrids, separated by IEF/SDS–PAGE and silver-stained. The same amount of protein (400 mg) was loaded, and about 1150 protein spots are detected in 2-DE profile.

5.2   iTRAQ Procedures and Their Application

The iTRAQ workflow is summarized as follows: quantitating proteins, denaturing and reducing the sample, blocking cysteines, digesting proteins with trypsin, labeling digested proteins with iTRA reagents, combining and vortex-mixing the contents of each iTRAQ reagent-labeled sample in one tube, and finally spinning for LC/MS/MS or MALDI TOF/TOF MS analysis (Figure 4).

5.2.1   iTRAQ Procedures

1. Select different methods to extract the proteins;
2. Protein quantitation;
3. Add each group of samples to pre-cooled acetone and precipitate at $-20\,°C$ for 30 min to 4 h; centrifuge at 12,000 rpm for 15 min, and remove the supernatant;
4. Protein denaturation and reduction;
5. Fully dissolve the samples by vortex spinning in 20 μL dissolution buffer. Add 1 μL denaturant to denature the samples by vortex-mixing and add 2 μL reducing reagent. Reduce the samples by vortex-mixing. Centrifuge the samples and incubate them in a $60\,°C$ water bath for 1 h.

6. Cysteine blocking: Add the reduced samples to 1 μL cysteine blocking reagent, mix them by vortex spinning, centrifuge, and incubate at room temperature for 10 min;
7. *Enzyme digestion*: Add pancreatic enzyme at a ratio of 50:1 (sample:pancreatic enzyme). Dissolve the pancreatic enzyme with ultrapure water and incubate at $37\,°C$ for 12 h or overnight;
8. *iTRAQ labeling*: Centrifuge a tube of labeling reagent to gather the reagent at the bottom of the tube, add 70 μL ethanol or isopropanol, and uniformly mix them. Next, add the solution obtained in the previous step into each tube of sample, react at room temperature for 1 h, and add triple the volume in water to dissolve the labeling regent;
9. After labeling, mix each labeled sample separately; the sample separation process is a two-dimensional liquid phase separation process (first strong cation exchange (SCX) separation and subsequently RP separation);
10. *Fraction collection*: Start collecting from the third minute and collect a tube of fraction every 2 min, in which some of the fractions based on their chromatograms may be combined. It is standard to prepare 10–15 fractions for RP separation. Freeze-dry each SCX fraction, dissolve
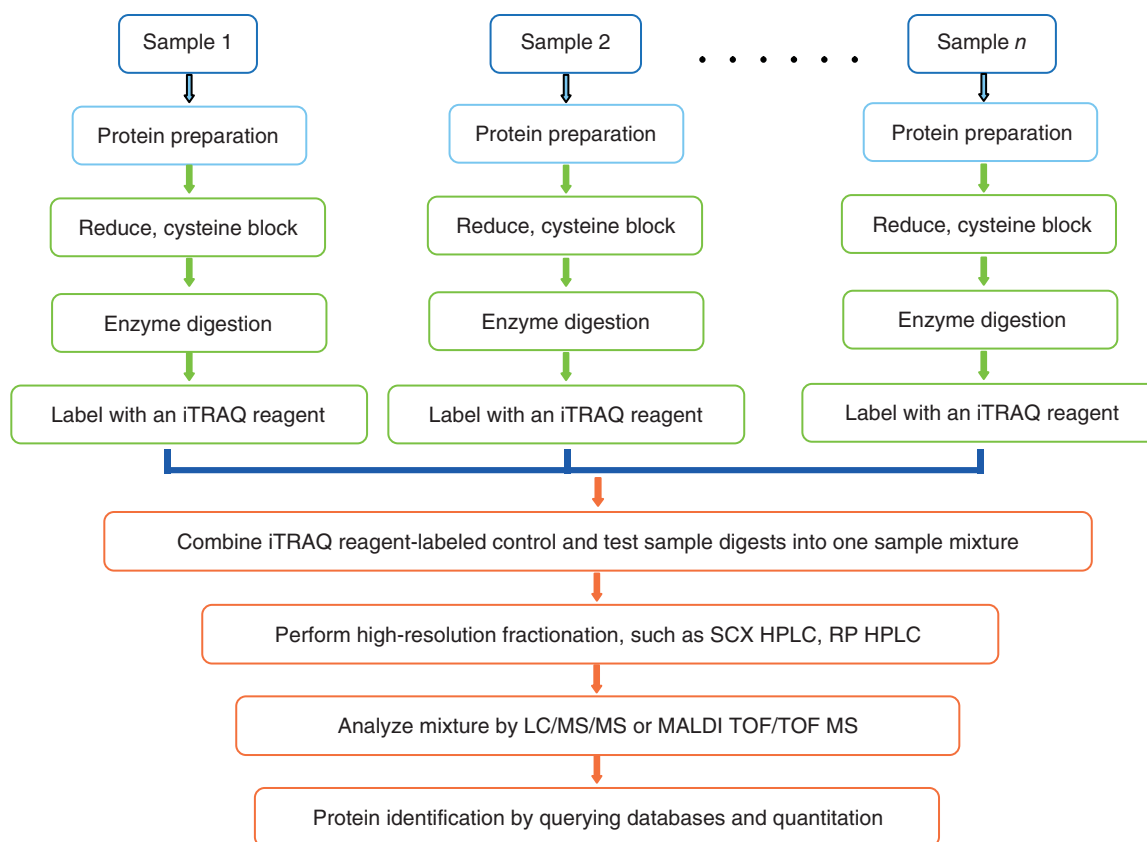
Figure 4 Schematic workflow of iTRAQ Technique. Protein samples are extracted from different plant materials. Each sample is reduced, cysteine blocked, digested, and labeled in a single tube. All iTRAQ reagent-labeled samples are subsequently combined into one sample mixture for LC/MS/MS analysis. Proteins are identified by querying protein databases and quantitated. Combine iTRAQ reagent-labeled control and test sample digests into one sample mixture. Perform high-resolution fractionation, such as SCX HPLC and (RP) reversed-phase HPLC HPLC. Protein identification by querying databases and quantitation. Analyze mixture by LC/MS/MS or MALDI TOF/TOF MS.

with RP A phase, and load after centrifuging at 10,000 rpm for 15 min;

11. Conduct MS analysis, database search, and analysis.

5.2.2 iTRAQ Technique to Analyze Differentially Expressed Proteins Mediated by OsICE1 Gene in Rice Leaves under Cold Stress

iTRAQ are widely used in proteomics to study quantitative changes during different developmental stages of plants or responding to diverse environmental factor (Han *et al*., 2010; Yang *et al*., 2011a,b; Melo-Braga *et al*., 2012; Nguyen *et al*., 2012; Lim *et al*., 2013). It is considered to be more accurate in quantitating protein abundances. In our laboratory, iTRAQ technique is used to analyze differentially expressed proteins mediated by *OsICE1* (Inducer of CBF Expression1) gene in rice leaves under cold stress. Over-expressing and antisense *OsICE1* transgenic rice lines were obtained by transgenic technology. Four-leaf-stage seedlings of *OsICE1* transgenic rice lines and the control were grown in Hogland's nutrient solution in a growth chamber where leaf samples were harvested at four points: 0 (control), 6, 12, and 24 h after 4 °C chilling stress treatment for protein extraction.

The leaves of different rice genotypes were collected and smashed in a mortar-grinder with liquid nitrogen. The powder was added to lysis buffer (2 *M* thiourea, 7 *M* urea, and 4% CHAPS, pH = 8.5) and mixed thoroughly by vortex spinning. The mixture
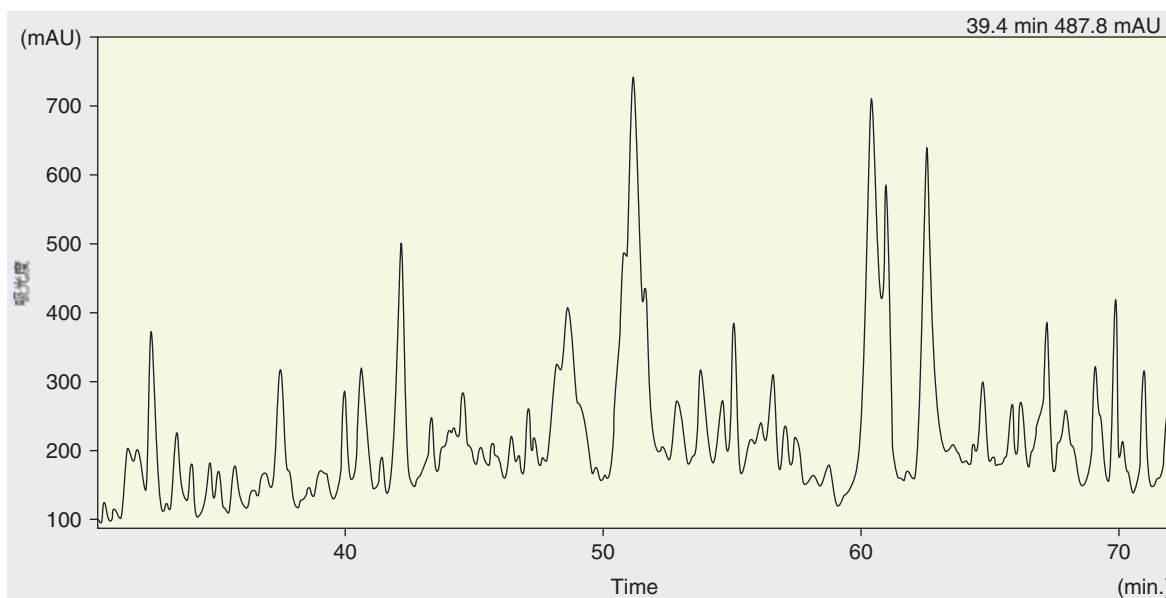
Figure 5 Profile of cation exchange chromatography of eight-plex iTRAQ-labeled peptides. Profile of cation exchange chromatography of eight-plex iTRAQ-labeled peptides derived from untreated rice seedlings (iTRAQ tags 113, 114, 115, or 116) and cold-treated rice seedlings at different time course (iTRAQ tags 117, 118, 119, or 121). Sixteen fractions were collected for downstream reversed-phase chromatographic separation.

was then centrifuged at 40,000g for 30 min and the pellet was discarded. The supernatant was added to acetone containing 10% trichloroacetic acid and precipitated proteins at $-20\,^{\circ}$C. The precipitated proteins were collected by centrifugation at 40,000g for 30 min and washed with acetone for two or three times until there is no color on the pellet. After being dried in vacuum, the protein pellet was dissolved in 8 M urea supplemented with 10 mM DTT, pH = 8.5 and the concentration was determined by the Bradford assay.

The proteins were dissolved, denatured, alkylated, and digested with trypsin (Promega) at 37 $^{\circ}$C for 18 h. To label peptides with iTRAQ reagent, one unit of label (defined as the amount of reagent required to label 100 g of protein) was thawed and reconstituted in 70 µL of ethanol. In addition, a SCX column was used to separate the mixed peptides. The labeled peptides were diluted in buffer A and injected at a flow rate of 0.7 mL min^{-1} into a high resolution of SCX column (4.6 × 250 mm^2 5 µm; Thermo BioBasic, USA). The SCX column and C$_{18}$ pre-column were flushed with a three-step gradient sodium chloride solution (0, 50, and 100 mM) for 70 min after loading and the pH of the diluted

sample was checked to be between 2.5 and 3.3 (Figure 5). The diluted sample mixture was loaded into the cation exchange cartridge, and the peptides were eluted with 500 µL of buffer-elute (10 mM K$_2$HPO$_4$ in 25% (v/v) acetonitrile, 350 mM KCl at pH = 3.0) after washing with 10 column volumes of buffer-load. The eluate of the cation exchange was desalted with an Agilent 1100 series HPLC system equipped with an autosampler, 2/6 valve, and a diode array detector (220 nm) (Agilent, Waldbronn, Germany) afterward and 30 fractions were collected. Before LC–MS/MS, each fraction was desalted by SP-10 pre-column.

The fractionated peptides were analyzed by the Triple Quadrupole TOF® 5600 System (AB SCIEX, Concord, ON) fitted with a Nanospray® III source (AB SCIEX, Concord, ON) and a pulled quartz tip as the emitter (New Objectives, Woburn, MA). ProteinPilot™ software 4.0 (AB SCIEX, Foster City, CA) was used to interpret raw data files produced by MS. Protein quantitation was also performed by the ProteinPilot™ software, which automatically calculated the relative abundance of iTRAQ-labeled peptides and the corresponding proteins.
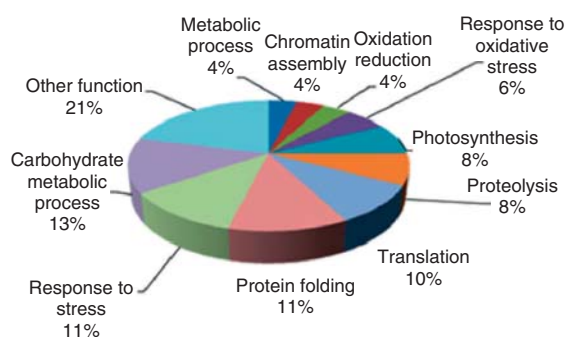
Figure 6 Functional classification of cold-responsive proteins mediated by rice *OsICE1* gene. The pie chart shows the distribution of non-redundant proteins according to their functional classes in percentages.

With high throughput MS analysis, 3258 proteins were identified from rice seedlings after searching in protein database, and 72 differentially expressed proteins (up-regulated or down-regulated profile) with >1.5-fold variations during low temperature treatment. Functional distribution of cold-responsive proteins identified above by Gene Ontology revealed that most of the differentially expressed proteins were involved in metabolic processes, redox, photosynthesis, chromatin assembly, oxidative stress, proteolysis, translation, protein folding, stress response, and carbohydrate metabolism (Figure 6).

6 CONCLUSION

Plant proteomics research has made great progress in protein separation and identification as well as interaction between proteins after its development in recent years, and it is widely applied in every field in life science research.

The challenges that proteomics research and application of medicinal plant are facing fall into three categories: (i) complete protein identification, including protein isomers, protein modifications, and localizations; (ii) functional studies of protein biochemistry and cytology; and (iii) analyses of protein regulations and their relationships with regulatory networks (Bertone and Snyder, 2005). Just as in other fields of life science, important breakthroughs in proteomics research of medicinal plant depend on technological innovations.

REFERENCES

Aebersold, R. and Mann, M. (2003) *Nature*, **422**, 198–207.

Bai, X., Yang, L., Yang, Y., *et al*. (2011) *J. Proteome Res.*, **10**, 4349–4364.

Bertone, P. and Snyder, M. (2005) *Plant Physiol.*, **138**, 560–562.

Chen, J., Cheng, T., Wang, P., *et al*. (2012a) *J. Proteomics*, **75**, 5226–5243.

Chen, J., Tian, L., Xu, H., *et al*. (2012b) *Plant Omics*, **5**, 194–199.

Cordwell, S. J., Basseal, D. J., Bjellqvist, B., *et al*. (1997) *Electrophoresis*, **18**, 1393–1398.

Debnath, M., Pandey, M. and Bisen, P. S. (2011) *OMICS*, **15**, 739–762.

Decker, G., Wanner, G., Zenk, M. H., *et al*. (2000) *Electrophoresis*, **21**, 3500–3516.

Domon, B. and Aebersold, R. (2010) *Nat. Biotechnol.*, **28**, 710–721.

Espejo, A., Cote, J., Bednarek, A., *et al*. (2002) *Biochem. J.*, **367**, 697–702.

Fields, S. (2001) *Science*, **291**, 1221–1224.

Görg, A., Weiss, W. and Dunn, M. J. (2004) *Proteomics*, **4**, 3665–3685.

Gygi, S. P., Rist, B., Gerber, S. A., *et al*. (1999) *Nat. Biotechnol.*, **17**, 994–999.

Gygi, S. P., Corthals, G. L., Zhang, Y., *et al*. (2000) *Proc. Natl. Acad. Sci. U. S. A.*, **97**, 9390–9395.

Hall, D. A., Zhu, H., Zhu, X., *et al*. (2004) *Science*, **306**, 482–484.

Han, B., Chen, S., Dai, S., *et al*. (2010) *J. Integr. Plant Biol.*, **52**, 1043–1058.

Hew, C. S. and Gam, L. H. (2011) *Appl. Biochem. Biotechnol.*, **165**, 1577–1586.

Jacobs, D. I., Rijssen, M. S. van, Heijden, R. van der, *et al*. (2001) *Proteomics*, **1**, 1345–1350.

Jacobs, D. I., Gaspari, M., Greef, J. van der, *et al*. (2005) *Planta*, **221**, 690–704.

Kersten, B., Feilner, T., Kramer, A., *et al*. (2003) *Plant Mol. Biol.*, **52**, 999–1010.

Kersten, B., Possling, A., Blaesing, F., *et al*. (2004) *Anal. Biochem.*, **331**, 303–313.

Klose, J. (1975) *Hum. Genet.*, **26**, 231–243.

Klose, J. (2009) *Electrophoresis*, **30**, S142–S149.

Koay, S. Y. and Gam, L. H. (2011) *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.*, **879**, 2179–2183.

Kofler, M., Motzny, K. and Freund, C. (2005) *Mol. Cell. Proteomics*, **4**, 1797–1811.

Koller, A., Washburn, M. P., Markus-Lange, B., *et al.* (2002) *Proc. Natl. Acad. Sci. U. S. A.*, **99**, 11969–11974.

Kopf, E., Shnitzer, D. and Zharhary, D. (2005) *Proteomics*, **5**, 2412–2416.

Lim, S., Borza, T., Peters, R. D., *et al.* (2013) *J Proteomics*, DOI: 10.1016/j.jprot.2013.03.010.

Lopez-Casado, G., Covey, P. A., Bedinger, P. A., *et al.* (2012) *Proteomics*, **12**, 761–774.

Mackey, A. J., Haystead, T. A. and Pearson, W. R. (2002) *Mol. Cell. Proteomics*, **1**, 139–147.

Melo-Braga, M. N., Verano-Braga, T., León, I. R., *et al.* (2012) *Mol. Cell. Proteomics*, **11**, 945–956.

Michaud, G. A., Salcius, M., Zhou, F., *et al.* (2003) *Nat. Biotechnol.*, **21**, 1509–1512.

Millioni, R., Puricelli, L., Sbrignadello, S., *et al.* (2012) *Amino Acids*, **42**, 1583–1590.

Nguyen, T. H., Brechenmacher, L., Aldrich, J. T., *et al.* (2012) *Mol. Cell. Proteomics*, **11**, 1140–1155.

Ning, K., Fermin, D. and Nesvizhskii, A. I. (2012) *J. Proteome Res.*, **11**, 2261–2271.

O'Farrell, P. H. (1975) *J. Biol. Chem.*, **250**, 4007–4021.

Oldham, J. T., Hincapie, M., Rejtar, T., *et al.* (2010) *J. Proteome Res.*, **9**, 4337–4345.

Patton, W. F. and Beechem, J. M. (2001) *Curr. Opin. Chem. Biol.*, **6**, 63–69.

Ross, P. L., Huang, Y. N., Marchese, J. N., *et al.* (2004) *Mol. Cell. Proteomics*, **3**, 1154–1169.

Sabidó, E., Selevsek, N. and Aebersold, R. (2012) *Curr. Opin. Biotechnol.*, **23**, 591–597.

Steinberg, T. H., Agnew, B. J., Gee, K. R., *et al.* (2003) *Proteomics*, **3**, 1128–1144.

Templin, M. F., Stoll, D., Schwenk, J. M., *et al.* (2003) *Proteomics*, **3**, 2155–2166.

Tonge, R., Shaw, J., Middleton, B., *et al.* (2001) *Proteomics*, **1**, 377–396.

Uetz, P., Giot, L., Cagney, G., *et al.* (2000) *Nature*, **403**, 623–627.

Unlu, M., Morgan, M. E. and Minden, J. S. (1997) *Electrophoresis*, **18**, 2071–2077.

Washburn, M. P., Wolters, D. A. and Yates, J. R. III, (2001) *Nat. Biotechnol.*, **19**, 242–247.

V. C. Wasinger, S. J. Cordwell, A. Cerpa-Poljak, *et al.* (1995) *Electrophoresis*, **16**, 1090–1094.

Wu, X., Xiong, E., An, S., *et al.* (2012) *PLoS One*, **7**, e50497.

Yang, L., Tian, D., Luo, Y., *et al.* (2011a) *Afr. J. Agric. Res.*, **6**, 808–816.

Yang, Y., Qiang, X., Owsiany, K., *et al.* (2011b) *J. Proteome Res.*, **10**, 4647–4660.

Yang, L., Tian, D., Todd, C. D., *et al.* (2013) *J. Proteome Res.*, **12**, 1316–1330.

Zeng, J., Liu, Y., Liu, W., *et al.* (2013) *PLoS One*, **8**, e53409.

Zhou, H., Liu, Y., Chui, J., *et al.* (2007) *Arch. Biochem. Biophys.*, **459**, 70–78.

Zhu, H., Bilgin, M., Bangham, R., *et al.* (2001) *Science*, **293**, 2101–2105.

Multivariate Data Analysis

Bieke Dejaegher

Department of Analytical Chemistry and Pharmaceutical Technology (FABI), Vrije Universiteit Brussel (VUB), Brussels, Belgium

1 INTRODUCTION

Herbal plants and their extracts have been used in traditional medicine (TM) for thousands of years in order to prevent or treat diseases. While initially the use of herbal plants and their extracts was mainly practiced in Asian and African countries, nowadays, Western countries are also interested in their use.

A decent quality control of herbal plants, their extracts, or their derived medicines is of the utmost importance to have an effective prevention or treatment of the disease without the occurrence of side effects or, in the worst case, toxic effects. However, because of the complexity of this type of samples, such quality control is difficult and reliable methods are needed.

In the past, *pharmacopeial monographs* (Chinese Pharmacopoeia, 2010; European Pharmacopoeia, 2010; United States Pharmacopoeia, 2010) were published, where the identification and quality control of herbal material is performed using *microscopic and macroscopic techniques* and by the evaluation of the so-called markers. However, one might wonder whether evaluation of only one or several markers is enough to describe the complex samples. Indication of the presence of the markers does not imply that no other components are present that cause side effects or, in the worst case, toxic effects. Moreover, often the therapeutic outcome is caused by several components with synergic effects.

Therefore, since the beginning of this century, several regulatory bodies suggest and accept (WHO, 2000; FDA Guidelines, 2004; EMA Guidelines, 2006) or even demand (Chinese SFDA Guidelines, 2002) the use of the so-called fingerprint analysis for the quality control of herbal plants, their extracts, or their derived medicines. A fingerprint is a complete and characteristic profile or pattern that represents the composition of a complex sample. In the fingerprint analysis, *three steps* can be distinguished, that is, the fingerprint development, validation, and data handling or extraction of the desired information from the fingerprint.

Fingerprints can be *developed* using *spectroscopic*, *mass spectrometric* (MS) (Lee, Kim, and Jang, 2012; Zeng *et al.*, 2013), *chromatographic*, or *electrodriven techniques*, such as capillary electrophoresis (CE) (Sturm, Seger, and Stuppner, 2007) or pressurized *c*apillary *e*lectro*c*hromatography (pCEC) (Pieters *et al.*, 2011). Spectroscopic fingerprints can, for example, be developed with ultraviolet–visible (UV–vis) (Ni *et al.*, 2009; Palacios-Morillo *et al.*, 2013), infrared (IR) (Mao and Xu, 2006), near infrared (NIR) (Mao and Xu, 2006; Lucio-Gutiérrez, Coello, and Maspoch, 2011), or Raman (Mao and Xu, 2006) spectroscopy or with nuclear magnetic resonance (NMR) (Frédérich *et al.*, 2011). Chromatographic fingerprints can, for instance, be obtained with thin-layer chromatography (TLC), high performance thin-layer chromatography (HPTLC) (Tian, Xie, and Liu, 2009),

high performance liquid chromatography (HPLC) (van Nederkassel *et al*., 2005; Daszykowski, Vander Heyden, and Walczak, 2007; Grata *et al*., 2007; Dumarey *et al*., 2008; Møller, Hansen, and Cornett, 2009; Tian, Xie, and Liu, 2009; Tistaert *et al*., 2009), ultrahigh performance liquid chromatography (UHPLC) (Grata *et al*., 2007; Kong *et al*., 2009), or gas chromatography (GC) (Jumtee, Bamba, and Fukusaki, 2009; Parastar *et al*., 2012). For chromatographic and electrodriven techniques, the applied detection techniques are UV absorption (van Nederkassel *et al*., 2005; Daszykowski, Vander Heyden, and Walczak, 2007; Dumarey *et al*., 2008; Tian, Xie, and Liu, 2009; Pieters *et al*., 2011), diode array detection (DAD) (Kong *et al*., 2009; Møller, Hansen, and Cornett, 2009; Tistaert *et al*., 2009), fluorescence detection (Møller, Hansen, and Cornett, 2009), MS (Sturm, Seger, and Stuppner, 2007; Grata *et al*., 2007; Møller, Hansen, and Cornett, 2009; Jumtee, Bamba, and Fukusaki, 2009; Parastar *et al*., 2012), evaporative light scattering detection (ELSD) (Tian, Xie, and Liu, 2009), electrochemical detection (ED) (Møller, Hansen, and Cornett, 2009), or flame ionization detection (FID) (Jumtee, Bamba, and Fukusaki, 2009).

Spectroscopic and MS fingerprints will display a profile of the whole sample at one time point. In addition, spectroscopic fingerprints will not allow obtaining information from individual components. Fingerprints obtained with chromatographic or electrodriven separation techniques, on the other hand, will result in a characteristic profile or pattern in which as many components as possible are separated and will, therefore, allow revealing information about individual compounds.

The sample preparation and fingerprint development of all the above-mentioned case studies are considered outside the scope of this chapter and are, therefore, not discussed.

Each fingerprint represents a signal, that is, a response or an output from an analytical instrument, as a function of many variables (*n*). After the development of the fingerprinting method, fingerprints are recorded for all samples (*m*). Then, the fingerprinting method should be validated. Finally, the desired information should be extracted from the fingerprints.

To *validate the entire fingerprint*, the classically described method validation (Eurachem, 1998; ICH, 2005) cannot be used as such, because those guidelines are applicable for univariate methods, where one signal is related to the concentration of one component.

For chromatographic and electrodriven fingerprints, individual components can be separated and for all peaks that are pure, the univariate classic method validation approach can be applied, usually using standard compounds. This was, for example, done by Kong *et al*. (2009) for five alkaloids present in Rhizoma Coptidis. The International Conference of Harmonisation guidelines (2005) were followed for the validation of the ultra performance liquid chromatography-diode array detector (UPLC-DAD) method. Evaluated parameters were the linearity, the limit of detection and quantification, the repeatability, the time-different intermediate precision, and the accuracy.

When peak purity cannot be evaluated and/or when no standard compounds are available, alternative approaches to validation of fingerprints can be performed. For example, in the study by Tian, Xie, and Liu (2009), the guidelines of the *Chinese Pharmacopoeia* (Chinese Pharmacopoeia Commission, 2010) were followed to validate the HPLC-ELSD fingerprints of *Bupleurum* species. The %relative standard deviation (%RSD) of the retention time and the peak area of a given saikosaponin was calculated from five replicate measurements. The sample stability was evaluated using the %RSD of the peak area of the given saikosaponin at different time points. Alaerts *et al*. (2007) used a setup to simultaneously determine the repeatability and the time-different intermediate precision of the responses retention time, peak area, and peak height of four peaks in HPLC-UV fingerprints of *Liquorice*, *Cascara*, *Curcuma*, and *Artichoke* samples. For this purpose, three replicates of each sample were measured for 7 days in order to calculate the repeatability and the time-different intermediate precision.

Validation of spectroscopic and MS fingerprints, on the other hand, is not so straightforward, as a profile of the whole sample is given at one time point. Zeng *et al*. (2013) preliminarily validated their direct analysis in real time time of flight mass spectrometric (DART-TOF-MS) fingerprints by investigating the repeatability of the absolute and relative peak areas of three replicate measurements of the same *Danshen* alkaline precipitation sample. The %RSD was lower when using the relative peak areas. Therefore, for further data analysis, the relative peak areas were used.

The Pharmaceutical Analytical Sciences Group (PASG), which is an association of analytical

chemists working in the research-based pharmaceutical industry in the United Kingdom, has defined guidelines to validate NIR fingerprints (Broad *et al.*, 2002). The considered parameters for method validation are the same as those from the International Conference of Harmonisation (2005), but they were adapted to validate NIR methods. The required validation parameters of the ICH guidelines and their interpretation for NIR methods are given in Table 1.

This chapter primarily focuses on the last step in fingerprint development. To *extract the desired information from the fingerprints*, an $m \times n$ data matrix **X** of all fingerprints is, therefore, constructed, with m the number of samples and n the number of variables. The information that is desired depends on the goal of the study. Regardless of the goal, chemometric techniques are usually mandatory to extract the information from the multivariate data. Such techniques are divided into *unsupervised and supervised methods*. Before all techniques, the *data* is often *pretreated* (Section 2) to ensure adequate data analysis.

Unsupervised methods only extract the information contained in the $m \times n$ data matrix **X** (Figure 1a, Section 3), whereas supervised techniques try to relate the information contained in the data matrix **X** with independent variables to an $m \times 1$ response vector **y** with one dependent measure (Figure 1b, Section 4) (Alaerts *et al.*, 2010; Tistaert, Dejaegher, and Vander Heyden, 2011). When dealing with more than one dependent measure, multivariate data analysis techniques that allow evaluating the relation between two multivariate data sets, such as an $m \times p$ data matrix with independent variables and an $m \times q$ data matrix **Y** with multiple dependent measures, should be used (Vandeginste *et al.*, 1998; Brown, Tauler, and Walczak, 2009a).

2 DATA PRETREATMENT

Before any chemometric technique, the data is usually pretreated. In general, the goal of *preprocessing the data* is to remove unwanted variability while maintaining the relevant variation. Depending on the type of fingerprint, different pretreatment techniques can be applied (Brown, Tauler, and Walczak, 2009a; Alaerts *et al.*, 2010; Tistaert, Dejaegher, and Vander Heyden, 2011). In general, these techniques can be divided into unsupervised and supervised methods.

The former methods only require the data matrix **X** for the data pretreatment, whereas the latter methods also require a response vector **y**. Most preprocessing methods, however, are unsupervised in nature (Brown, Tauler, and Walczak, 2009a).

For *spectroscopic fingerprints* (UV–vis, IR, NIR, NMR, etc.), the often-applied preprocessing methods are, for example, filtering methods, baseline correction (BC), first- and second-order derivatives (FSDs), wavelet transform (WT), column centering (CC), normalization, autoscaling, Pareto-scaling, standard normal variate (SNV), multiplicative signal correction (MSC), and direct orthogonal signal correction (DOSC).

Filtering methods are used to filter out the noise from the data in order to reveal only important signals or information. They can be divided into nonrecursive and recursive filtering approaches. Examples of the former approaches are moving-average filtering, boxcar filtering, median filtering, polynomial or Savitzky–Golay filtering, asymmetric filtering, and matched filtering. An example of the latter approach is Kalman filtering. *BC* methods will remove from each row its corresponding baseline. *FSDs* transform the data such that smaller differences between samples are more emphasized. Derivatives often smooth the signal, thereby denoising the data, improving the peak resolution, and correcting for a baseline or trend. *WT* transforms the data, with the goal of denoising and compressing the data. Unimportant regions of the data will be deleted while important regions will be selected. *CC* removes the column mean from each corresponding column element and is very often applied before further data analysis.

Normalization consists of row scaling, that is, dividing each row element by the corresponding row standard deviation. *Autoscaling* performs column centering, followed by column scaling, that is, dividing each column element by its corresponding column standard deviation. *Pareto-scaling* consists of column centering followed by a division of each column element by its corresponding column square root of standard deviation. *SNV* transforms the data by performing row centering, followed by row scaling, that is, dividing each row by its corresponding standard deviation. *MSC* is a response correction method that removes baseline shift by subtracting from each row the average baseline of different fingerprints. In contrast to the above-mentioned methods, *DOSC* is a supervised preprocessing method, in which from the

Table 1 Required validation parameters of the ICH guidelines and their interpretation for NIR methods.

ICH Q2A validation parameter	NIR validation parameter and interpretation	Type of NIR procedure			
		Identification	Qualification	Quantification	
				Major components	Impurities/minor components
Specificity[a]	As ICH Q2A and Q2B	+	+	+	+
Linearity[b]	With NIR spectra, the data are typically multidimensional as opposed to one-dimensional data seen with conventional analytical methodology. Therefore, with NIR, the equivalent of linearity is the mapping of a calibration surface/volume as opposed to the construction of a single univariate calibration line. The validation for NIR, therefore, involves the demonstration of correlated NIR response to samples distributed throughout the defined range of the calibration model	−	−	+	+
Range[b]	Defined by coverage of the product, process, and material variability, which needs to be accommodated in the NIR method	−	−	+	+
Accuracy	As ICH Q2A and Q2B. Usually demonstrated for NIR methods by correlation of NIR results with analytical reference data. NIR is often constrained (e.g., particularly for intact solid dosage forms) by the nonfeasibility of performing recovery experiments	−	−	+	+
Precision					
Repeatability	As ICH Q2A and Q2B	−[c]	+[c]	+	+
Intermediate precision	As ICH Q2A and Q2B, encompassing different analysts and different days but not necessarily instruments	−	−	+[d]	+[d]
Robustness[e]	Robustness is inherently built into an NIR method during development by correct and appropriate sample selection/presentation (see technical guidelines), but still needs to be demonstrated in a similar way to conventional methods as described in ICH guidelines Q2A and Q2B	+	+	+	+
Detection limit	As Q2A and Q2B	−	−	−	−
Quantification limit	As Q2A and Q2B, but constrained by lowest level available in sample calibration set	−	−	−	+

− signifies that this characteristic is not normally evaluated.

+ signifies that this characteristic is normally evaluated.

[a] Lack of specificity of the NIR procedure could be compensated by other supporting analytical procedures.

[b] Both linearity and range of an NIR method will be dependent on the availability of samples representing product and process variations, in contrast to the fixed range (e.g., 80−120%), applied in validation of conventional methodology.

[c] Not normally required for identification methods. For qualification methods, repeatability is addressed in order to demonstrate that the acceptance thresholds established provide reliable discrimination between acceptable and unacceptable materials; the approach is, therefore, conceptually different for NIR methods compared with conventional methods.

[d] In cases where reproducibility has been performed, intermediate precision is not needed.

[e] Robustness is not listed in this table in ICH Q2A; for conventional method validation, robustness is frequently assessed after the method has been developed, and may not be built in during method development.

Source: N. Broad *et al*. (2002). Reproduced from John Wiley & Sons, Ltd.
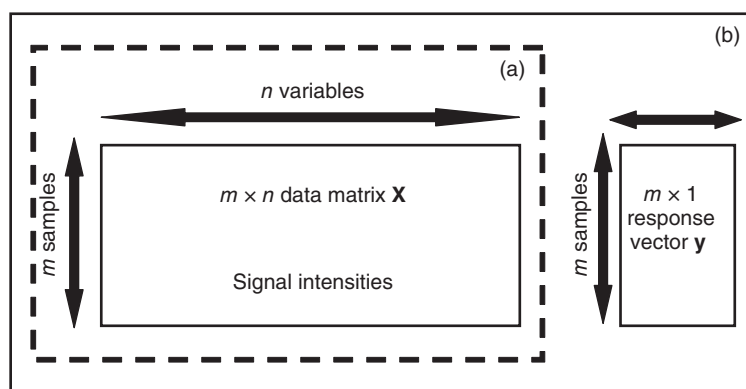
Figure 1 (a) Unsupervised and (b) supervised data analysis.

data matrix **X**, the variation orthogonal to, that is, not related to, the response **y** is removed.

When comparing NMR, IR, NIR, Raman, and MS spectra of different samples, it is seen that sometimes *spectral shifts* take place. These shifts can be caused by differences in temperature, pressure, viscosity, and pH or by matrix effects or instrumental shortcomings. To overcome these spectral shifts, the spectral fingerprints should preferably be aligned because in order to obtain an adequate data analysis, the information about a given peak should for all samples be enclosed in the same column of the data matrix **X**. Alignment of spectral data can be done with different *alignment techniques*, such as dynamic time warping (DTW) (Kassidas, MacGregor, and Taylor, 1998; Pravdova, Walczak, and Massart, 2002), parametric time warping (PTW) (Eilers, 2004; van Nederkassel *et al*., 2006), semiparametric time warping (STW) (Eilers and Marx, 1996; van Nederkassel *et al*., 2006), and correlation optimized warping (COW) (Vest Nielsen, Carstensen, and Smedsgaard, 1998; Pravdova, Walczak, and Massart, 2002; van Nederkassel *et al*., 2006). Specifically for NMR and MS spectra, bucketing or peak-picking methods can also be used to align. Moreover, these latter methods reduce the number of variables in the data, which leads to less computation time needed for future data analysis. Detailed information about these different algorithms can be found in the literature (Brown, Tauler, and Walczak, 2009a).

For *chromatographic* (HPTLC, HPLC, UPLC, GC, etc.), pCEC, and CE fingerprints, the most important data pretreatment consists of applying *peak aligning techniques* because very often *retention factor* R_f or *time* t_R *shifts* can occur. These shifts can, for example, be caused by column aging or by small variations in mobile phase composition, flow rate, or temperature. Again, to ensure that the information about a given peak for all samples is found in the same column of the data matrix **X**, it is usually mandatory to align the corresponding peaks in the fingerprints. To align HPLC, UPLC, GC, pCEC, or CE fingerprints, different techniques can be used, for example, DTW, PTW, STW, COW, bucketing, and peak-picking methods. This alignment is shown in Figure 2. When MS detection is performed, usually bucketing or peak-picking methods are applied to align the corresponding peaks. To align HPTLC fingerprint images, an image-warp procedure can be used. As an example, an algorithm of an image-warp procedure, which was applied by Komsta *et al*. (2011), was described by Arganda-Carreras *et al*. (2006). Besides these aligning techniques, other preprocessing methods often applied for this type of fingerprints are BC, CC, normalization, SNV, and so on.

3 UNSUPERVISED DATA ANALYSIS

Unsupervised data analysis (Figure 1a) can be used for identification and quality control, for exploratory data analysis purposes, and in the context of curve resolution (Brown, Tauler, and Walczak, 2009a; Alaerts *et al*., 2010; Tistaert, Dejaegher, and Vander Heyden, 2011).

In the context of *identification and quality control*, very often, the correlation coefficient r $(-1 < r < 1)$
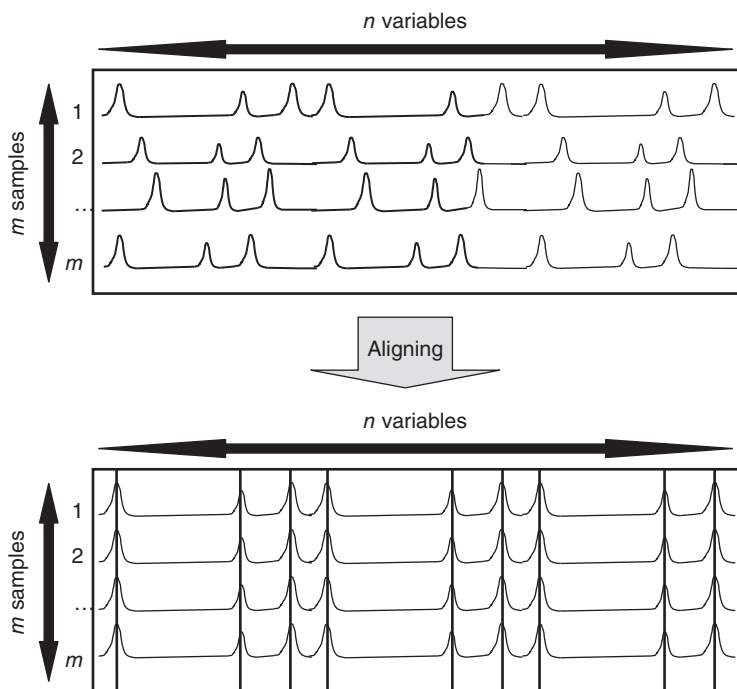
Figure 2 Alignment of corresponding peaks in HPLC, UPLC, GC, pCEC, and CE fingerprints.

or the congruence coefficient c $(0 < c < 1)$ between two fingerprints is calculated and evaluated. The correlation coefficient between two fingerprints is the scalar product of the normed mean-centered fingerprints, whereas the congruence coefficient is the cosine of the angle between two fingerprints or the scalar product of the normed fingerprints. In the computer-aided similarity evaluation (CASE) software (Gong *et al.*, 2005), c is called the *similarity index*. The closer r or c is to 1, the more similar the fingerprints are.

Exploratory data analysis methods can also be applied in the context of identification and quality control. They are also very often performed before a supervised data analysis. Principal component analysis (PCA), robust principal component analysis (rPCA), independent component analysis (ICP), projection pursuit (PP), or (hierarchical) cluster analysis ((H)CA) can be performed. Such methods allow visualization and insight in the multivariate data. They allow revealing a given structure or clustering tendency in the data, for example, according to a given activity, species, geographical origin, and harvest time.

Curve resolution methods resolve overlapping peaks, in order to obtain the individual spectra of the components, the individual concentration profiles of the components, and/or the number of components in the overlapping peaks. Applied methods are (iterative) target transformation factor analysis ((I)TTFA), evolving factor analysis (EFA), window factor analysis (WFA), heuristic evolving latent projection (HELP), orthogonal projection approach (OPA), multivariate curve resolution-alternating least squares (MCR-ALS), and extended multivariate curve resolution-alternating least squares (E-MCR-ALS).

4 SUPERVISED DATA ANALYSIS

As mentioned in Section 1, *supervised data analysis* tries to relate the information contained in the data matrix $\mathbf{X}$ to an $m \times 1$ response vector $\mathbf{y}$ (Figure 1b). When the response vector $\mathbf{y}$ is categorical, that is, contains classes, for example, high/low activity, different species, and different geographical origins, pattern recognition, discrimination or

classification models are built. When **y** is continuous, for example, an activity that is expressed on a continuous scale, multivariate calibration models are constructed (Brown, Tauler, and Walczak, 2009b; Alaerts *et al*., 2010; Tistaert, Dejaegher, and Vander Heyden, 2011).

Pattern recognition, discrimination, or classification methods will discriminate herbal samples into classes, according to activity, species, geographical origin, and so on. Discrimination or classification rules will be determined from a set of samples belonging to *a priori* known classes. Discrimination rules are then used to discriminate new samples into one of the possible classes, whereas classification rules are used to classify new samples into none, one, or several of the classes.

Pattern recognition methods include, for example, canonical discriminant analysis (CDA), linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), *K*-nearest neighbors (*k*-NN), classification and regression trees (CART), (probabilistic) partial least squares-discriminant analysis ((p)PLS-DA), orthogonal projections to latent structures-discriminant analysis (OPLS-DA), artificial neural networks (ANNs), soft independent modeling of class analogy (SIMCA), and support vector machines for classification (SVM-C). All these methods are discriminative in nature, except for SIMCA, which is a classification technique.

Multivariate calibration methods, on the other hand, model a property of interest, such as a quality rank or an antioxidant/cytotoxic/antimicrobial activity, from a set of samples with *a priori* known **y** values. Then, the model is used to predict the property of interest and/or to indicate components potentially responsible for the property of interest. Examples of multivariate calibration techniques are, for instance, stepwise multiple linear regression (Step-MLR), principal component regression (PCR), robust principal component regression (rPCR), partial least squares (PLS), robust partial least squares (rPLS), uninformative variable elimination partial least squares (UVE-PLS), orthogonal projection to latent structures (OPLS), and support vector machines for regression (SVM-R).

Supervised analysis techniques generally use a *training or calibration set*, of which the classes or the **y** values are *a priori* known, to build the discrimination/classification or calibration model, respectively. To build the model, first, the *optimal model complexity*, that is, the number of components or latent variables, should be determined. This is often done using a cross-validation (CV) procedure, such as a leave-one-out cross-validation (LOOCV), a leave-N-out cross-validation, or a Monte Carlo cross-validation (MCCV). Then, the root mean squared error of cross-validation (RMSECV) is determined for and plotted as a function of models with different complexity, that is, different numbers of components or latent variables. The model with the (nearly) lowest RMSECV is then selected as optimal. For pattern recognition techniques, besides evaluating the RMSECV, the % correct classification rate of the training set using cross-validation (%CCR_CV) can also be determined and plotted as a function of models with different complexity. Then, the model with the (nearly) highest %CCR_CV is considered to be optimal. Sometimes, the RMSECV or the %CCR for the training or calibration set without CV is also reported.

A discrimination/classification or calibration model can be evaluated using *three different criteria*, that is, its predictive ability, simplicity, and interpretability. Depending on the goal of the study, one, two, or all characteristics will be important and assessed.

To evaluate the *predictive ability* of a discrimination/classification or calibration model, preferably an independent test or prediction set is used, for which the classes or the **y** values, respectively, are also *a priori* known. Division of the data into a training/calibration and test/prediction set can be done using, for example, a uniform selection depending on the **y** values of the samples, the Duplex algorithm (Snee, 1977), or the Kennard and Stone algorithm (Kennard and Stone, 1969). As a rule of thumb, often two-thirds of the data is used for training/calibration, whereas the remaining one-third part will be used to evaluate the predictive ability of the model. The latter is done by calculating the root mean squared error of prediction (RMSEP). For pattern recognition methods, the % correct classification rate of the prediction set (%CCR_PRED) can also be calculated. However, when there are a limited number of samples, which is often the case for herbal samples, a division into a training/calibration and test/prediction set is not recommended. Then, often a CV procedure is used, where the RMSECV or the %CCR_CV will be used to evaluate the predictive ability of the model. Anyway, the model with the best predictive ability is that with the lowest RMSEP or RMSECV or the highest %CCR_CV or %CCR_PRED.

For discrimination methods, the % sensitivity for class i is related to the true positive rate and is the number of true positives (TP_i) for that class divided by the number of true positive and false negative (FN_i) predictions for that class. The % β-error for each class is equal to $100\% - \%$ sensitivity for each class and is linked to the presence of false negatives. The % specificity for class i is related to the true negative rate and is the number of true negative (TN_i) predictions for that class divided by the number of true negative and false positive (FP_i) predictions for that class. The % α-error for each class is equal to $100\% - \%$ specificity for each class and is linked to the occurrence of false positives. The % average and % overall specificity and sensitivity for each class i can consequently be calculated.

The second criterion is the model *simplicity*, which is related to the model complexity. The simplest model is that with the lowest number of components or latent variables. Finally, the *interpretability* of a model can also be important, especially when one wants to evaluate the contribution of the original variables to the final model. For this purpose, the regression coefficients are plotted as a function of the original variables to indicate which of the latter contribute most to the model.

Multivariate data analysis techniques that allow evaluating the relation between *two multivariate data sets*, such as, an $m \times p$ data matrix with independent variables and an $m \times q$ data matrix **Y** with multiple dependent measures, are, for example, procrustes analysis, canonical correlation analysis (CCA), multivariate least squares regression, reduced rank regression, multivariate PCR, multivariate PLS, multivariate regression trees, or multivariate consensus regression trees (Vandeginste *et al*., 1998; De'ath, 2002; Hancock, 2006; Brown, Tauler, and Walczak, 2009a). More detailed information on methods and their underlying algorithms can be found in the references mentioned earlier.

5   CASE STUDIES

In this section, some case studies of multivariate data analysis of herbal fingerprints, developed with spectroscopic, mass spectroscopic, chromatographic, or electrodriven techniques, are discussed. The number of examples is, however, not exhaustive.

5.1   Spectroscopic Fingerprints

In the study by Ni *et al*. (2009), three preparations, that is, an aqueous extract, a volatile oil, and a mixture of both, of four herbs, that is, *Radix bupleuri* (RB), *Flos lonicerae* (FL), *Fructus forsythlae* (FF), and *Radix isatidis* (RI), are studied to evaluate their antipyretic effects. Besides, a patented Shuanghuanglian™ (SHL) oral liquid, containing FL and FF aqueous extracts, and a mixture of this oral liquid with FL and FF volatile oils were also investigated. Ibuprofen was selected as control. The antipyretic effect of all the samples was determined using results from animal experiments by measuring the temperature at seven different time points after administration of the samples, resulting in a 15×7 matrix **Y**. *UV–vis fingerprints* (250–900 nm) were recorded for all 15 samples, resulting in a 15×1300 data matrix **X**. As data pretreatments, first derivative smoothing and CC were applied. PCA was then used to compress the original data. It was concluded that 14 principal components (PCs) described 100% of the data variance. These first 14 PCs, forming a 15×14 data matrix **Z**, together with the antipyretic data matrix **Y**, were then subjected to CCA, in order to indicate the correlation between both data matrices. It was found that the first canonical variable of the UV spectra could reflect the antipyretic potency of the samples. This canonical variable was then used to predict the antipyretic activity of four test set samples, which were mixtures of 50% aqueous extract and 50% volatile oil of the four herbs in a given ratio. Predictions were found to be sufficient to use this first canonical variable as spectral index to evaluate the antipyretic effect.

UV–vis fingerprints (250–800 nm, $n = 551$) were also developed by Palacios-Morillo *et al*. (2013) to distinguish $m = 90$ tea samples of a different variety. The UV–vis spectra already allowed observing a difference in absorbance regions of the different tea varieties, that is, black tea, green tea, and Pu-erh tea. The fingerprints were then autoscaled and PCA was applied. The PC1–PC2–PC3 score plot allowed distinguishing the three different varieties, although the black and green tea samples showed some overlap. From the plot of the loadings on the first five PCs, it could be determined which absorbance ranges are correlated with which PC.

These first five PCs were then used as input variables to build several discrimination models, that is, LDA, SVM-C with a Sigmoid kernel, SVM-C with a radial basic function (RBF) kernel, RBF-ANN, and

multilayer perceptrons (MLP)-ANN. The data set was divided into a training (two-third of the samples) and a test set (one-third). To optimize the models, a jackknife CV procedure was applied. Models were evaluated based on their % overall sensitivity and % overall specificity.

The LDA plot of the second discriminant function as a function of the first allowed to completely distinguish Pu-erh tea samples from the black and green tea samples, but these latter samples showed overlap. The % overall sensitivity and the % overall specificity for the test set were 97.7% and 98.9%, respectively. The same values were found when applying sigmoid-SVM-C. The RBF-SVM-C model leads to somewhat better results for the % overall sensitivity (98.9%) and the % overall specificity (99.4%) than the LDA and sigmoid-SVM-C models. The same was observed for the RBF-ANN model (98.2% and 98.6%, respectively). The MLP-ANN model was found to be the best as both 100% overall sensitivity and overall specificity were obtained.

NIR fingerprints (1100–2498 nm, $n = 700$) were used for the identification and quality control of *Eleutherococcus senticosus* samples, which exhibit various therapeutic properties (Lucio-Gutiérrez, Coello, and Maspoch, 2011). For this purpose, fingerprints were developed for 20 *E. senticosus* samples and 32 samples of 8 other herbs, which were considered as possible adulterants (Figure 3a and b). From these eight herbs, some belong to the Araliaceae family, as *E. senticosus* (ES), whereas others are not related. The herbs belonging to the Araliaceae family are *Panax ginseng* (PG), *Panax quinquefolium* (PQ), and *Panax notoginseng* (PN). The unrelated herbs are *Lepidium meyenii* (LM), *Angelica sinensis* (AS), *Withania somnifera* (WS), *Pfaffia paniculata* (PP), and the cooking spice Ginger. Also, NIR fingerprints for 10 mixtures of *E. senticosus* with different amounts of adulterants were recorded. Each sample or mixture was recorded thrice.

Pretreatment consisted of performing SNV, FSDs, and CC and removing water bands from the NIR spectra. PCA on $m = 117$ of 156 randomly selected fingerprints of the samples (not mixtures) was performed. From the PC1–PC2 score plot (Figure 3c), it was observed that the *E. senticosus* samples could be well separated from all other samples along PC1. Four samples were indicated as being suspicious. To evaluate whether these samples have to be considered as outliers or not, an influence plot was drawn

(Figure 3d). Outlying samples will present both high leverage and high residual variance values. From this plot, one *E. senticosus* sample could clearly be considered an outlier and was omitted for the subsequent modeling. The three other suspicious samples were not considered to be outliers and were not deleted from the data.

Then, two discrimination techniques, that is, LDA and PLS-DA, and one classification technique, that is, SIMCA, were performed. Only three classes of herbs belonging to the Araliaceae family were considered to build the models, that is, ES, PG, and PQ, because their discrimination was the authors' main interest. Seventy-five samples belonging to the above three classes were considered as training set, and the 81 remaining herbal samples and the 30 mixtures were considered as test set. The outlying ES sample was included in the test set and was considered to belong to the class of the other herbal samples, that is, the non-ES samples.

Because CDA cannot be used if the number of variables (n) is higher than the total number of samples (m), PCA was applied before CDA to reduce the number of variables. In this study, the first 10 PCs were considered as input for CDA. From these, eight PCs were selected to build the CDA model. All ES test set samples were correctly discriminated as being ES samples. Except for the outlying ES sample, which was probably incorrectly labeled as belonging to the ES class, all non-ES test set samples were correctly discriminated as being non-ES samples. None of the 30 mixtures was correctly identified as adulterated samples.

For PLS-DA, a score plot indicated that the first latent variable was sufficient to distinguish the ES samples from the other two classes of PG and PQ samples. From an influence plot, no outlying calibration samples could be detected. The predicted **y** value was also plotted as a function of the measured **y** value, in order to evaluate the predictive ability of the model. From the 15 ES test set fingerprints, four were mislabeled as belonging to the non-ES class. On the other hand, all non-ES test set samples were correctly discriminated as non-ES samples. PLS-DA allowed indicating 25 of the 30 mixtures as adulterated samples.

For SIMCA, respectively, six, four, and three PCs were needed to describe the three classes. Then, a Coomans plot was drawn (Figure 3e), in which the distance of validation samples from the PG class is plotted as a function of validation samples to the
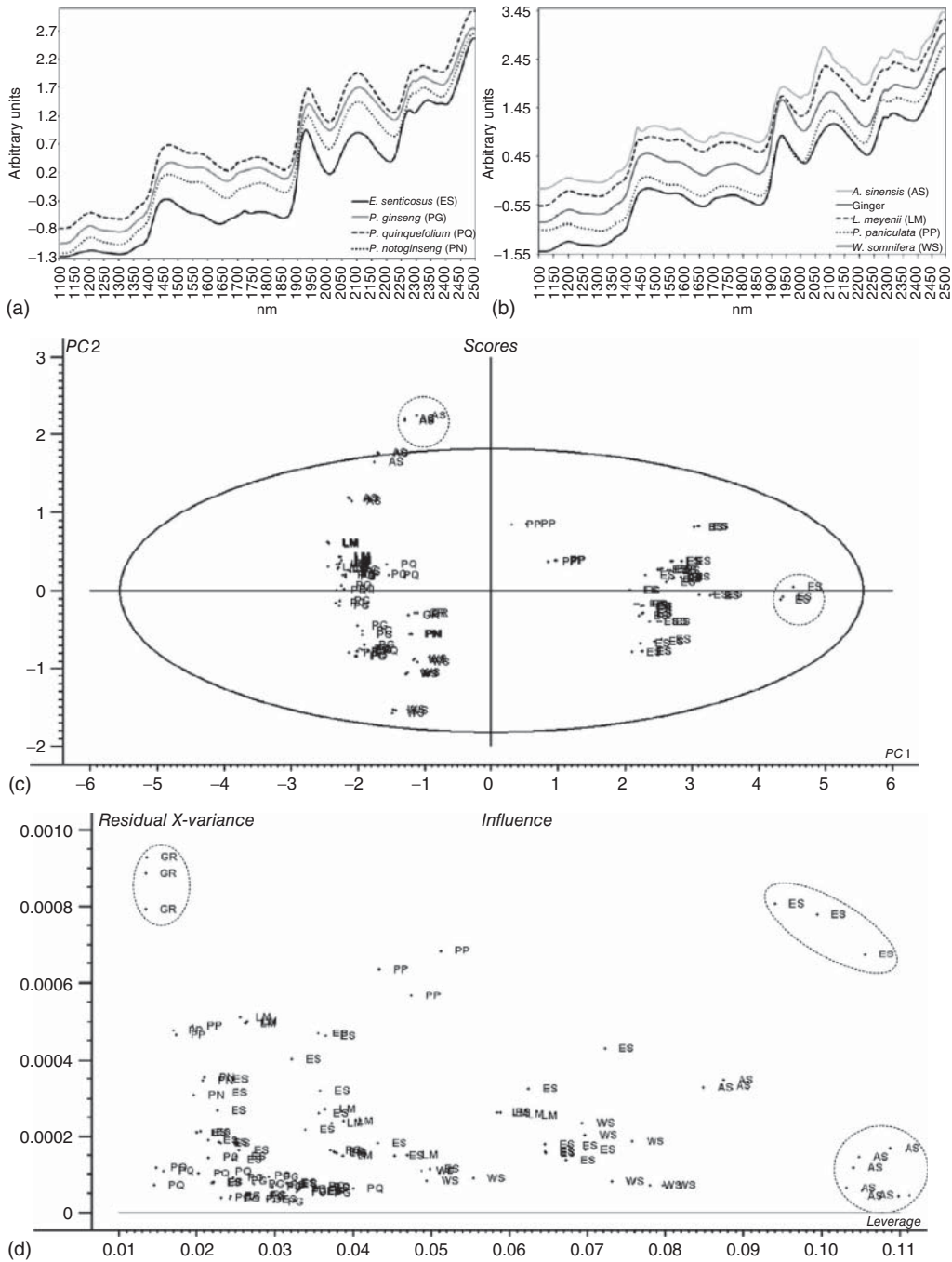
Figure 3 (a) NIR fingerprints from *Eleutherococcus senticosus* samples and other herbs of the Araliaceae family. (b) NIR fingerprints for herbs not belonging to the Araliaceae family. (c) PC1–PC2 score plot with 95% confidence ellipse. (d) Influence plot. (e) Coomans plot. (f) Distance of validation samples from the ES class as a function of leverage values. Anomalous samples were encircled. ES, *Eleutherococcus senticosus*; PG, *Panax ginseng*; LM, *Lepidium meyenii*; PQ, *Panax quinquefolium*; AS, *Angelica sinensis*; WS, *Withania somnifera*; PP, *Pfaffia paniculata*; PN, *Panax notoginseng*; and Ginger, Ginger. (*Source*: Reprinted from Food Research International, 44, Lucio-Gutiérrez J. R., Coello J., Maspoch S., Application of near infrared spectral fingerprinting and pattern recognition techniques for fast identification of Eleutherococcus senticosus, 557–565, Copyright 2011, with permission from Elsevier.)
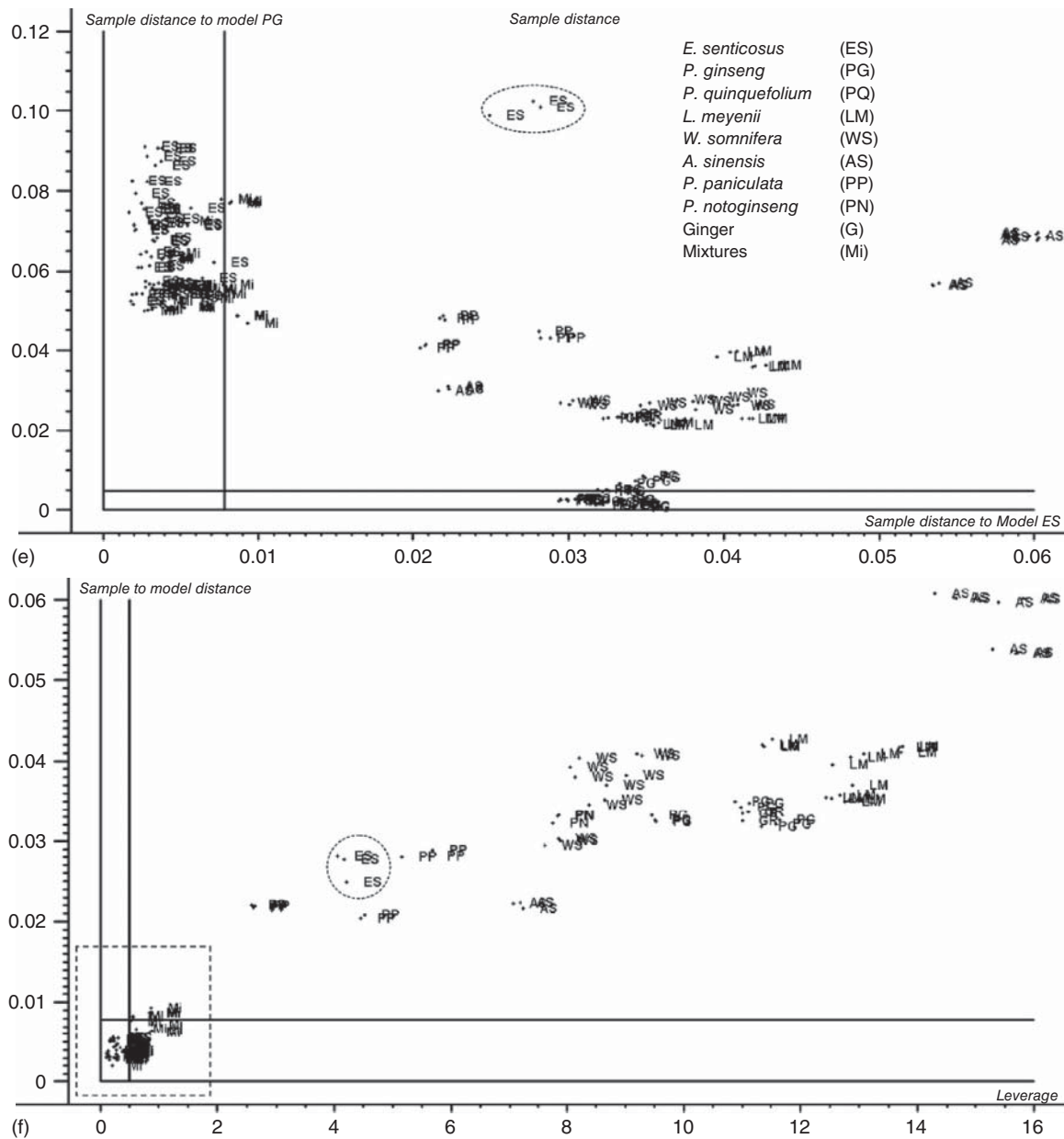
Figure 3 *(continued)*

ES class. From this plot, it was clear that all ES and PG validation samples will correctly be assigned to the ES and PG class, respectively, except for the outlying ES sample. Also, the distance of validation samples from the ES class was plotted as a function of leverage values (Figure 3f). From this plot, it was concluded that all ES validation samples are correctly discriminated as belonging to the ES class, except for the outlying ES sample. From the plot, all other validation samples clearly do not belong to the ES class. All ES and non-ES test set samples were correctly discriminated as being ES and non-ES samples, respectively. Only 12 of the 30 mixtures were correctly labeled as adulterated samples.

In conclusion, CDA, PLS-DA, and SIMCA allowed distinguishing ES samples from other samples, both those belonging and not belonging to the Araliaceae family. However, to detect adulterated samples, PLS-DA was found to be the best pattern recognition method.

In the study by Mao and Xu (2006), *Fourier transform IR* (4000–$400\,cm^{-1}$, $n = 900$), *Fourier transform NIR* ($12,000$–$4000\,cm^{-1}$, $n = 1000$), and *Fourier transform Raman* (3700–$100\,cm^{-1}$, $n = 450$) *fingerprints* were developed for three ginseng types,

that is, Radix Ginseng from China (RG1–RG5), Radix Ginseng Rubra (RGR1–RGR5 from China and KRGR from Korea), and Radix Panacis Quinquefolii (RPQ1–RPQ2 from China and RPQ3 from North-America), and two pseudo-ginseng types, that is, Radix Codonopsis (RC) and Radix Platycodi (RP). These latter two TMs are very similar to ginseng regarding appearance and smell, but their therapeutic effect is very different to ginseng. The goal of the study was the quality control of the ginseng samples. Figure 4a shows the Fourier transform
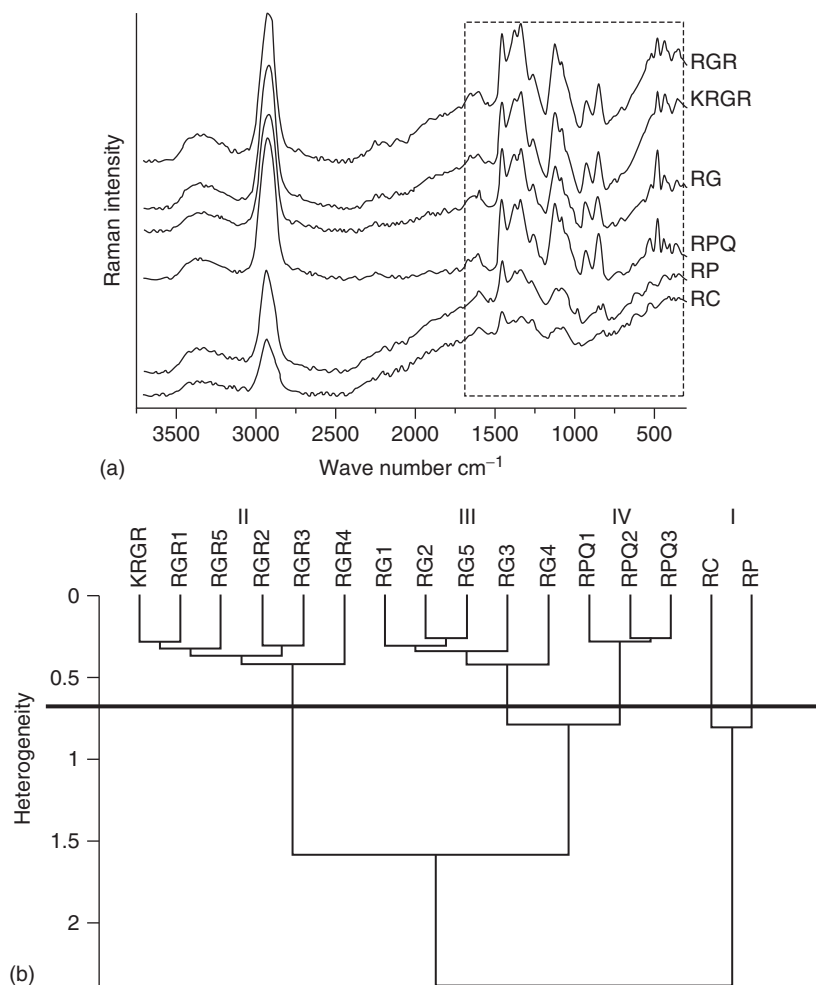
Figure 4 (a) Fourier transform Raman spectra of Radix Ginseng (RG), Radix Ginseng Rubra from China (RGR), Radix Ginseng Rubra from Korea (KRGR), Radix Panacis Quinquefolii (RPQ), Radix Codonopsis (RC), and Radix Platycodi (RP). The characteristic absorption region is indicated. (b) Dendrogram obtained using Euclidean distance as the (dis)similarity criterion and Ward's method as the linkage technique. A line is drawn that allows distinguishing four clusters: (i) pseudo-ginsengs RC and RP, (ii) RGR and KRGR ginsengs, (iii) RG ginsengs, and (iv) RPQ ginsengs. (*Source*: Reprinted from Spectrochimica Acta Part A, 65, Mao J., Xu J., Discrimination of herbal medicines by molecular spectroscopy and chemical pattern recognition, 497–500, Copyright 2006, with permission from Elsevier.)

Raman fingerprints of the six different sample types. The characteristic absorption region is indicated. The Raman fingerprints of the pseudo-ginseng samples are different from those of ginseng samples. However, within the ginseng samples, no differences among the different types were visible. The same could be observed from the Fourier transform IR fingerprints. On the other hand, the Fourier transform NIR fingerprints are similar for both ginsengs as pseudo-ginseng samples.

After preprocessing using normalization and second derivatives, high content analysis (HCA) was used to group the samples, based on a selected region of each type of spectra. Figure 4b shows the obtained dendrogram of the HCA, based on the Raman fingerprints and using the Euclidean distance as the (dis)similarity criterion and Ward's method as the linkage technique. A line is drawn that allows distinguishing four clusters. The first split in the dendrogram divides the pseudo-ginseng samples (group I in Figure 4b) from the ginseng samples (groups II, III, and IV). The second split leads to a group of RGR and KRGR samples (group II) and a cluster of RG and RPQ samples (groups III and IV). A third split then further divides the latter cluster into two clusters, that is, one with RG samples (group III) and one with RPQ samples (group IV). Similarly, dendrograms were also drawn based on the IR and NIR fingerprints. However, the best result was obtained using the Raman fingerprints.

Frédérich *et al.* (2011) developed ^1H *NMR fingerprints* of 29 *Polygonum cuspidatum* Siebold & Zucc. (PC) and 28 *Polygonum multiflorum* Thunb. (PM) samples (Figure 5a) in order to distinguish both the species. Each sample was measured twice. Common primary metabolites, such as sucrose, glucose, amino acids, and stilbene and anthraquinone derivatives, could be identified by evaluating the NMR spectra.

As data pretreatment, the spectra were baseline corrected, normalized, and reduced to integrated regions of equal width (0.04 ppm) within the region of δ 0.40–10. Then, regions of residual water signals were removed, resulting in a data matrix with $m = 57 \times 2 = 114$ samples and $n = 220$ variables. PCA was then performed. The PC1–PC2 score plot allowed discriminating both species. Only the two replicates of the PC26 sample were not found in the PC cluster but in the PM cluster. This, however, was confirmed by pharmacopeial identification tests and an HPLC-UV assay. It was suspected that the PC26 sample was mislabeled and was probably a PM

sample. Then, a PC1 loading line plot was drawn, in which the loadings on PC1 were plotted as a function of δ. On the PC1 loading line plot (Figure 5b), different stilbene derivatives contributing to the discrimination between the two species could be indicated.

5.2 Mass Spectrometric Fingerprints

DART-TOF-MS fingerprints (Figure 6a) were developed for 40 herbal samples from four species, that is, *Angelica tenuissima* (1–1 until 1–10), *Angelica gigas* (2–1 until 2–10), *Angelica dahurica* (3–1 until 3–10), and *Cnidium officinale* (4–1 until 4–10) (Lee, Kim, and Jang, 2012). All samples were measured five times and the average mass spectrum was calculated for the data analysis. As can be seen in Figure 6a, there is a clear difference in the DART-TOF-MS spectra of the four species. Each species had a specific major peak. For *A. tenuissima*, this peak was found at $m/z = 191.11623$ and corresponded to ligustilide. For *A. gigas*, the peak at $m/z = 329.13589$ was caused by decursin. For *A. dahurica*, the $m/z = 317.09142$ peak corresponded to byankangelicol. For *C. officinale*, senkyunolide A was found at $m/z = 193.11751$.

To decrease the number of variables, only those ion intensity values with the same integer m/z value were selected. Pareto-scaling and CC were then used to preprocess the data. Afterward, OPLS-DA was used to discriminate the four species. The PLS factor 2 (t2)–PLS factor 3 (t3) score plot (Figure 6b) permits a clear discrimination between the four species. One sample (sample 1–4) falls outside the 95% confidence Hotelling T^2 ellipse on the score plot, which indicates one outlier in the *A. tenuissima* group. The outlying sample might be caused by the biological variation of the herbal material, the degradation of secondary metabolites during storage, and so on. The PLS factor 2 (pq2)–PLS factor 3 (pq3) loading plot (Figure 6c) indicated four compounds as important chemical markers, that is, ligustilide at $m/z = 191.11623$ for *A. tenuissima*, decursin at $m/z = 329.13589$ for *A. gigas*, byankangelicol at $m/z = 317.09142$ for *A. dahurica*, and senkyunolide A at $m/z = 193.11751$ for *C. officinale*.

OPLS-DA was then also used to distinguish only two species, that is, *A. tenuissima*, and *C. officinale*. The PLS factor 1 (t1)–PLS factor 3 (t3) score

plot now allowed a more clearer discrimination between these two species, and the *A. tenuissima* sample 1–4 is not considered as an outlier anymore. The PLS factor 1 (pq1)–PLS factor 3 (pq3)

loading plot revealed ligustilide at $m/z = 191.11623$ as a marker for *A. tenuissima* and senkyunolide A at $m/z = 193.11751$ as a marker for *C. officinale*.
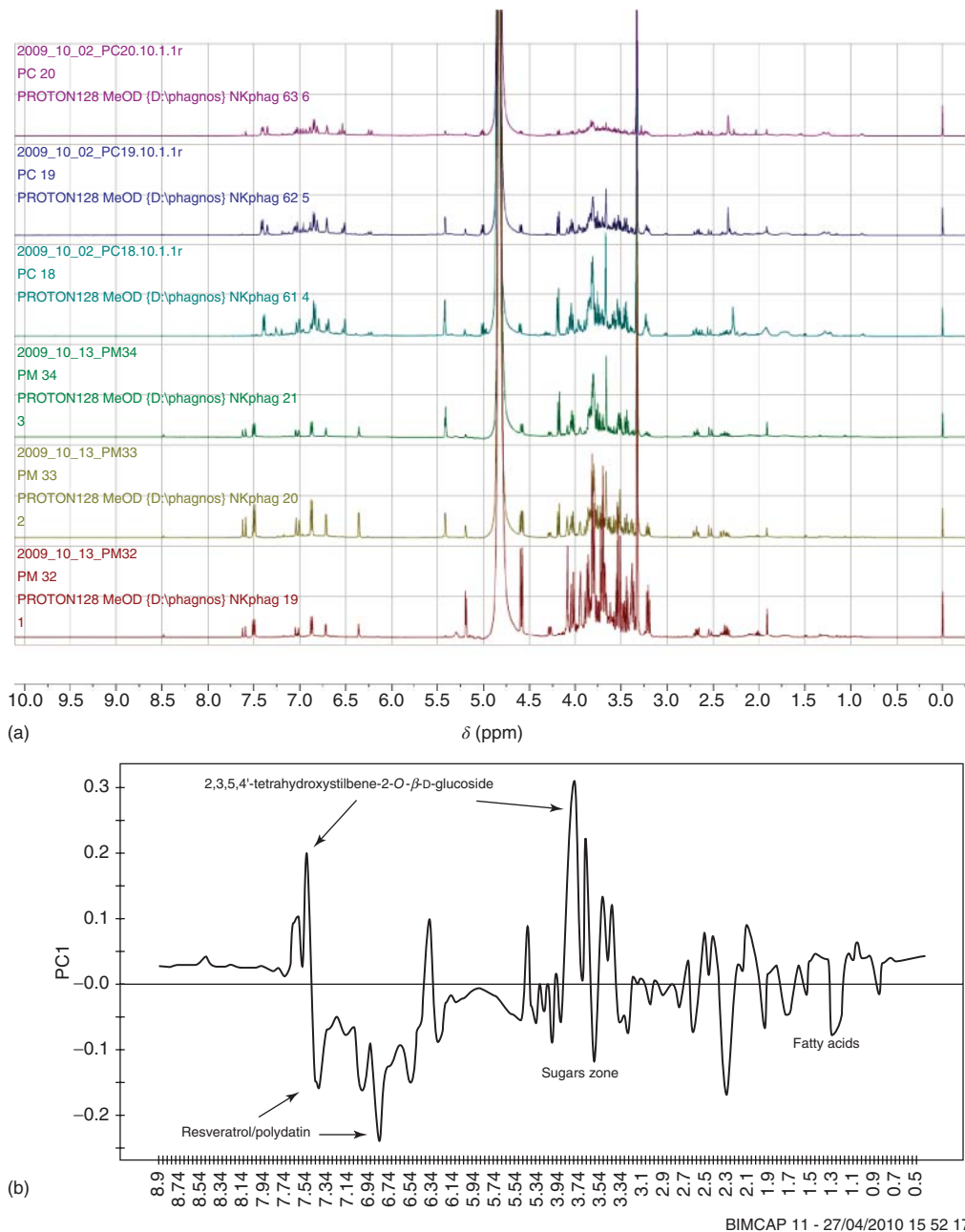
Figure 5 (a) ^{1}H NMR fingerprints of *Polygonum cuspidatum* Siebold & Zucc. (PC) and *Polygonum multiflorum* Thunb. (PM) samples. (b) PC1 loading line plot. The 95% confidence Hotelling T^2 ellipse is indicated on the score plot. (*Source*: M. Frédérich *et al*. (2011). Reproduced by permission of Thieme.)
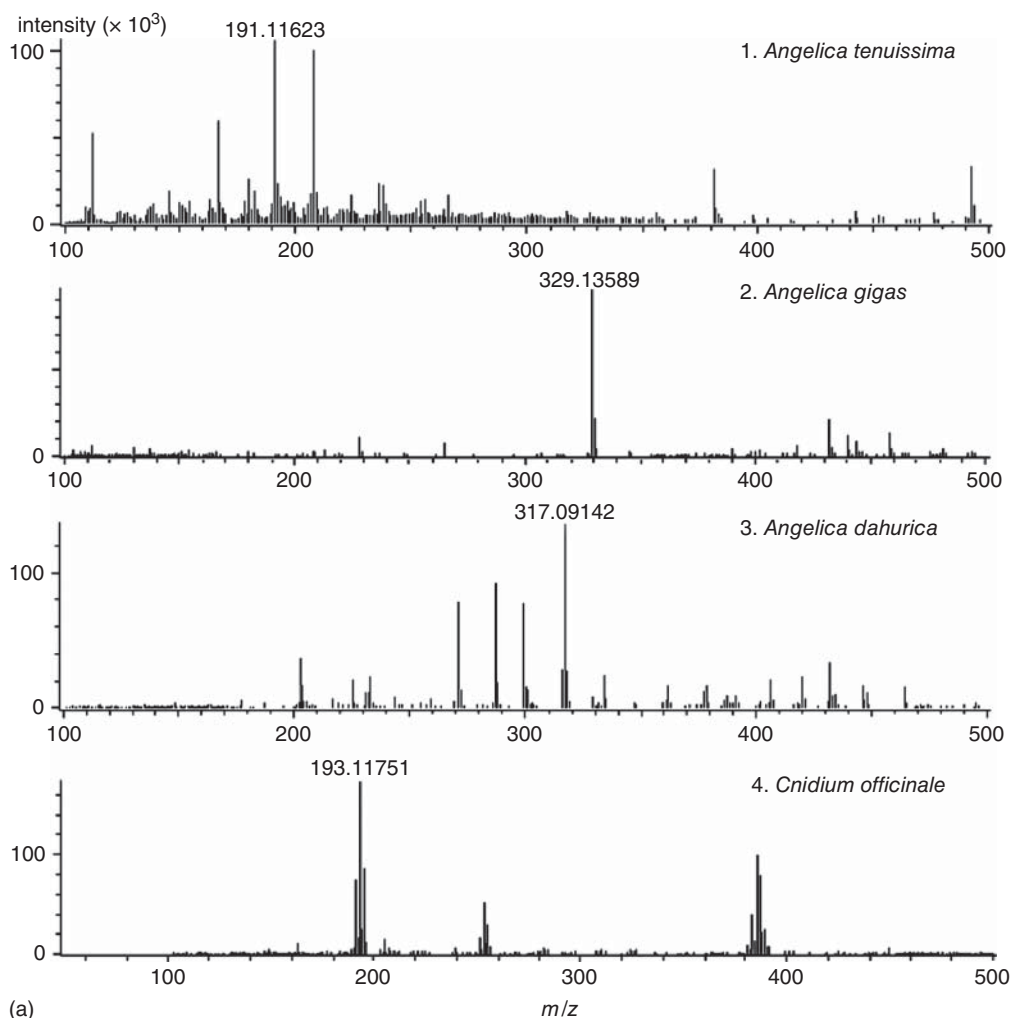
Figure 6 (a) DART-TOF-MS fingerprints of *Angelica tenuissima* (1–1 until 1–10), *Angelica gigas* (2–1 until 2–10), *Angelica dahurica* (3–1 until 3–10), and *Cnidium officinale* (4–1 until 4–10) samples. (b) PLS factor 2 (t2)–PLS factor 3 (t3) score plot of the OPLS-DA model. The 95% confidence Hotelling T^2 ellipse is indicated on the score plot. *A. tenuissima* (1–1 until 1–10, ▲), *A. gigas* (2–1 until 2–10, ■), *A. dahurica* (3–1 until 3–10, ♦), and *C. officinale* (4–1 until 4–10, ✱). (c) PLS factor 2 (pq2)–PLS factor 3 (pq3) loading plot of OPLS-DA model. Important chemical markers are indicated (♦). (*Source*: S.M. Lee *et al.* (2012). Reproduced by permission of John Wiley & Sons, Ltd.)

DART-TOF-MS fingerprints of 27 *Danshen* alkaline precipitation samples or batches were developed to monitor batch-to-batch consistency (Zeng *et al.*, 2013). The first group consisted of 15 samples (A1–A15), which were prepared traditionally. These samples were considered as reference group. The second group had 12 samples (B1–B3, C1–C3, D1–D3, and E1–E3). These samples were taken as test group. Samples B1–B3 were prepared traditionally, whereas all others were prepared in three different ways (C, D, and E). To decrease the number of variables in the

DART-TOF-MS fingerprints, only relative peak areas of $n = 15$ individual ions, visible on the selective ion monitoring spectra, were selected. Each sample was measured thrice and the average relative peak areas were calculated for data analysis. All 15 ions of the DART-TOF-MS fingerprints could be identified with standards by comparing the diagnostic ions and corresponding mass fragmentation patterns. The samples contained two important types of compounds, that is, salvianolic acids and monosaccharides.

Figure 6 *(continued)*

After CC, PCA was performed on the 15 samples of the reference group. On the PC1–PC2–PC3 score plot (Figure 7a), the 95% confidence Hotelling T^2 ellipse is indicated. All the scores of all the reference batches (A1–A15) fell inside this ellipse, which means that there is a good batch-to-batch consistency between them. This consistent product quality was also confirmed by HPLC-UV and HPLC-ELSD assays of five components. The PC1–PC2–PC3 loading plot (Figure 7b) was used to indicate correlating variables, such as the ions at $m/z = 260$ and $m/z = 377$, both of which originate from salvianic acid A.

The suitability of the PCA model for the purpose of monitoring batch-to-batch consistency was then tested using the test group samples. Batches B1–B3 fell within the indicated 95% confidence ellipse in Figure 7a, which is as expected because these test batches were prepared in the same traditional way as the reference batches A1–A15. The HPLC-UV and HPLC-ELSD assays of five components confirmed this.

Batches C1–C3 were used to evaluate the influence of input material on batch consistency. They were not included in the ellipse, indicating that these batches are out of control. According to the contribution of the ions in the MS fingerprints, the reason for the bad quality of these batches was considered to be caused by several components. This was only

Figure 7 (a) PC1–PC2–PC3 score plot. The 95% confidence Hotelling T^2 ellipse is indicated on the score plot. (b) PC1–PC2–PC3 loading plot. The plots were drawn based on DART-TOF-MS fingerprint data of 15 reference *Danshen* alkaline precipitation samples or batches (A = ▼). The different test samples are also indicated (B = ▼, C =▼, D = ▼, E = ▼). (*Source*: Reprinted from *J. Pharm. Biomed. Anal.*, **76**, Zeng S., Chen T., Wang L., Qu H., Monitoring batch-to-batch reproducibility using direct analysis in real time mass spectrometry and multivariate analysis: A case study on precipitation, 87–95, Copyright 2013, with permission from Elsevier.)

partly confirmed by the HPLC assays. The difference between MS and HPLC results is probably caused by the matrix effect that occurs during DART ionization.

Batches D1–D3 were used to evaluate the effect of a process mistake. Although they presented higher T^2 values, they were still included in the ellipse. This was not confirmed by the HPLC assays, which indicated a difference in composition between the reference batches A1–A15 and the test batches D1–D3. Possibly, the difference in composition was too small to result in a major change in the total quality, which may cause these batches to fall inside the 95% confidence ellipse.

Batches E1–E3 were used to evaluate the influence of another process mistake. Batches E1 and E2 fell outside the ellipse, as expected. On the other hand, batch E3 was found inside the ellipse, because it deviates less from the reference batches A1–A15, which was confirmed by the HPLC assays. In conclusion, the proposed strategy can be used to evaluate and monitor batch-to-batch consistency and variability.

5.3 Chromatographic Fingerprints

Tian, Xie, and Liu (2009) developed both *HPTLC-UV* and *HPLC-ELSD fingerprints* of the TM Chaihu (Bupleuri Radix). According to the Chinese Pharmacopoeia Commission (2010), genuine Chaihu corresponds to the root of *Bupleurum chinense* or *B. scorzonerifolium*. Besides, other *Bupleurum* species were also examined, that is, *B. falcatum*, *B. bicaule*, *B. longiradiatum*, and *B. marginatum* var. *stenophyllum*. The quality of these 31 lots of authenticated Chaihu samples and 33 lots of commercial samples was investigated in this study.

For the HPTLC-UV images, first, an image-warping procedure was performed to correct for retention factor R_f shifts of the bands corresponding to the different components. However, the exact algorithm was not specified. The characteristic HPTLC-UV fingerprints of the different *Bupleurum* species are shown in Figure 8a. Then, a 64×512 pixel matrix **X** was generated, which was converted from RGB color space to a grayscale image. Finally, this data matrix was normalized.

For the HPLC-ELSD fingerprints, the corresponding peaks in the different chromatograms were aligned using the nearest retention time-matching algorithm to remove retention time t_R shifts. The characteristic 60-min HPLC-ELSD fingerprints of the different *Bupleurum* species are shown in Figure 8b. A peak areas table with aligned peaks was then constructed. Thus, the data matrix **X** for further data analysis here consists of $m = 64$ samples and $n = 12$ peak areas.

For both the fingerprints, the same data analysis techniques were applied and their results were compared. First, a similarity analysis using the correlation coefficient r as a parameter was performed. Figure 8c shows the similarity analysis results of the 31 authenticated *Bupleuri* samples. It was, however, not specified whether these were obtained from the HPTLC or HPLC fingerprints. For both the fingerprints, the common pattern fingerprint of *B. chinense* was generated. Both HPTLC and HPLC similarity thresholds were then determined to decide on the (dis)similarity between the HPTLC or HPLC fingerprint of a new sample and the HPTLC or HPLC common pattern fingerprint, respectively. When r between those fingerprints is equal to or higher than the threshold, the new sample is considered to belong to *B. chinense*. When r between those fingerprints is lower than the threshold, the new sample is considered to be different from *B. chinense*. The 33 commercial samples were, thus, discriminated into qualified and nonqualified samples. For the HPTLC fingerprints, 85.7% of the commercial samples were correctly discriminated, whereas for the HPLC fingerprints, 97.6% were correctly discriminated.

Secondly, different discrimination (k-NN and ANN) and classification (SIMCA) models were also constructed and evaluated for both the fingerprints, in order to distinguish genuine and fake *B. chinense* samples and to discriminate superior quality samples from inferior quality samples. SIMCA models were found to have a very low correct classification rate and are, therefore, not further discussed in this chapter. For k-NN, 100% of the commercial samples were correctly discriminated from both the HPTLC and the HPLC fingerprints. For the HPTLC fingerprints, a three-layer feed-forward back-propagation ANN structure was used, leading to 88.1% correctly classified commercial samples. For the HPLC fingerprints, a two-layer ANN was applied, resulting in 100% correct classification rate. In conclusion, the HPLC fingerprints generally lead to models with a higher correct classification rate than the HPTLC fingerprints.

HPLC-UV fingerprints of 55 green tea samples (data matrix **X**) were developed in order to predict
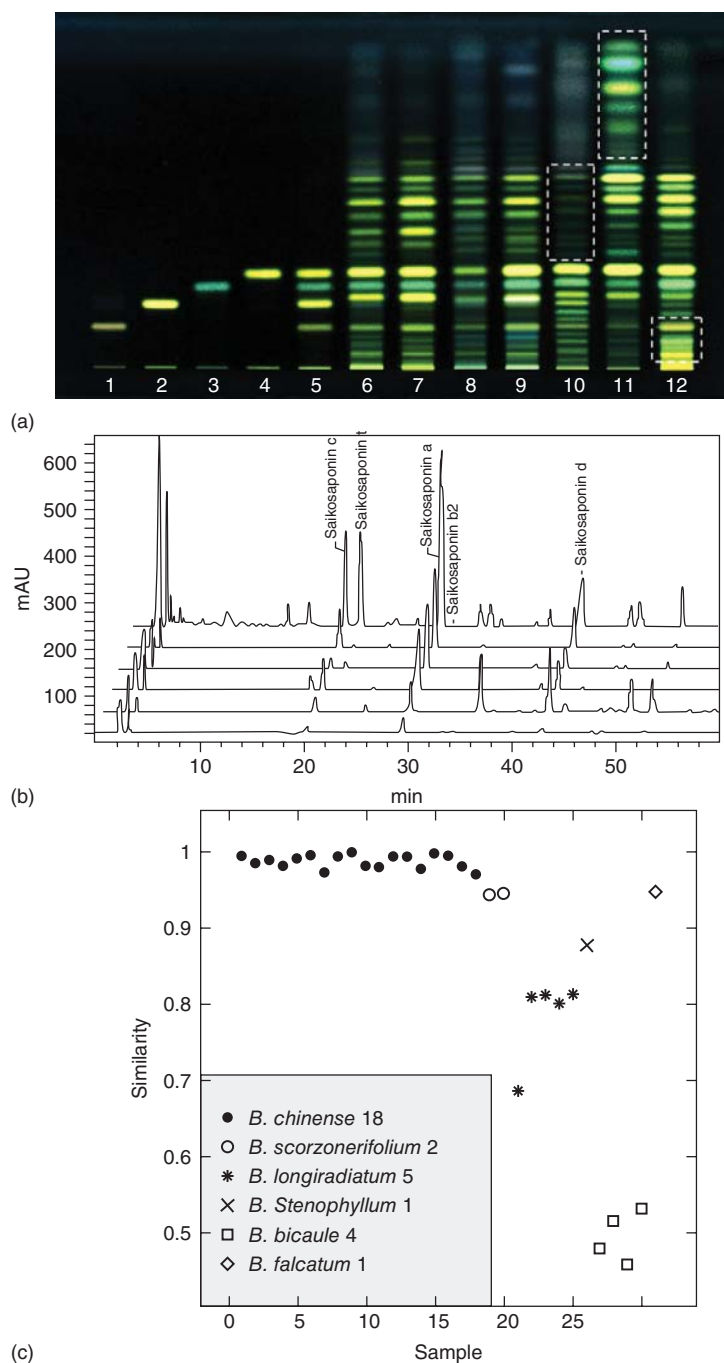
Figure 8 (a) Characteristic HPTLC-UV fingerprints of the different *Bupleurum* species. 1–4, saikosaponin references; 5, reference mixture; 6–7, *B. chinense*; 8, *B. scorzonerifolium*; 9, *B. falcatum*; 10, *B. longiradiatum*; 11, *B. bicaule*; and 12, *B. marginatum* var. *stenophyllum*. (b) Characteristic HPLC-ELSD fingerprints of the different *Bupleurum* species. From top to bottom: *B. marginatum* var. *stenophyllum*, *B. falcatum*, *B. chinense*, *B. longiradiatum*, *B. bicaule*, and *B. scorzonerifolium*. (c) Similarity analysis results of the 31 authenticated *Bupleuri* samples. (*Source*: Reprinted from *J. Chromatogr. A*, **1216**, Tian R-T., Xie P-S., Liu H-P., Evaluation of traditional Chinese herbal medicine Chaihu (Bupleuri Radix) by both high-performance liquid chromatographic and high-performance thin-layer chromatographic fingerprint and chemometric analysis, 2150–2155, Copyright 2009, with permission from Elsevier.)

the antioxidant activity of the samples (response vector **y**) (van Nederkassel *et al.*, 2005). Both 11- and 2-min (Figure 9a) fingerprints were recorded, and each sample was analyzed in duplicate. Because retention time t_R shifts were observed between the fingerprints (Figure 9a), the chromatograms were aligned using COW. The aligned 2-min fingerprints are shown in Figure 9b.

To detect outliers in the data matrix **X** of the fingerprints, rPCA was applied. A score diagnostic plot (Figure 9c) was drawn in which the distance of a sample from the PCA model space, that is, the orthogonal distance, is plotted as a function of the distance of a sample to the data majority, that is, the robust distance. For both distances, a cutoff value is calculated, which divides the plot into four quadrants that allows visualizing three different types of outlying objects. Ordinary objects do not exceed both cutoff values (Quadrant III). When objects exceed the cutoff value of the orthogonal distance, but not that of the robust distance, they are called *orthogonal outliers* (Quadrant IV). Objects that exceed the cutoff value of the robust distance but not that of the orthogonal distance are defined as good leverage objects (Quadrant II). Bad leverage objects exceed both cutoff values (Quadrant I). From Figure 9c, it was decided to remove the two indicated orthogonal outliers. To evaluate the response vector **y** with antioxidant activity values for outliers, a histogram of the values (Figure 9d) was drawn. From Figure 9d, it was decided to remove two outlying activity values, that is, one extremely low and one extremely high value. Thus, three samples were removed from the data. Besides the alignment and the outlier removal, the performed data preprocessing before modeling is, however, not discussed in the paper of van Nederkassel *et al.* (2005).

The data set was divided into calibration and test sets. The calibration set was selected by uniformly selecting samples according to their sorted antioxidant activity values. The antioxidant activity was modeled from the 11- and 2-min fingerprints using several multivariate calibration techniques, that is, Step-MLR, PCR, PLS, UVE-PLS, and OPLS (Dumarey *et al.*, 2008). Another paper modeled the activity from the 2-min fingerprints using PLS without outlier removal, PLS with outlier removal, and rPLS without outlier removal (Daszykowski, Vander Heyden, and Walczak, 2007). The results are presented in Table 2.

The goal of this study was to predict the antioxidant activity for new samples from their measured fingerprints and the obtained calibration models. Thus, the models should primarily be evaluated based on their predictive ability. Step-MLR, PCR, PLS, UVE-PLS, and OPLS models were found to have a similar predictive ability (Dumarey *et al.*, 2008). PLS without outlier removal was found to have a poor predictive ability, whereas PLS after outlier removal and rPLS without outlier removal resulted in a similar predictive ability. Thus, rPLS enables obtaining a good activity prediction even with outliers in the data set (Daszykowski, Vander Heyden, and Walczak, 2007).

Besides the predictive ability, which is the most important evaluation criterion for the goal of this study, Dumarey *et al.* (2008) also studied the interpretability of the regression coefficients of the models, in order to indicate peaks with compounds that contribute most to the final calibration models and, thus, to the antioxidant activity. Because the higher the **y** value of a sample was, the higher its antioxidant activity was, peaks with positive regression coefficients correspond to either antioxidant compounds, for example, (−)-epigallocatechin gallate peak, or compounds for which the concentration is correlated with the antioxidant activity, for example, caffeine peak. The Step-MLR model did not allow indicating such peaks because variables in the tail of the peaks were selected. PCR was also not recommended because the model included a large amount of variation, not correlated with the activity. The PLS model led to regression coefficients that were not easy to interpret. UVE-PLS led to an easier interpretation, but when repeating the method, different variables will be selected. The best and easiest model to interpret was found to be OPLS, where compounds contributing to the activity could be indicated.

HPLC-DAD fingerprints of 39 *Mallotus* samples from various species (data matrix **X**) were developed (Tistaert *et al.*, 2009) in order to indicate potentially interesting components responsible for the antioxidant activity (response vector **y**). Fingerprints with three different lengths (60, 35, and 22.5 min) were developed.

Aligning corresponding peaks was not evident, given the large diversity of the samples. An in-depth discussion about the need and/or obligation or not for alignment is given by Tistaert *et al.* (2012). Different data preprocessing methods, that is, CC, normalization followed by CC, and SNV followed
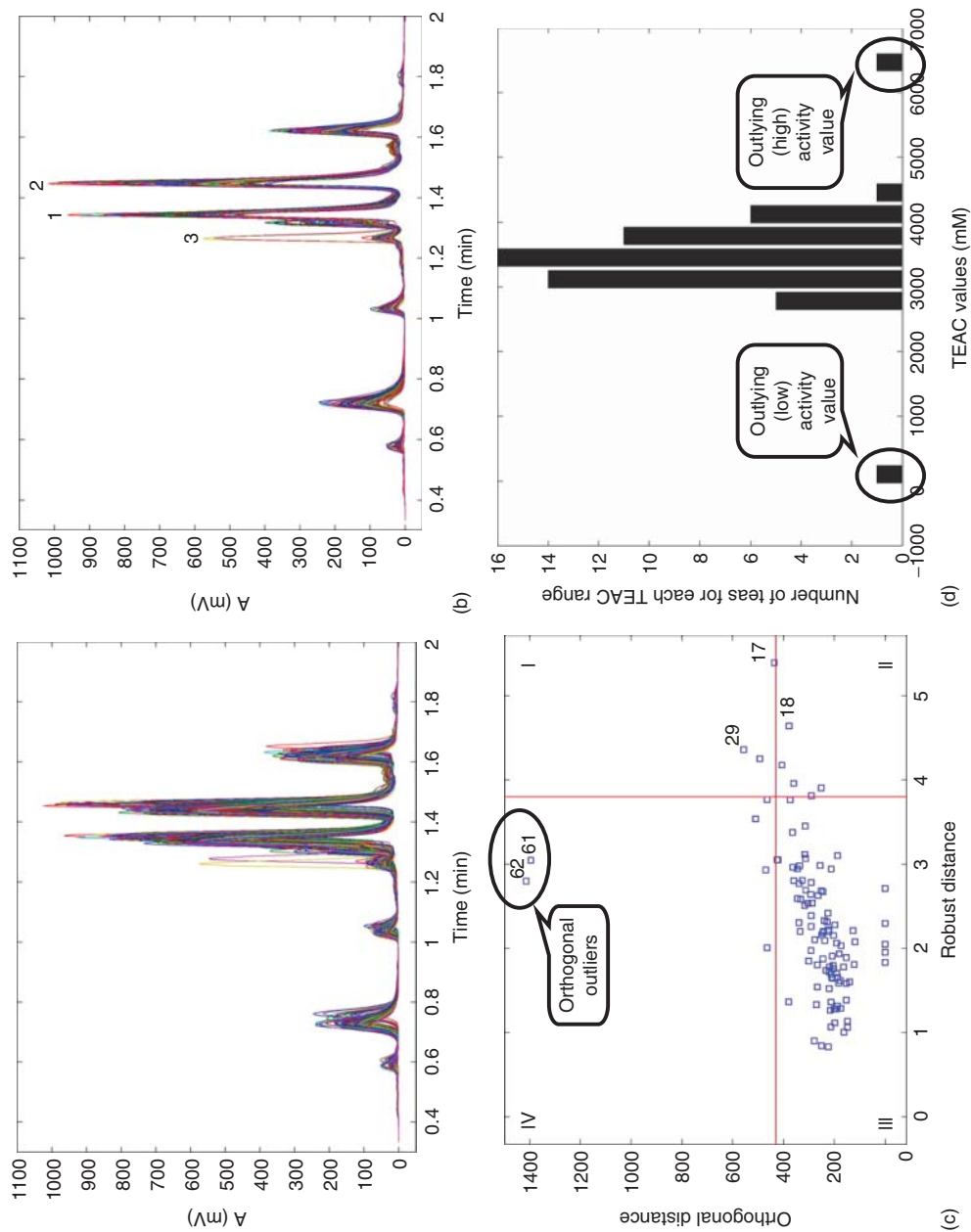
Figure 9 (a) Raw HPLC-UV and (b) HPLC-UV fingerprints aligned using COW of 55 duplicated green tea samples. Peak 1 = caffeine, peak 2 = (−)-epigallocatechin gallate, and peak 3 = unidentified substance. (c) Score diagnostic plot of the 2-min chromatograms. The removed orthogonal outliers are indicated. (d) Histogram of the 55 antioxidant activity values. Removed outliers with atypical activity values are indicated. (*Source:* Reprinted from *J. Chromatogr. A*, **1096**, van Nederkassel A. M., Daszykowski M., Massart D. L., Vander Heyden Y., Prediction of total green tea antioxidant capacity from chromatograms by multivariate modeling, 177–186, Copyright 2005, with permission from Elsevier.)

Table 2 Multivariate calibration models to predict the antioxidant activity of green tea samples from both 2- and 11-min HPLC-UV fingerprints.

Multivariate calibration technique	Short (2 min) fingerprints			Long (11 min) fingerprint		
	RMSECV	RMSE	RMSEP	RMSECV	RMSE	RMSEP
Step-MLR after outlier removal	214	162	186	182	140	86
PCR after outlier removal	216	189	192	227	194	227
PLS after outlier removal	206	177	177	159	80	174
PLS without outlier removal	/	721	350	/	/	/
rPLS without outlier removal	/	172	186	/	/	/
UVE-PLS after outlier removal	215	195	208	158	105	198
OPLS after outlier removal	209	177	176	166	80	168

/, not specified.

Source: M. Daszykowski, *et al.* (2007) and M. Dumarey, *et al.* (2008). Adapted from Elsevier.

Table 3 Multivariate calibration model results to indicate potentially antioxidant active components in *Mallotus* samples from (a) 22.5, 30, and 60-min HPLC-DAD fingerprints and (b) pCEC-UV fingerprints, after using normalization and column centering as preprocessing.

Multivariate calibration technique	22.5-min HPLC fingerprints		30-min HPLC fingerprints		60-min HPLC fingerprints		pCEC fingerprints	
	No. of components	RMSECV	No. of components	RMSECV	No. of components	RMSECV	No. of components	RMSECV
Step-MLR (LOOCV)	10	15.3	14	16.3	9	16.2	/	/
PCR (LOOCV)	3	21.8	7	15.1	6	15.1	/	/
PLS (LOOCV)	6	11.4	7	13.7	5	12.7	3	19.4
PLS (MCCV)	4	10.6	4	11.9	4	12.3	/	/
UVE-PLS (LOOCV)	4	7.5	4	9.0	3	11.6	/	/
UVE-PLS (MCCV)	4	7.0	4	8.2	4	10.0	/	/
OPLS (LOOCV)	1(1*)	13.4	1(1*)	14.2	1(1*)	13.8	1(1*)	27.0

For OPLS, the number of removed orthogonal components is given between brackets.

/, not performed.

Source: Reprinted from *Anal. Chim. Acta*, **652**, Tistaert, C., Dejaegher, B., Nguyen Hoai, N., Chataigné G., Rivière, C., Nguyen Thi Hong, V., Chau Van, M., Quetin-Leclercq, J., Vander Heyden, Y., Potential antioxidant compounds in Mallotus species fingerprints. Part I: Indication, using linear multivariate calibration techniques, 189–197, Copyright 2009, with permission from Elsevier.

by CC, were evaluated in combination with PCA and different multivariate calibration techniques. The best preprocessing was obtained using normalization followed by CC. On the PC1-PC2 score plot, the antioxidant samples are situated together, but unfortunately, they cannot be clearly distinguished from the nonactive samples.

To indicate compounds potentially responsible for the activity, different multivariate calibration techniques were applied and compared, that is, Step-MLR, PCR, PLS, UVE-PLS, and OPLS. The interpretability of the regression coefficients of the model was the most important evaluation criterion to achieve the goal of this study. Besides, the predictive ability of the different models was also determined. The number of components in the model and the RMSECV results of the different techniques are

presented in Table 3. For all methods, an LOOCV procedure was applied. In addition, for PCR, PLS, and UVE-PLS, an MCCV procedure was also applied. The model with the lowest RMSECV was found to be UVE-PLS.

Indicating potentially interesting components responsible for the antioxidant activity was done by localizing peaks in the fingerprints responsible for the antioxidant activity from evaluating the regression coefficients of different multivariate calibration models. The regression coefficients from Step-MLR, PCR, PLS, UVE-PLS, and OPLS are shown in Figure 10. In this figure, the fingerprints of all antioxidant active samples and of a mixture of reference compounds were also plotted. Because the lower the **y** value of a sample was, the higher its antioxidant activity was, peaks with negative regression
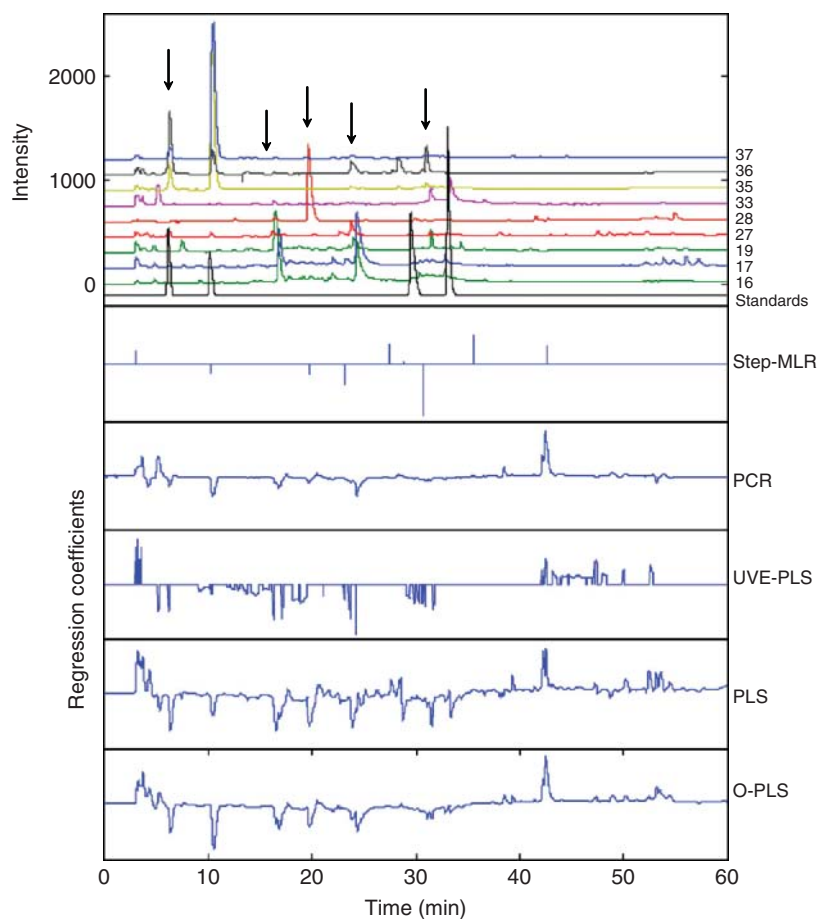
Figure 10 Sixty-minute HPLC-DAD fingerprints of the antioxidant *Mallotus* samples and standard compounds mallonanoside B (first peak of standards), mallonanoside A (second peak), myricetin (third peak), and quercetin (fourth peak). The bottom plots present the regression coefficients from Step-MLR, PCR, PLS, UVE-PLS, and OPLS after normalization and column centering as preprocessing. (*Source*: Reprinted from *Anal. Chim. Acta*, **652**, Tistaert C., Dejaegher B., Nguyen Hoai N., Chataigné G., Rivière C., Nguyen Thi Hong V., Chau Van M., Quetin-Leclercq J., Vander Heyden Y., Potential antioxidant compounds in Mallotus species fingerprints. Part I: Indication, using linear multivariate calibration techniques, 189–197, Copyright 2009, with permission from Elsevier.)

coefficients (indicated with arrows in Figure 10) are either antioxidant compounds or compounds for which the concentration is correlated with the antioxidant activity. Step-MLR and PCR gave no clear view on the peaks' individual contribution to the regression model. PLS and UVE-PLS presented a somewhat clearer view, but interpretation remained quite difficult and not straightforward. As was the case for the green tea fingerprints, OPLS gave the clearest and easiest interpretable view on the individual contributions of the peaks to the final regression model. Then, the indicated peaks were further examined using HPLC-DAD-MS, in order to identify the compounds under these peaks (Tistaert *et al.*, 2012).

UPLC-DAD fingerprints were developed for the quality control of 10 Rhizoma Coptidis samples from different origins, that is, from three different cultivation locations (Kong *et al.*, 2009). Eleven characteristic peaks were detected in a run of 3 min. Using UPLC over HPLC, thus, leads to a huge decrease in analysis time and an increase in peak capacity.

Similarity analysis was performed using a professional software called *Similarity Evaluation System for Chromatographic Fingerprint of Traditional Chinese Medicine* of the Chinese Pharmacopoeia Committee (Chinese Pharmacopoeia Commission, 2010) and recommended by the SFDA of China (Chinese SFDA Guidelines, 2002). A mean
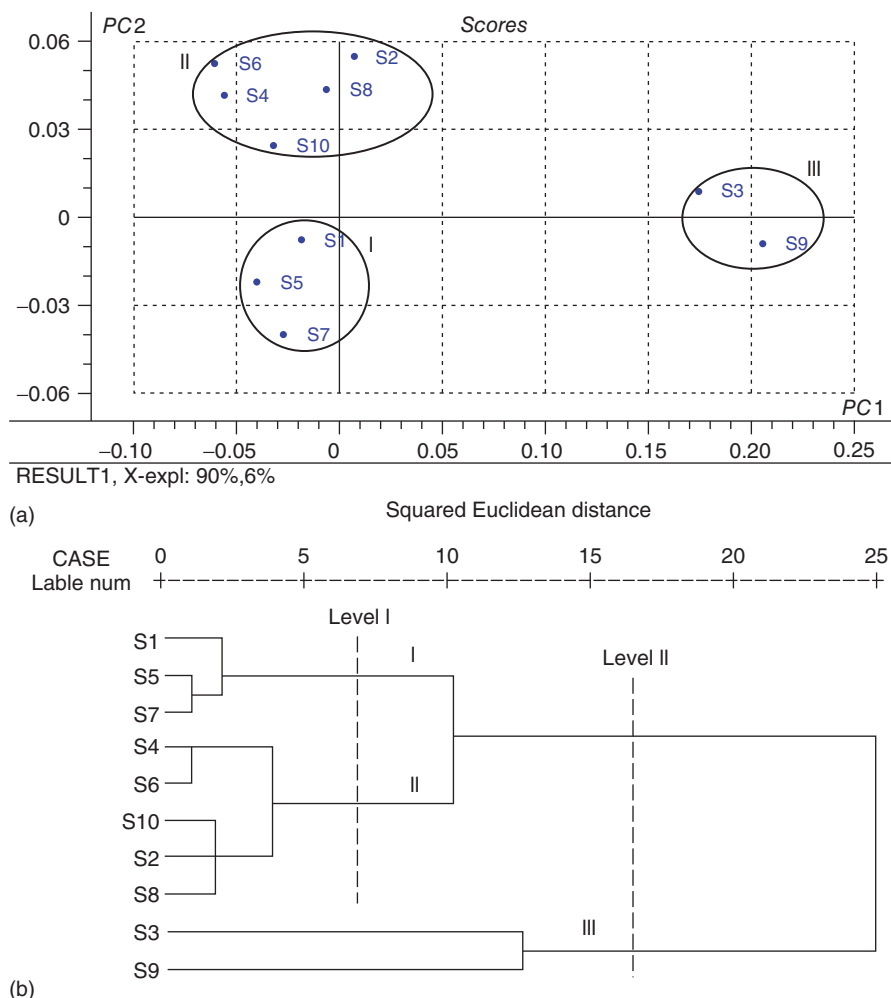
Figure 11   (a) PC1–PC2 score plot. (b) Dendrogram obtained using squared Euclidean distance as the (dis)similarity criterion and Ward's method as the linkage technique. The plots were drawn based on UPLC-DAD fingerprints of 10 Rhizoma Coptidis samples from three different cultivation locations (groups I–III). (*Source*: Reprinted from *Phytomedicine*, **16**, Kong W-J., Zhao Y-L., Xiao X-H., Jin C., Li Z-L., Quantitative and chemical fingerprint analysis for quality control of Rhizoma Coptidischinensis based on UPLC-PAD combined with chemometrics methods, 950–959, Copyright 2009, with permission from Elsevier.)

chromatogram was calculated and used as a representative fingerprint. The correlation coefficient r was used as a parameter and was calculated between each chromatogram and the representative fingerprint. Samples from the same origin had similar r values. Sample S1 had the highest r value and sample S9 the lowest. This is in accordance with the UPLC assay results of five alkaloids in all samples. S1 has the highest total content, whereas S9 had the lowest.

A relative peak areas table was then formed. For PCA and HCA, the data matrix **X**, thus, consists of $m = 11$ samples and $n = 11$ variables. The PC1–PC2 score plot (Figure 11a) revealed a clear distinction between the three cultivating locations (groups I–III). Groups I and II are located near each other while both groups are situated far away from group III, indicating that samples from the cultivating regions of groups I and II are closely related and had a similar quality.

Figure 11b shows the dendrogram obtained using the squared Euclidean distance as the (dis)similarity criterion and Ward's method as the linkage technique. A first split (level II in Figure 11b) separates the samples of group III from those of groups I and II.

A second split (level I) then separates the latter two groups. Thus, similar conclusions could be drawn from PCA and HCA, and both methods could be used to determine the geographical origin and the quality of Rhizoma Coptidis samples.

Grata *et al*. (2007) developed a two-step approach to determine metabolite modifications when wounding leaf extracts of *Arabidopsis thaliana*. A fast fingerprinting step using HPLC-MS was, therefore, followed by a metabolite profiling step using high resolution UPLC-MS.

In the first step, a *fast HPLC-MS screening* of 48 *A. thaliana* samples (24 control and 24 wounded plant samples) was performed in order to indicate metabolites supposedly involved in wound inducing. For this purpose, a short column (RP18, 20×2.1 mm ID, 5 µm) and a ballistic gradient were used, resulting in a total analysis time of 7 min. All chromatograms were summed, collapsing the retention time information and resulting in total mass spectra (TMS) (Figure 12a). After a linear normalization procedure and reduction to rounded unit resolution, autoscaling of the data was performed. Then, PCA was applied after orthogonal signal correction (OSC) as preprocessing. The PC1–PC2 score plot allowed a clear discrimination between control and wounded samples. From the PCA loadings, variables responsible for the discrimination between the two classes could be indicated. Then, HCA, using the Euclidean distance as the (dis)similarity criterion and complete linkage as the linkage technique, was applied on the first five OSC-PCA PCs. This allowed indicating two outlying wounded samples, which were removed for future analysis. Moreover, HCA indicated similar nonwounded and wounded samples, which were, respectively, pooled to obtain sufficient plant material for the subsequent UPLC-MS analysis.

In the second step, a long *high resolution UPLC-MS* method was used to analyze the pooled samples, in order to indicate the components induced by wounding of *A. thaliana* samples. For this purpose, a long column (C18, 150 mm $\times 2.1$ mm I.D., 1.7 µm) and a shallow gradient of 100 min were applied. The obtained TMS spectra from the UPLC-MS analyses were compared to those from the fast HPLC-MS screening analyses. From the TMS spectra, four ions were clearly indicated as being present in the wounded samples, and hardly at all in the control samples. These were identified as four oxylipins, that is, jasmonic acid (*m/z* 209), hydroxyjasmonic acid (*m/z* 225), 3-oxo-2(2[Z]-pentenyl)-cyclopentane-1-butanoic acid (*m/z* 237), and dinor-oxo-phytodienoic acid (*m/z* 263). In addition, the formate adduct of jasmonic acid was also found in the wounded samples at *m/z* 255. These wound biomarkers were then localized in the UPLC-MS total ion current (TIC) traces (Figure 12b). The traces at *m/z* 209 and 263 showed single sharp peaks. The trace at *m/z* 237 indicated two partially separated peaks, which correspond to the two isomers of 3-oxo-2(2[Z]-pentenyl)-cyclopentane-1-butanoic acid.

The proposed two-step approach, thus, allowed indicating metabolite changes when wounding leaf extracts of *A. thaliana*, and the wound biomarkers could be precisely indicated.

Conventional *GC-FID* (38 min, $n = 5455$), *GC-MS* (28 min, $n = 4831$), and *fast GC-FID* (14 min) *fingerprints* (data matrices **X**) were developed from green tea samples in order to predict their quality (Jumtee, Bamba, and Fukusaki, 2009). The quality ranking ranged between 0 and 60 (response vector **y**), and was judged by professional tea tasters. The fingerprints were normalized, baseline corrected, and aligned.

Then, PLS and OPLS calibration models were constructed from the conventional GC-FID and GC-MS fingerprints and the quality ranking. Both fingerprints had similar chromatographic patterns for a given sample. For both methods, the GC-FID fingerprints lead to a slightly better predictive ability than the GC-MS fingerprints. For both types of fingerprints, the OPLS model showed a lower RMSEP, and thus, a better predictive ability than the PLS model. The best model, obtained with OPLS from the GC-FID fingerprints, had an RMSEP of 1.660. The predictions were found to be very accurate.

Then, fast GC-FID fingerprints were developed and OPLS was used to construct a calibration model that would allow prediction of the quality ranking of a green tea sample. The RMSEP was found to be 3.464, which is worse than the OPLS models obtained from the conventional GC-FID (RMSEP = 1.660) and the GC-MS (RMSEP = 1.913) fingerprints. Figure 13 shows the observed ranking as a function of the predicted ranking for the OPLS model from the fast GC-FID fingerprints. Predictions were found to be less accurate as those from the two other types of fingerprints, but they were still found to be acceptable.

In the study by Parastar *et al*. (2012), *GC-MS fingerprints* of 18 citrus fruit peels were developed to evaluate their distribution of secondary metabolites in order to indicate the differences between the
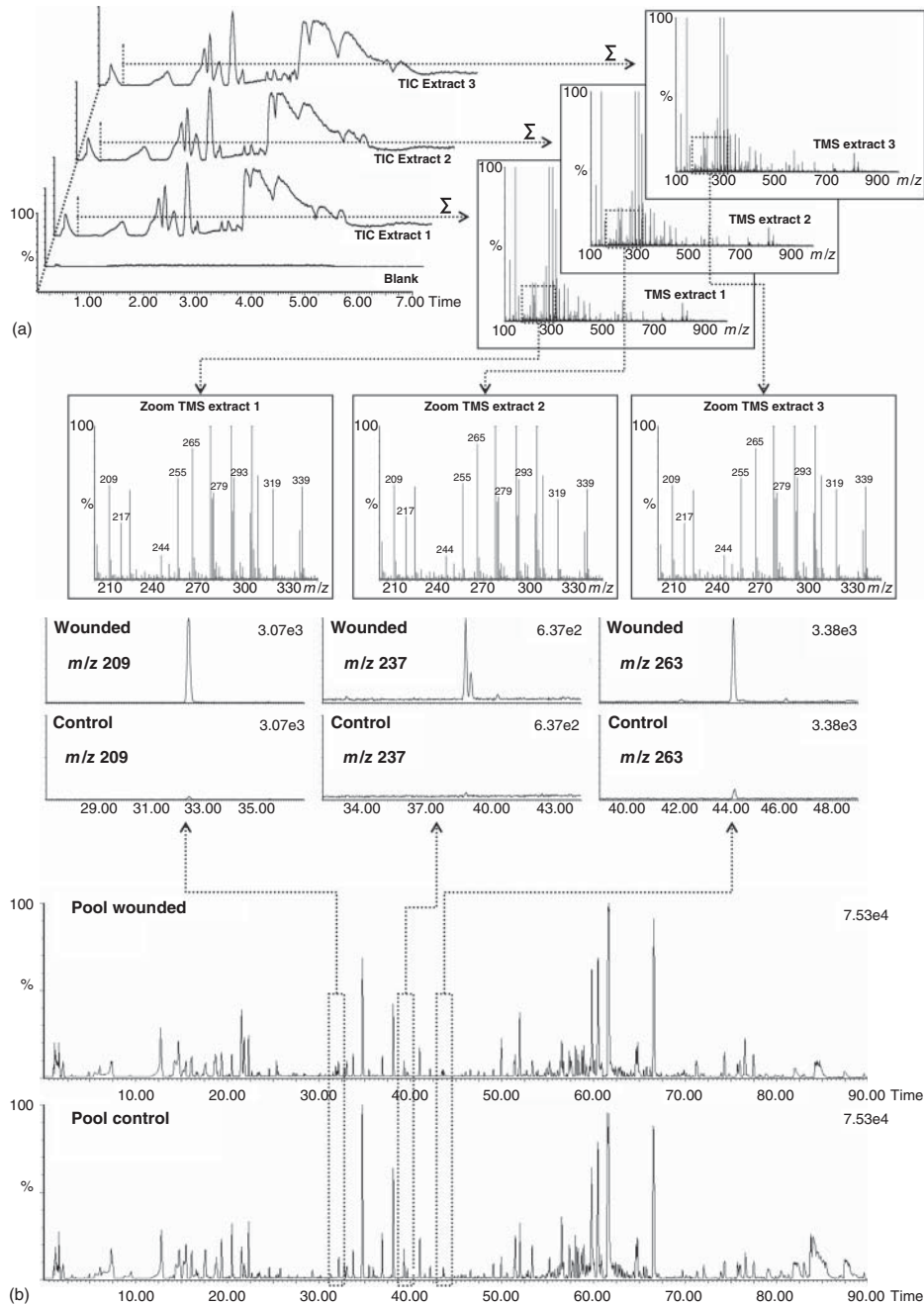
Figure 12 (a) Fast HPLC-MS screening of *A. thaliana* samples: TIC chromatograms and their corresponding TMS obtained for three extractions of the same sample. (b) High resolution UPLC-MS metabolite profiling chromatograms of nonwounded and wounded plants. (*Source*: Reproduced with permission from Grata, E., Boccard, J., Glauser, G., Carrupt, P.-A., Farmer, E. E., Wolfender, J.-L., Rudaz, S. 2007. Development of a two-step screening ESI-TOF-MS method for rapid determination of significant stress-induced metabolome modifications in plant leaf extracts: The wound response in Arabidopsis thaliana as a case study. *J. Sep. Sci.* **30** 2268–2278. © 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.)
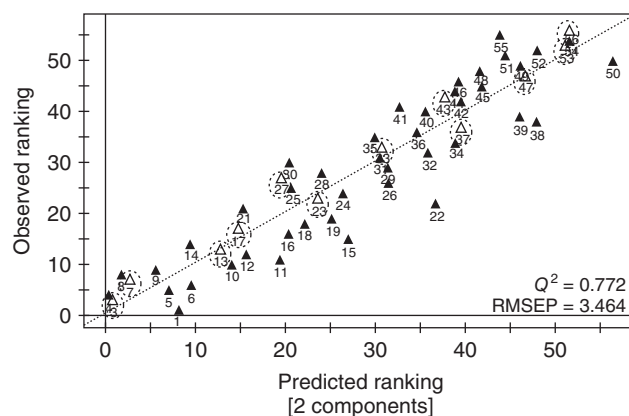
Figure 13 Observed versus predicted ranking results from an OPLS model to model the quality ranking (between 0 and 60) of green tea samples as a function of their fast GC-FID fingerprints. Both samples from the calibration (▼) and the test (Δ) are shown. (*Source*: Reproduced with permission from Jumtee, K., Bamba, T., Fukusaki, E. 2009. Fast GC-FID based metabolic fingerprinting of Japanese green tea leaf for its quality ranking prediction. *J. Sep. Sci.* **32** 2296–2304. © 2009 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.)
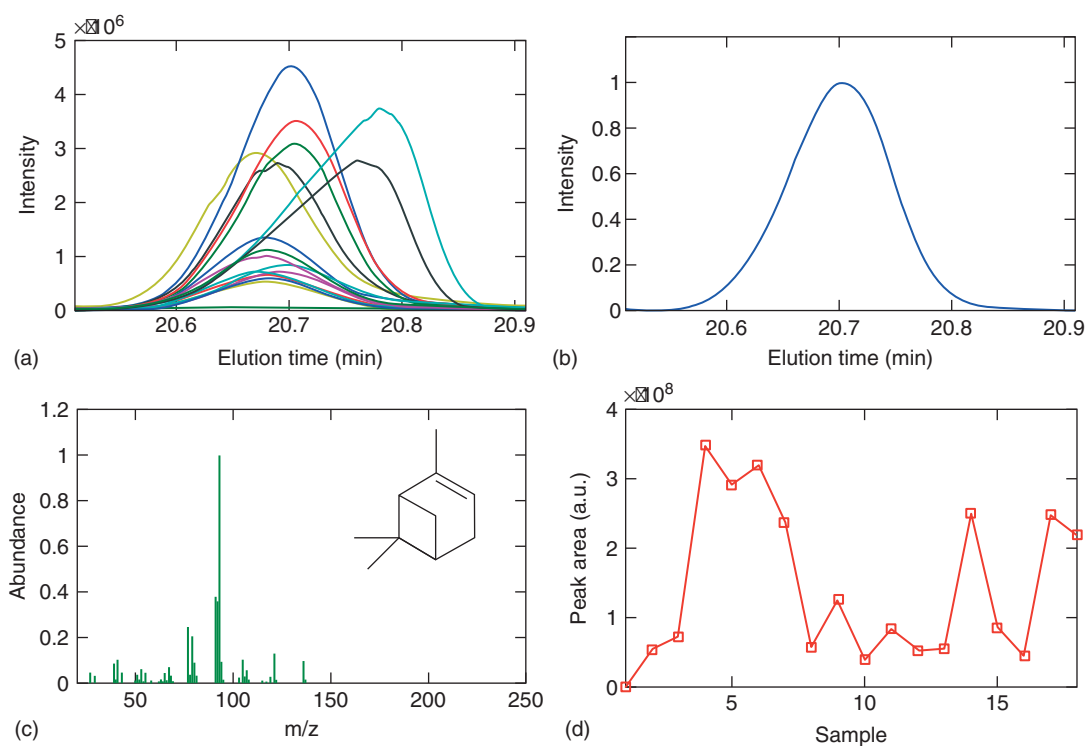
Figure 14 Results from E-MCR-ALS for the GC-MS segment between 20.5 and 20.9 min. (a) Overlaid TIC chromatograms of 18 citrus fruit peel samples, (b) resolved elution profile, (c) resolved mass spectral profile, and (d) relative concentration of this resolved component in the 18 samples. (*Source*: Reprinted from *J. Chromatogr. A*, **1251**, Parastar, H., Jalili-Heravi, M., Sereshti, H., Mani-Varnosfaderani, A., Chromatographic fingerprint analysis of secondary metabolites in citrus fruits peels using gas chromatography-mass spectrometry combined with advanced chemometric methods, 176–187, Copyright 2012, with permission from Elsevier.)

samples. Each sample was analyzed in triplicate. The chromatograms were different for different types of samples. The TIC chromatograms were then divided into some regions, which were then columnwise augmented with elution times as columns and *m/z* values as rows. This will allow maintaining the bilinear model assumption of MCR. A baseline correction, denoising, and smoothing was also performed before applying E-MCR-ALS. This method allowed obtaining the pure elution (Figure 14b) and mass spectral (Figure 14c) profiles of all components in each chromatographic segment (Figure 14a) and their relative concentrations (Figure 14d).

These relative concentrations of 37 resolved compounds were then examined using PCA, *k*-NN, and counterpropagation (CP) ANN, after applying autoscaling as data preprocessing. The PC1–PC2 score plot revealed that the grapefruit samples could be clearly distinguished from the other citrus fruit peel samples, which showed a quite large overlap and could not be separated from each other. The *k*-NN method revealed that only 5 of the 37 compounds show different distributions among the samples and 4 clear clusters of samples were obtained. CP-ANN was then used to discriminate the samples into these four classes, based on only the information from these five compounds. The method allowed indicating the main components responsible for the discrimination into the four classes.

5.4 Fingerprints Obtained Using Electrodriven Techniques

CE-MS fingerprints were recorded for six different *Corydalis* species, that is, *Corydalis cava* from two different regions (CM and CB), *Corydalis pumila* (PU), *Corydalis intermedia* (IN), and *Corydalis solida* from two different regions (SG and SE) (Sturm, Seger, and Stuppner, 2007). From the CE-MS fingerprints, $n = 39$ selected ion trace-derived peak areas were selected as variables, and normalized, and column centered before PCA and HCA. The goal of the study was to discriminate the six species.

The PC1–PC2–PC3 score plot allowed distinguishing the six *Corydalis* species samples. From plotting the loadings on the most important PCs as a function of the 39 variables, the compounds that allowed the differentiation between the 6 clusters of samples were indicated. An HCA dendrogram was then constructed using the squared Euclidean distance as the

(dis)similarity parameter and Ward's method as the linkage technique. The first split separates the *Corydalis cava* CM and CB samples from all other samples. Then, the *Corydalis pumila* (PU) samples are isolated from the remaining samples. The final split distinguishes the *Corydalis intermedia* (IN) samples from the *Corydalis solida* samples from two different regions (SG and SE). In conclusion, both PCA and HCA allowed a clear clustering of the six *Corydalis* species samples.

The earlier developed 60-min HPLC-DAD fingerprints of 39 *Mallotus* samples from various species (Tistaert *et al*., 2009) were transferred to *pCEC-UV* (Pieters *et al*., 2011). The goal of the study was to see whether HPLC fingerprints could be transferred to pCEC *fingerprints*. Secondly, it was evaluated whether these pCEC fingerprints allowed extracting the same information as the HPLC fingerprints, concerning components possibly responsible for an antioxidant activity. The transferred 60-min pCEC-UV fingerprints displayed characteristic chromatographic patterns of the different samples, as was also the case for the HPLC-DAD fingerprints.

Aligning corresponding peaks was again not done because of the large diversity of the samples. Preprocessing consisted of normalization followed by CC. In contrast to the PCA results from the HPLC-DAD fingerprints, the PC1–PC2 score plot from the pCEC-UV fingerprints did allow revealing two clear clusters, that is, one with the antioxidant active samples and the other with the nonactive samples (Figure 15).

To indicate compounds potentially responsible for the activity, PLS and OPLS were applied. The interpretability of the regression coefficients of the model was the most important evaluation criterion. Besides, the predictive ability of the different models was also determined. The number of components in the model and the RMSECV results of the PLS and OPLS are presented in Table 3. For all methods, an LOOCV procedure was applied. It could be decided that the RMSEP values from the pCEC fingerprints are higher than those from the HPLC fingerprints. The latter, thus, have a better predictive ability.

Nevertheless, the interpretability of the regression coefficients is more important. Again, peaks with negative regression coefficients are either antioxidant compounds or compounds for which the concentration is correlated with the antioxidant activity. It was found that both the PLS and the OPLS model from the pCEC fingerprints indicated a similar
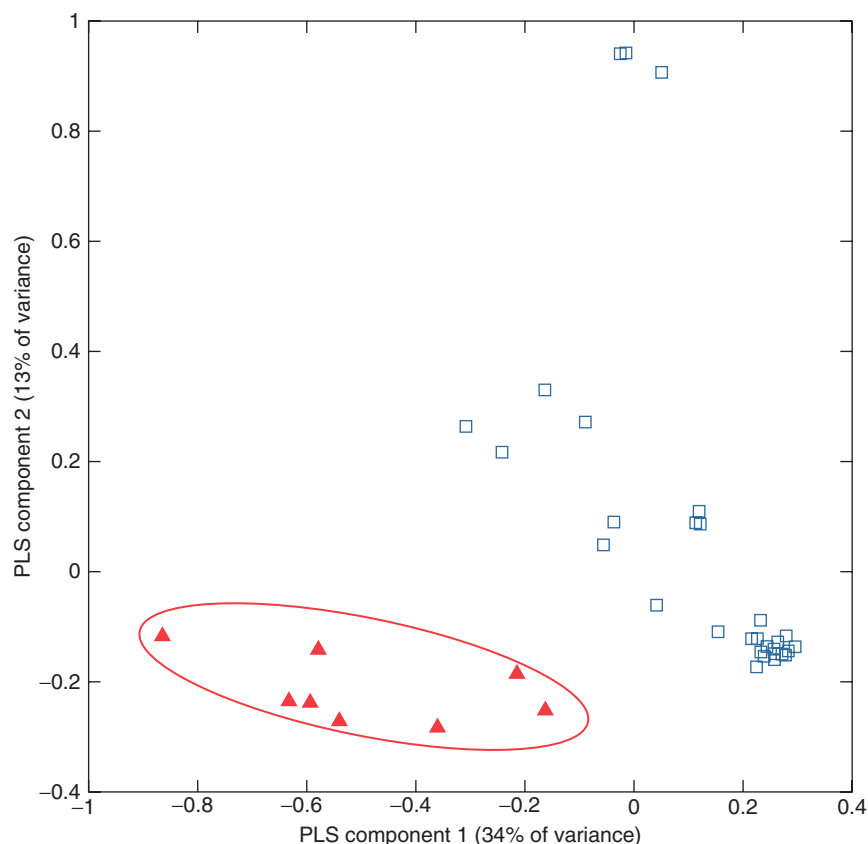
Figure 15 PC1–PC2 score plot drawn from the pCEC-UV fingerprints of 39 *Mallotus* samples from various species, after normalization and column centering as preprocessing. ▲, antioxidant active samples; □, nonantioxidant active samples. (*Source*: Reprinted from *Talanta*, **83**, Pieters, S., Tistaert, C., Alaerts, G., Bodzioch, K., Mangelings, D., Dejaegher, B., Rivière, C., Nguyen Hoai, N., Chau, Van M., Quetin-Leclerq, J., Vander Heyden, Y., Pressurized capillary electrochromatography in a screening for possible antioxidant molecules in Mallotus fingerprints: Challenges, potentials and prospects, 1188–1197, Copyright 2011, with permission from Elsevier.)

number of peaks potentially responsible for the antioxidant activity than the OPLS model from the HPLC-DAD fingerprints. Thus, pCEC could be used as an alternative technique for HPLC to develop fingerprints of herbal samples. However, pCEC does have some drawbacks and some precautions should be taken when using pCEC as the analytical technique to develop fingerprints for herbal samples.

developed with spectroscopic, mass spectrometric, chromatographic, or electrodriven analytical techniques.

Multivariate data analysis techniques proved to be useful when gathering information from herbal samples based on their fingerprints in various contexts. Depending on the goal of the study, different types of techniques should be applied.

6 CONCLUSIONS

This chapter mainly focused on the multivariate data analysis of fingerprints of herbal samples,

ACKNOWLEDGMENT

REFERENCES

Alaerts, G., Matthijs, N., Smeyers-Verbeke, J., *et al*. (2007) *J. Chromatogr. A*, **1172**, 1–8.

Alaerts, G., Dejaegher, B., Smeyers-Verbeke, J., *et al*. (2010) *Comb. Chem. High Throughput Screen.*, **13**, 900–922.

Arganda-Carreras, I., Sorzano, C. O. S., Marabini, R., *et al*. (2006) *Lect. Notes. Comput. Sci.*, **4241**, 85–95.

Broad, N., Graham, P., Hailey, P., Hardy, A., Holland, S., Hughes, S., Lee, D., Prebble, K., Salton, N., Warren, P. (2002) Guidelines for the development and validation of near Infrared (NIR) spectroscopic methods. In *Handbook of Vibrational Spectroscopy*. Chalmers, J. M. and Griffiths Pr (Eds.) John Wiley & Sons, Ltd, Chichester, UK.

Brown, S. D., Tauler, R., and Walczak, B. eds. (2009a) Comprehensive Chemometrics, Elsevier, Amsterdam, vol. **2**.

Brown, S. D., Tauler, R., and Walczak, B. eds. (2009b) Comprehensive Chemometrics, Elsevier, Amsterdam, vol. **3**.

Chinese State Food and Drug Administration (SFDA) (2002) *Requirements for Studying Fingerprints of Traditional Chinese Medicine Injection*, SFDA, Beijing, China, http://eng.sfda.gov.cn/(accessed 30 December 2012).

Chinese Pharmacopoeia Commission (2010) *Pharmacopoeia of the People's Republic of China*, People's Medical Publishing House, Beijing, China–English edition.

Daszykowski, M., Vander Heyden, Y., and Walczak, B. (2007) *J. Chromatogr. A*, **1176**, 12–18.

De'ath, G. (2002) *Ecology*, **83**, 1105–1117.

http://www.esapubs.org/archive/ecol/E083/017/, accessed 30 December 2012.

Dumarey, M., van Nederkassel, A. M., Deconinck, E., *et al*. (2008) *J. Chromatogr. A*, **1192**, 81–88.

Eilers, P. H. C., and Marx, B. D. (1996) *Statist. Sci.*, **11**, 89–121.

Eilers, P. H. C. (2004) *Anal. Chem.*, **76**, 404–411.

Eurachem (1998) *The Fitness for Purpose of Analytical Methods: A Laboratory Guide to Method Validation and Related Topics*, LGC (Teddington) Ltd, 1998, pp. 1–75.

European Medicines Agency (EMA) Committee for medicinal products for human use (CHMP) (2006) *Guideline on Quality of Herbal Medicinal Products/Traditional Herbal Medicinal Products, CPMP/QWP/2819/00 Rev 1, EMEA/CVMP/814/00 Rev 1*, EMA, London, UK, http://www.ema.europa.eu/ accessed 30 December 2012.

European Directorate for the Quality of Medicines & Healthcare: Strasbourg, France (2010) *European Pharmacopoeia Online 6*, http://online.edqm.eu/entry.htm accessed 30 December 2012.

Food and Drug Administration (FDA) (2004) Guidance for Industry: Botanical Drug Products, Food and Drug Administration. FDA, Silver Spring, Maryland, USA. http://www.fda.gov/, accessed 30 December 2012.

Frédérich, M., Wauters, J.-N., Tits, M., *et al*. (2011) *Planta Med.*, **77**, 81–86.

Gong, F., Wang, B. T., Chau, F. T., *et al*. (2005) *Anal. Lett.*, **38**, 2475–2492.

Grata, E., Boccard, J., Glauser, G., *et al*. (2007) *J. Sep. Sci.*, **30**, 2268–2278.

ICH Guidelines prepared within the International Conference on Harmonisation of Technical Requirements for the Registration of Pharmaceuticals for Human Use (ICH), Validation of Analytical Procedures: Text and Methodology, Q2(R1) (2005) 1–13, http://www.ich.org/, accessed 30 December 2012.

Hancock, T. (2006) Multivariate consensus trees: tree-based clustering and profiling for mixed data types. Ph.D. Thesis. James Cook University: Townsville, Australia.

Jumtee, K., Bamba, T., and Fukusaki, E. (2009) *J. Sep. Sci.*, **32**, 2296–2304.

Kassidas, A., MacGregor, J. F., and Taylor, P. A. (1998) *AIChE J.*, **44**, 864–875.

Kennard, R. W., and Stone, L. A. (1969) *Technometrics*, **11**, 137–148.

Komsta, Ł., Cieśla, Ł., Bogucka-Kocka, A., *et al*. (2011) *J. Chromatogr. A*, **1218**, 2820–2825.

Kong, W.-J., Zhao, Y.-L., Xiao, X.-H., *et al*. (2009) *Phytomedicine*, **16**, 950–959.

Lee, S. M., Kim, H.-J., and Jang, Y. P. (2012) *Phytochem. Anal.*, **23**, 508–512.

Lucio-Gutiérrez, J. R., Coello, J., and Maspoch, S. (2011) *Food Res. Int.*, **44**, 557–565.

Mao, J., and Xu, J. (2006) *Spectrochim. Acta Part A*, **65**, 497–500.

Møller, C., Hansen, S. H., and Cornett, C. (2009) *Phytochem. Anal.*, **20**, 231–239.

Ni, L.-J., Zhang, L.-G., Hou, J., *et al*. (2009) *J. Ethnopharmacol.*, **124**, 79–86.

Palacios-Morillo, A., Alcázar, Á., de Pablos, F., *et al*. (2013) *Spectroc. Acta Pt A – Molec. Biomolec. Spectr.*, **103**, 79–83.

Parastar, H., Jalili-Heravi, M., Sereshti, H., *et al*. (2012) *J. Chromatogr. A*, **1251**, 176–187.

Pieters, S., Tistaert, C., Alaerts, G., *et al*. (2011) *Talanta*, **83**, 1188–1197.

Pravdova, V., Walczak, B., and Massart, D. L. (2002) *Anal. Chim. Acta*, **456**, 77–92.

Snee, R. D. (1977) *Technometrics*, **19**, 415–428.

Sturm, S., Seger, C., and Stuppner, H. (2007) *J. Chromatogr. A*, **1159**, 42–50.

Tian, R.-T., Xie, P.-S., and Liu, H.-P. (2009) *J. Chromatogr. A*, **1216**, 2150–2155.

Tistaert, C., Dejaegher, B., Nguyen Hoai, N., *et al*. (2009) *Anal. Chim. Acta*, **652**, 189–197.

Tistaert, C., Dejaegher, B., and Vander, H. Y. (2011) *Anal. Chim. Acta*, **690**, 148–161.

Tistaert, C., Dejaegher, B., Nguyen Hoai, N., *et al*. (2012) *Anal. Chim. Acta*, **721**, 35–43.

United States Pharmacopoeia 33 – National Formulary 28 (2010) *United States Pharmacopoeia: Rockville*, Maryland, USA.

Vandeginste, B. G. M., Massart, D. L., Buydens, L. M. C., *et al*. (1998) *Handbook of Chemometrics and Qualimetrics. Part B*, Elsevier, Amsterdam.

van Nederkassel, A. M., Daszykowski, M., Massart, D. L., *et al*. (2005) *J. Chromatogr. A*, **1096**, 177–186.

van Nederkassel, A. M., Daszykowski, M., Eilers, P. H. C., *et al*. (2006) *J. Chromatogr. A*, **1118**, 199–210.

Vest Nielsen, N.-P., Carstensen, J. M., and Smedsgaard, J. (1998) *J. Chromatogr. A*, **805**, 17–35.

World Health Organization (WHO) (2000) *General Guidelines for Methodologies on Research and Evaluation of Traditional Medicines*, WHO, Geneva, Switzerland, http://www.who.int/ accessed 30 December 2012.

Zeng, S., Chen, T., Wang, L., *et al.* (2013) *J. Pharm. Biomed. Anal.*, **76**, 87–95.

Part Five

Drugs from Plants

In Silico-Guided Strategies for the Discovery and Rationalization of Bioactive Natural Products

Petra H. Pfisterer, Daniela Schuster, Judith M. Rollinger and Hermann Stuppner

Institute of Pharmacy/Pharmacognosy and Pharmaceutical Chemistry, and Center for Molecular Biosciences Innsbruck, University of Innsbruck, Innsbruck, Austria

1 INTRODUCTION

Historically, natural products (NPs) and their derivatives have been the major source of lead compounds in drug development. The isolation of the alkaloid morphine from opium by Friedrich Wilhelm Sertürner in 1804 initiated the use of NPs as single, pharmaceutical entities in the modern twentieth-century medicine. Further prominent discoveries, such as the accidental one of penicillin by Ian Fleming in 1928, stimulated the systematic search for naturally derived compounds. Even today, NPs play a dominant role for the development of drugs, which was demonstrated recently by Newman and Cragg. They circumstantiated that of all newly approved drugs in the time frame from January 1981 to December 2010, only 29% can be classified as being truly synthetic in origin (Newman and Cragg, 2012).

Despite the documented impact of NPs on drug discovery, many pharmaceutical companies have scaled down or eliminated their NP research in recent decade. The underlying reasons are, on the one hand, problems related directly to NPs themselves such as reliable access and supply of the natural source, small quantities of ingredients, complex structural characterization, synergistic activities, intellectual property concerns of local governments, and the Rio Convention on Biodiversity (Li and Vederas, 2009). On the other hand, pharmaceutical companies are confronted with great pressure to produce "blockbuster drugs" in order to overcome the relatively short patent protection time and the competition from generic drug manufacturers, which are generally not involved in drug discovery. High throughput screening and combinatorial chemistry were supposed to be promising strategies to accelerate the discovery process; however, the results are disappointing. In 2010, the US Food and Drug Administration (FDA) approved only 15 new molecular entities and six biologic license applications, which is the lowest number of newly approved small molecules since 1996 (Mullard, 2010).

In recent years, a new wave of interest in NPs can be observed (Rouhi, 2003). Advances in separation technologies and in the speed and sensitivity of structure elucidation overcome the practical difficulties of NP research. In addition, novel technologies are expected to provide discoveries accelerating the development of new pharmaceutical agents (Li and Vederas, 2009; Schuster and Wolber, 2010). One methodology that is more and more used is

the application of computational (virtual) screening methods. This technique offers the efficient, time- and cost-saving identification of unexploited, biologically active chemical scaffolds, and the rationalization of bioactivities on a molecular level (Newman, 2008).

This chapter aims to give an overview of virtual screening in the field of NPs. In addition, selected examples and application strategies are described in detail.

2 CHARACTERISTICS OF NATURAL PRODUCTS

Biological, ecological, and chemical rationales are the reasons why NPs are a reliable source of structures for the design of new drugs. Considering the immobility of parent organisms, NPs are produced as chemical defense or attack system. For the fulfillment of this role, they have been coevolved in a long selection process for optimal ligand functionality (Verdine, 1996). In addition, NPs are directly involved in the growth regulation and may therefore serve as templates for modulation processes (e.g., tumor growth). It is consequently not surprising that their influence as structural source is especially obvious in certain therapeutic areas such as the anti-infective and the anticancer ones (Newman and Cragg, 2012). However, their impact is also significant in therapeutic areas for which they might not seem relevant such as antidementia (e.g., galanthamine), lipid-control (e.g., lovastatin and analogs), and antiemetic (e.g., Δ-9-tetrahydrocannabinol) ones.

A further evolutionary success story of parent organisms is the use of multicomponent systems and their functional promiscuity to defend themselves against pathogenic attack. For a long time, this multitarget character of NPs has been associated with "dirty drugs" in contrast to high affinity, single-target drugs, particularly because the inhibition of a variety of targets is often combined with relatively weak inhibition patterns. However, the clinical outcome of recent two decades challenges this hypothesis: the specific binding to two or more molecular targets, the so-called polypharmacology, has attracted interest and its input in modulating metabolic and signaling pathways has already proved to be effective (Hopkins, 2008; Csermely, Agoston, and Pongor,

2005). Furthermore, the multitarget character offers the possibility not only to prevent salvage or escape mechanisms but also to reduce the range of side effects and toxicity, thereby increasing the likelihood of successful therapies (Korcsmaros *et al*., 2007). For instance, potential uses of multitarget drugs are chronic, refractory, and complex diseases such as cancer or acquired immune deficiency syndrome (AIDS) replacing therapeutic regimens. The application of herbal remedies and formulae, such as traditional medicines, which contain a plethora of secondary metabolites, is even more complex. Besides the pharmacological profile of each individual compound, interindividual effects may become noticeable (e.g., synergism, addition, and competition).

Besides the pharmacological characteristics of NPs, there are several studies in literature, which analyze their molecular structures from a chemoinformatics point of view (Henkel *et al*., 1999; Stahura *et al*., 2000; Lee and Schneider, 2001; Grabowski and Schneider, 2007; Koch *et al*., 2005). Feher and Schmidt compared the molecular properties between NPs, molecules from combinatorial synthesis, and drug molecules. This study revealed that the number of chiral centers, the presence of aromatic rings, the introduction of complex ring systems, the degree of the saturation of the molecules, and the number and ratios of different heteroatoms are the most important differences among the three groups (Feher and Schmidt, 2003). They concluded that the failure of combinatorial libraries to generate novel drug leads may be based on the restricted chemical space covered by the combinatorial synthetic molecules. Ertl, Roggo, and Schuffenhauer (2008) introduced a method to calculate an NP-likeness score, allowing the determination of how similar molecules are in comparison to the structural space covered by NPs. The aim of this score is the guidance of molecular design toward interesting regions of chemical space, which have been identified as "bioactive" by natural evolution. Recently, Rosén *et al*. (2009) used the chemical space navigation tool ChemGPS-NP to explore biologically relevant chemical space. Analyzing Euclidean distances, they divided the chemical space into smaller sections and neighborhoods revealing that almost all drugs have NP neighbors. The use of these near neighbors is supposed to be a good starting point for further drug discovery.

3 NATURAL PRODUCT DATABASES

A three-dimensional (3D) structure database containing energetically minimized chemical compounds is the basic requirement for virtual screening.

A comprehensive survey of NP databases, suppliers, and manufactures has been published by Füllbeck *et al*. (2006); an overview concerning literature, commercial, and focused NP databases has been given by Schuster and Wolber (2010). The most important and recently published NP databases are listed in the following sections.

3.1 Commercial NP Databases

There are many compendia of NP distributed commercially; however, the availability and supply of the substances are not necessarily contained. Some vendors offer additionally generated NP-based combinatorial structure libraries. They are delivered mainly on CD-ROM or DVD or are accessible via internet accounts for a fee (special academic prices are sometimes offered on request).

Examples for commercial providers of NPs including structure data files for virtual screening are AnalytiCon (MEGxm, MEGxp, and NATx libraries),[1] Specs,[2] BioFocusDPI,[3] GreenPharma,[4] InterBioScreen,[5] TimTec,[6] and Vitas-M.[7] The Dictionary of Natural Product Database (DNP, more than 243,000 entries) is launched by Chapman & Hall and is updated every 6 months with current literature data (Buckingham, 2007).

3.2 NP Databases from Literature Sources

NP databases are often described by NP research groups who collect chemical structure data from literature sources and transform them into a 3D multiconformational database.

- The SuperNatural database (approximately 50,000 entries) contains NPs that are available from different suppliers; it is updated twice a year (Dunkel *et al*., 2006).
- The Natural Product Database (NPD, 146,790 entries) covers compounds from a broad panel of natural sources, including plants, fungi, marine organisms, and animals (Rollinger *et al*., 2004). As compounds of rare sources are also included, the access to a substance identified as virtual hit may be difficult.

Therefore, many NP databases focus on locally available material or specific parent organisms offering the advantages to have a facilitated access to the material. Furthermore, the use of ethnopharmacologically biased selections of database entries enables a focused investigation and rationalization of their traditional use by *in silico* techniques. A considerable increase of the success rate has been described (Rollinger *et al*., 2004).

- The DIOS database (9676 entries) consists of small molecular weight NPs from phytochemical reports of anti-inflammatory medicinal plants by the ethnopharmacological source "*De materia medica*" witnessed by Pedanius Dioscorides (First century A.D.) (Rollinger *et al*., 2004).
- Information of approximately 400 plants and their chemical constituents traditionally used in Austria has been collected by the University of Vienna, Institute of Pharmacognosy in the Austrian traditional herbal medicines database (VOLKSMED database, 2826 entries) (Gerlach, 2007).
- In recent decades, Traditional Chinese Medicine (TCM) has attracted much attention. On the basis of literature data, the Chinese Herbal Medicines Database (CHMD, 10,216 entries) (Fakhrudin *et al*., 2010), the Traditional Chinese Medicines Database (TCMD, 10,458 entries) (Wang *et al*., 2007), and the Chinese Herbal Constituents Database (CHCD, approximately 700 compounds) (Ehrman, Barlow, and Hylands, 2007a) have been published.
- The Marine Natural Product Database (MNPD, approximately 6000 entries) covers chemicals from marine-derived materials (Lei and Zhou, 2002). For drug discovery, they are highly interesting as they are a source of a large number of novel, diverse NP structures (Haefner, 2003). In addition, their chemical defense system usually dilutes rapidly in water and therefore is very potent.

Besides these NP databases, others contain entries of compounds that are directly linked to bioactivity or have been used as drugs.

- The Bioactive Plant Compound Database (BPCD, 2597 entries) contains information on the molecular targets of each compound, their ligand type, inhibition data (IC$_{50}$, K_d, K_i values), details on bioactivity, and the name(s) of the species in which the compound has been detected so far (Ehrman, Barlow, and Hylands, 2007a).

- The National Cancer Institute database (NCI, 260,071 entries) consists of natural and synthetic compounds, which have been evaluated for their anticancer activity.[8] NPs, such as bile acids, steroids, and flavonoids, are included.

- The recently published Benzylisoquinoline Alkaloids database (BIAdb, 846 entries) contains information about benzylisoquinoline alkaloids with entries in term of their source and type of function (Singla *et al*., 2010). It is tightly integrated with Drugpedia and available online.

- The Derwent World Drug Index (WDI, 80,000 entries in its current version) is often used by researchers for the quality assessment of theoretical models, as the compounds have been used as drugs or are reported to be biologically active.[9] NPs listed in this database are, for example, penicillin, testosterone, cortisone, and callyspongenols A–C.

3.3 Focused NP Databases

In literature, the generation of small, project-centered NP databases is also described. These focused databases are especially valuable, if not all the constituents of the investigated material are covered by large databases (prospective application). For example, Rollinger *et al*. (2009) generated a focused virtual library of the main constituents from *Ruta graveolens* L. (Rutaceae) in order to apply pharmacophore-based target fishing on this dataset. The group of Lee merged two libraries (BIOMOL International LP and Indofine Chemical Company) into a new NP database with the aim of identifying novel inhibitors of *Escherichia coli* β-ketoacyl-acyl carrier protein synthase (KAS) III (Lee *et al*., 2009b).

Alternatively, a focused database of the isolated and already biologically tested compounds offers the possibility to rationalize NP activity or identify the mechanisms of action for known biological effects of complex remedies (retrospective application). Recently, Schuster *et al*. (2010a) compiled the Prasaplai database, which was used to elucidate one mechanism of action of Prasaplai (Thai traditional medicine composed of 10 plants, sodium chloride, and camphor). The structural information on NPs from the plants present in Prasaplai was collected from literature. Virtual screening using five previously validated pharmacophore models for cyclooxygenase inhibitors (Schuster *et al*., 2010a) led to a list of Prasaplai constituents, which could act via this target and can thus be responsible for the anti-inflammatory and antiprostaglandin activities of this herbal remedy. Because 50% of the retrieved hits were already known cyclooxygenase inhibitors, it would be promising to explore the remaining, so far untested virtual hits for their biological activity (Waltenberger *et al*., 2011).

4   REPRESENTATION OF MOLECULES

In the virtual screening process, the biological activity of a compound cannot directly be calculated from the molecular structure. Therefore, the structure has first to be represented by mathematical structure descriptors. These are then used for modeling the biological activity, whereas the constitution, the 3D structure, and the molecular surface represent various degrees of sophistication for the structure representation. In order to represent all the investigated objects of a data set by the same number of descriptors, a mathematical transformation has to be performed. In addition, the structure representations are combined with calculated physicochemical properties of the atoms and bonds and are submitted to autocorrelation or atom radial distribution functions (Gasteiger *et al*., 2003). A comprehensive overview of structure representations (1D, 2D, and 3D descriptors) needed for solving the different problems encountered in drug design has been published by Gasteiger (2006).

The most basic level is associated with the 1D representation, such as the molecular weight or the log *P* value. To identify a molecule or to rapidly search for the presence or absence of certain predefined substructures or fragments, structural fingerprints (1D descriptor) can be used. The identification of the individual substructure is, however, not possible with this method.

The next level is the 2D representation, which is associated with the connectivity of the molecules

but not with stereochemistry. Fragment codes (2D descriptor) can retain the identity and quantity of present chemical substructures or fragments. They are mainly used for defining the similarity of structures by calculating the Tanimoto index. Topological autocorrelations (2D descriptor), which are applied for finding new lead structures, lead hopping, and for the comparison of libraries, consider the constitution of a molecule, its set of atoms, and how they are connected.

The highest accuracy level is the 3D representation, which conveys the conformation, the chirality of stereogenic centers, and the molecular surface properties. This can be achieved by the use of automatic 3D generators, radial distribution functions, and autocorrelation of the molecular electrostatic potential (3D descriptors). The most challenging representation is the molecular flexibility (3D descriptors) yielded by the rotation of single bonds. Ideally, the calculated 3D structure of the molecule should represent its bioactive conformer, which is the conformation of biological relevance (Schwab, 2010). As this bioactive conformer is not necessarily related to the global energy minimum conformer, it cannot be obtained by simply calculating the conformer with the lowest strain energy. However, the generation of a limited, but quite diverse set of energetically minimized conformations has proved to be a good representation for the bioactive conformer (Kirchmair *et al.*, 2005, 2006; Boström, 2001). Prominent algorithms for the calculation of conformational models are OpenEye's OMEGA,[10] Molecular Networks' ROTATE,[11] Schrodinger's ConfGen (Watts *et al.*, 2010), and Accelrys' ConFirm.[12]

In order to store molecules in suitable formatted files, a number of methods have been developed (Martinez-Mayorga and Medina-Franco, 2009). For example, molecules can be represented for 2D representations by fingerprints, SMILES, SMARTS, or InChI strings using a small amount of memory; for 3D representations by .sdf, .pdb, .mol, or .mol2, the formats contain additionally the 3D coordinates of the structures.

5   VIRTUAL SCREENING METHODS

Advances in computer processing capabilities and hardware solutions facilitated the application of computational (*in silico*) methods, especially virtual screening, in drug development. As virtual screening improves the time and cost efficiency of lead compound discovery (prospective application) as well as enables the rationalization of bioactivity on a molecular level (retrospective application), a steadily growing number of publications on this subject could be observed in recent decade (Rester, 2008). By definition of the International Union of Pure and Applied Chemistry (IUPAC), virtual screening is the "selection of compounds by evaluating their desirability in a computational model" (Maclean *et al.*, 1999). In contrast to experimental high throughput screening, this computational selection process is a knowledge-driven approach, which is used to identify the most promising candidates out of large databases for a focused testing in a relevant assay. For example, only 0.02% active hits (out of 400,000 compounds) have been discovered by a random high throughput screen for protein tyrosine phosphatase 1b inhibitors, whereas a compound selection of 365 substances assembled by virtual screening revealed 34.8% novel inhibitors (Doman *et al.*, 2002). It is noteworthy that such a high success rate of a virtual screening experiment is a best-case scenario. Usually, positive hit rates between 5% and 20% active compounds can be expected from properly conducted studies. Virtual screening is generally associated with a data mining or filtering process, reducing the number of *in vitro* test compounds to a promising fraction of all available substances. In general, there are two major concepts of virtual screening, which are described in the following sections in detail: target-based virtual screening and ligand-based virtual screening (Rester, 2008).

Virtual screening of a database with a data mining tool as search query results in a list of potential hits, the so-called virtual hit list. As different virtual screening tools can produce different virtual hit lists, the combination of various methods (e.g., ligand- and structure-based virtual screenings) can offer a complementary discovery of activities (Schuster *et al.*, 2010b). Usually, virtual screening studies begin with the least-demanding approaches and progress to the more accurate, time-consuming techniques in the last stages, where fewer compounds remain to be investigated (Talevi, Gavernet, and Bruno-Blanch, 2009).

5.1 Target-Based Virtual Screening

Target-based virtual screening methods rely on the spatial information of the target protein and its binding site obtained from X-ray crystallography or multidimensional nuclear magnetic resonance (NMR) spectroscopy. In addition, theoretical models of targets, the so-called homology models, can be used. The largest public repository of pharmacologically relevant macromolecules, many of them with cocrystallized ligands, is the Protein Data Bank (PDB, www.pdb.org) (Berman *et al.*, 2000). Currently, this database contains approximately 98,000 entries covering polypeptides, enzymes, receptors, and nucleic acids.

There are a lot of established and potential drug targets for which no experimental structural information is available. In this case, homology modeling can be used to construct a model of the target protein (Vallat, Pillardy, and Elber, 2008). On the basis of the assumption that important regions such as binding sites remain conserved in a protein family, related homologous proteins, for which 3D structures are available, are used as templates. The success of these models depends consequently on the sequence similarity between the target protein and the related homologous protein. The accuracy of a homology model has a profound influence on the outcome of virtual screening experiments relying on it, so its quality should be critically addressed before employing the model in a study.

5.1.1 Docking

The most established structure-based modeling technique is docking, as only the knowledge of the 3D structure of the target and the location of the binding site is necessary. Docking algorithms place molecules within the binding site of the target of interest and rank them according to their calculated binding affinity estimations (scoring). A large number of different programs and algorithms are available, whereas the most established ones are incremental ligand fragmentation and reconstruction (e.g., FlexX; Rarey *et al.*, 1996), volume- or shape-based algorithms (e.g., DOCK; Ewing *et al.*, 2001), genetic algorithms (e.g., GOLD; Jones *et al.*, 1997 and AutoDOCK; Morris *et al.*, 1998), systematic search (e.g., Glide; Halgren *et al.*, 2004), Monte Carlo

approaches (e.g., LigandFit; Venkatachalam *et al.*, 2003), and surface-based molecular similarity methods (e.g., SurflexDock; Jain, 2003). While docking has proved to be a valuable tool in drug development, the inaccuracy of the functions that are used to estimate the binding affinity between the ligand and the target, the so-called scoring functions, still remain an often-discussed topic. An overview of the very active and rapidly advancing area of research concerning these scoring functions is given by Rester (2008).

In literature, there are many docking studies that report the discovery of novel NPs as lead compounds (prospective application). For example, Nikolovska-Coleska *et al.* (2004) virtually screened an in-house 3D database containing 8221 small organic molecules with diverse chemical structures isolated from TCM herbs to identify small molecular weight X-linked inhibitor of apoptosis proteins (XIAP) baculovirus inhibitor of apoptosis proteins repeat (BIR) 3 inhibitors. The XIAP BIR3 domain is a promising target for the design of substances that sensitize chemoresistant cancer cells toward the treatment of conventional chemotherapeutics (chemosensitizers). The virtual hit embelin (**1**), originally isolated from the Japanese Ardisia herb (*Ardisia japonica*, Myrsinaceae), is slightly less potent than the natural ligand second mitochondria-derived activator of caspases (Smac). Recently, *in vivo* studies suggested embelin (**1**) for the treatment of hormone refractory prostate and colon cancer (Danquah *et al.*, 2009; Dai *et al.*, 2009). In addition, derivatives of embelin (**1**) have been designed showing a higher affinity for XIAP BIR3 (Chen *et al.*, 2006).

Celastrol (**2**), a triterpene from *Tripterygium wilfordii* (Celastraceae), has been known to regulate the heat shock protein (Hsp) 90 pathway (Westerheide *et al.*, 2004; Hieronymus *et al.*, 2006). Regulating the stability, folding, and activation state of a variety of oncogenic "client" proteins, the molecular chaperone Hsp90 is an attractive anticancer target. Besides the development of inhibitors targeting the ADP/ATP-binding site, a novel approach is the design of agents that block the interactions between Hsp90 and the cell division cycle 37 homolog (Cdc37). As Cdc37 recruits protein kinases to the Hsp90 machinery, the inhibition of the Hsp90–Cdc37 interaction would have therapeutic potential in kinase-dependent cancers (Pearl, Prodromou, and Workman, 2008). Using a combination of molecular docking, simulated annealing,

and molecular dynamics simulations, Zhang *et al.* (2008) suggested that celastrol (**2**) induces a major conformational change in the binding pocket that blocks an essential H-bond interaction between Arg167 of Cdc37 and Glu33 of Hsp90. Experimental evaluation efforts confirmed that celastrol disrupts the Hsp90–Cdc37 interaction in the superchaperone complex to exhibit antitumor activity *in vitro* and *in vivo*.

Miguet *et al.* (2006) employed homology modeling to build a 3D structure of 11β-hydroxysteroid dehydrogenase (HSD) type 1, which plays an important role in many human disorders such as metabolic syndrome, diabetes type 2, and obesity. This model was found to reproduce closely the crystal structure that became available in the final stages of the study. Screening a database of 114,000 natural molecules with both the model and the crystal structure as search queries revealed virtual hits that have already been shown to be potent inhibitors of 11β-HSD1.

Besides the use of docking as predictive tool, it is often applied to rationalize an already known bioactivity (retrospective application). In the search for novel antiviral drug leads, Grienke *et al.* (2010) have recently performed phytochemical investigations on the seed extract of *Alpinia katsumadai* (Zingiberaceae). Diarylheptanoids showed *in vitro* neuraminidase (NA) inhibitory activities in low micromolar ranges against human influenza virus A/PR/8/34 of subtype H1N1, whereas the most promising constituent, katsumadain A (**3**), also inhibited the NA of H1N1 swine influenza viruses. Katsumadain A (**3**) differing from known active NA inhibitors in terms of scaffold, shape, and size represented an interesting molecule for the exploration of the underlying binding mode. Considering the flexible loop regions of NA, molecular dynamics simulations and docking studies revealed well-established interactions of katsumadain A (**3**) and the protein, surface complementarity to the simulated binding pocket, and concordance with experimentally derived structure–activity relationship data.

Further retrospective docking studies have been performed in the field of antiasthma (Gouda *et al.*, 2009), anti-inflammatory (Avila *et al.*, 2009), and anticancer lead discovery (Lin *et al.*, 2008). In the latter study, hematoxylin (**4**), a well-known natural dye isolated from the heartwood of *Haematoxylon campechianum* (Leguminosae), was revealed to be one of the most potent inhibitors of cellular-Sarcoma

(c-Src) tyrosine kinase with an IC_{50} value of 440 n*M*. Additional experimental evaluations elucidated its multitarget character being a promising broad-spectrum protein tyrosine kinase inhibitor at both the enzymatic and the cellular levels. Docking studies provided insight into the interaction of hematoxylin (**4**) and the ATP-binding site of c-Src, indicating an ATP competitive mechanism that could be confirmed by enzyme kinetic studies.

5.1.2 Structure-Based Pharmacophore Modeling

A pharmacophore is the ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target structure and to trigger or block its biological response (Wermuth *et al.*, 1998). A structure-based 3D pharmacophore model allows for the direct investigation of key interactions between the ligand and the binding site of the protein on a generalized level. Not only the 3D structure of the target, but also the knowledge of the bioactive conformer of the ligand and its binding mode to the target is required for the generation. Software programs such as the program MOE by the Chemical Computing Group (www.chemcomp.com), Accelrys' Discovery Studio (www.accelrys.com), or Schrodinger's Phase (www.schrodinger.com) enable the manual construction of structure-based pharmacophore models. Inte:Ligand's program LigandScout (www.inteligand.com) allows for the automated perception of structure-based pharmacophores and supports the comparison, interpolation, and merging of several structure-based pharmacophores using a novel 3D alignment algorithm based on pattern-matching techniques (Wolber and Langer, 2005; Wolber, Dornhofer, and Langer, 2006).

The key interactions between the ligand and the binding site, which are thought to determine the bioactivity, are represented as points surrounded by a sphere of tolerance, the so-called features. As different structural motifs can express the similar chemical behavior and therefore the same biological effect, the features indicate only the region in space that should be occupied by a certain type of interaction such as H-bonding, electrostatic interaction,

or hydrophobic interaction. In addition, forbidden spheres, the so-called exclusion volume spheres, can be used to indicate sterically unfavorable regions for a mapped ligand conformation. As the pharmacophore model is only a binding model for the one ligand that is present in the structure, it has to be optimized. In an ideal case, different pharmacophore models, which are based on different ligands in complex with the same binding site of the macromolecule, can be brought to a consensus model. Alternatively, the pharmacophore model can be adjusted until other known bioactive ligands are recognized correctly.

A pharmacophore model should be able to identify as many biologically active molecules and to discard as many inactive ones from a structurally diverse database as possible. Therefore, it has to be validated carefully by screening a proper set of active and inactive molecules. Quantitative measures for its quality are the yield of actives, the enrichment factor, the sensitivity, and the correct assignment of inactive compounds (Schuster and Wolber, 2010). Besides this internal validation, also an external validation should be conducted by evaluating experimentally predicted compounds.

Structure-based pharmacophore models are mainly used as prospective screening tool in the early phase of drug discovery. Successful application examples within NPs have recently been published in the field of antibiotic treatment. Lee *et al.* (2009a) generated structure-based pharmacophore models for KAS III and validated their accuracy and specificity. KAS III initiates the bacterial fatty acid biosynthesis making it one of the most promising targets for the discovery of antibiotics with broad-spectrum activity. In a following study, the same group virtually screened an NP database consisting of 865 natural and flavonoid compounds (Lee *et al.*, 2009b). Four virtual hits, a flavone, two isoflavones, and a chalcone, were selected for experimental evaluation. According to NMR spectroscopy and fluorescence experiments, all four compounds bind to *E. coli* KAS III. One of the flavones, 3,6-dihydroxyflavone (**5**), showed antimicrobial activities against Gram-positive bacteria such as *Staphylococcus aureus* and *Enterococcus faecalis*, whereas the chalcone phloretin (**6**) displayed broad-spectrum activity against Gram-negative and Gram-positive bacteria.

An alternative strategy for the design of antibiotics is the prevention of virulence by inhibition of the bacterial cell-to-cell communication using the quorum-sensing inhibitor of RNAIII-inhibiting peptide (RIP). Hamamelitannin (**7**), isolated from the bark of *Hamamelis virginiana* (white hazel, Hamamelidaceae), emerged as a virtual hit when screening a database of commercially available small-molecule compounds with a RIP-based pharmacophore model as search query (Kiran *et al.*, 2008). It was found to inhibit the quorum-sensing regulator RNAIII *in vitro* and to prevent infections caused by methicillin-resistant *S. aureus* and *Staphylococcus epidermalis* strains *in vivo*.

Another field in which structure-based pharmacophore models were able to identify novel bioactive NPs is oncolytic research. Acylated flavonol monorhamnosides from *Eriobotrya japonica* Lindl. (Rosaceae) were virtually predicted as ligands of the XIAP BIR3 domain, a target to design chemosensitizers. The main constituent and virtual hit, kaempferol-3-O-α-L-(2″,4″-di-E-*p*-coumaroyl)-rhamnoside, was identified as XIAP BIR3 ligand with a dose-dependent affinity with an IC_{50} of 10.4 μM. Further, this compound induced apoptosis in XIAP-overexpressing Jurkat cells and activated caspase-9 in combination with etoposide (Pfisterer *et al.*, 2011).

Within the area of metabolic and anti-inflammatory compounds, structure-based models for the farnesoid X receptor (Schuster *et al.*, 2011) pointed toward triterpenes from *Ganoderma lucidum* as bioactive compounds. Five lanostanes (ergosterol peroxide, lucidumol A, ganoderic acid TR, ganodermanotriol, and ganoderiol F) out of 25 tested, isolated *Ganoderma* constituents dose-dependently induced FXR in low micromolar concentrations in a gene reporter assay (Figure 1) (Grienke *et al.*, 2011).

5.2   Ligand-Based Virtual Screening

Lead discovery projects are often started before detailed structural data of the protein target and the biological conformation of the ligand become available. Therefore, ligand-based virtual screening methods are advantageous as they rely on the assumption, the so-called similar property principle, that compounds with similar structures tend to express similar biological activity (Johnson and Maggoria, 1990). However, the aim is mainly the discovery of structurally diverse compounds having similar

Elucidation of protein–ligand interactions
between FXR and its agonist

Structure-based pharmacophore
model for FXR agonist

Virtual screening of natural products databases

OH

OH OH

O

H

Virtual hit
ganodermanotriol

In vitro
assay

EC$_{50}$
2.5 μ*M*

Dose-dependent
FXR induction
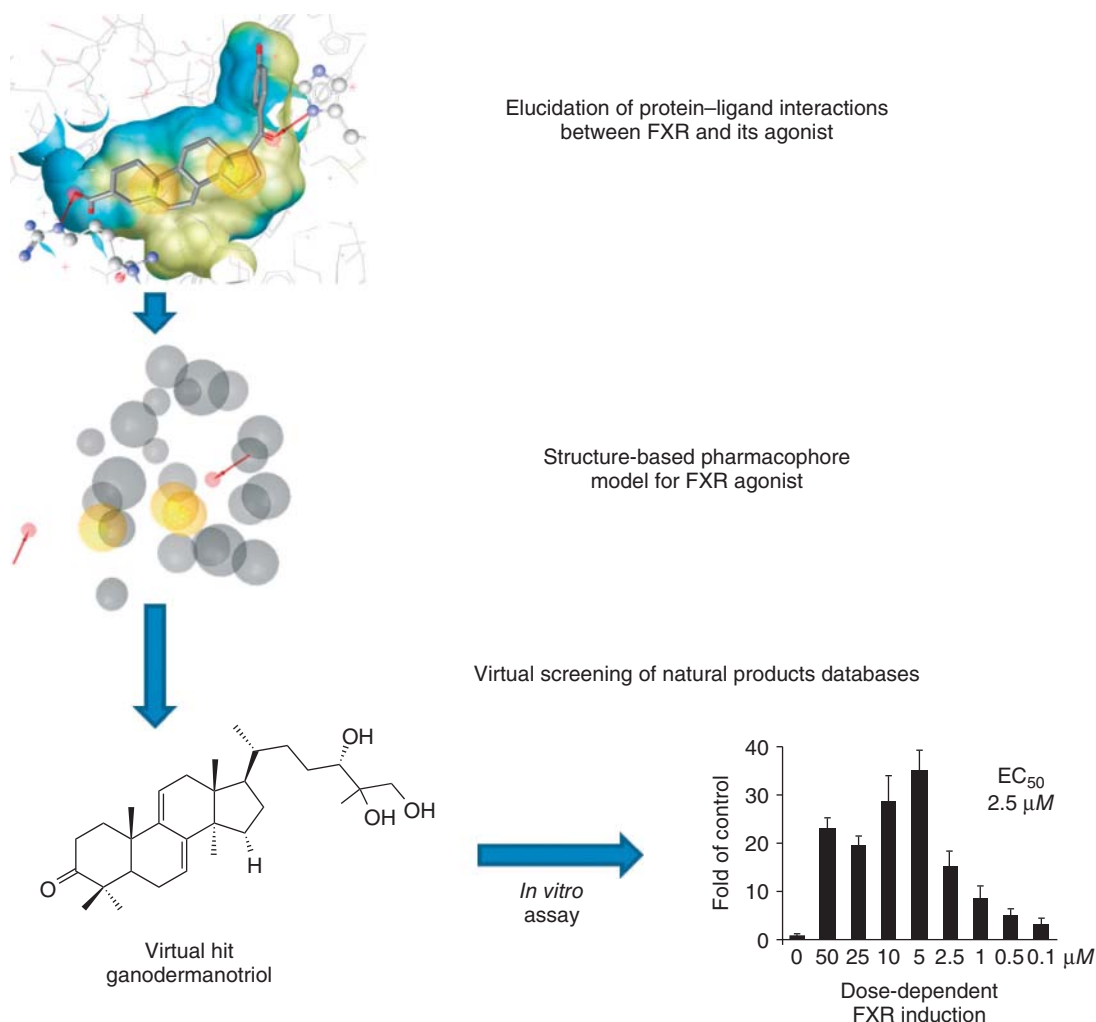
Figure 1 Example for a successful structure-based virtual screening study leading toward novel bioactive natural products. An X-ray crystal structure of FXR in complex with a steroidal agonist served as template for a structure-based pharmacophore model representing chemical features involved in protein–ligand interactions. Chemical features are color coded: yellow – hydrophobic, red arrow – hydrogen bond acceptor, and gray – exclusion volume sphere (forbidden area for the ligand). Virtual screening of natural compound databases suggested lanostane-type triterpenes from *Ganoderma lucidum* as potential FXR ligands. Subsequent isolation and *in vitro* studies confirmed FXR activity for 5 out of 25 tested *Ganoderma triterpenes*, exemplified by ganodermanotriol (Grienke *et al*., 2011).

bioactivities as the compounds used as input information (scaffold hopping).

Several ligand-based modeling and screening methods have been described in literature, among which similarity analysis, quantitative structure–activity relationships (QSARs), neural networks, and pharmacophore modeling are the most popular ones (Reddy *et al*., 2007).

5.2.1 Similarity Analysis

When only little information about the target and a few active compounds is available, a similarity search can stand at the beginning of a drug discovery project. First, a known bioactive molecule, the target structure, is identified. Using an objective measurement

of similarity, the structure of the bioactive molecule is then compared with each of the comparison structures present in a database. Finally, a similarity coefficient quantifies the degree of resemblance between the representations of the target and the comparison structures. Most molecular modeling software packages have similarity search algorithms included, for example, Schrodinger (www.schrodinger.com), MOE (www.chemcomp.com), or Accelrys Discovery Studio (www.accelrys.com). In addition, chemical literature databases such as SciFinder (www.cas.org) and the ChEMBL database (www.ebi.ac.uk/chembl) have integrated similarity search functions to explore compounds and information on their structurally close neighbors.

The 5-lipoxygenase (LO) pathway has been associated not only with inflammation but also with atherosclerosis, cancer, and osteoporosis. Using 5-LO pathway inhibitors as templates, Franke *et al.* (2007) virtually screened the MEGx and NATx compound libraries of AnalytiCon with a topological pharmacophore descriptor. Virtual hits were tested in intact polymorphonuclear leukocytes. The two most potent substances (IC_{50} values of $0.8\,\mu M$ and $0.9\,\mu M$, respectively) have different scaffolds in respect to their corresponding templates indicating successful scaffold hopping. In a second round, these molecules were used themselves as queries using different ligand-based virtual screening methods in order to retrieve potentially more potent inhibitors and to obtain preliminary structure–activity relationship (SAR) models. The screening of a focused NP-derived combinatorial library produced candidates, which potently suppressed 5-LO activity in intact cells representing a novel class of 5-LO inhibitors.

As S-adenosylhomocysteine hydrolase (SAH) plays an important role in the biological transmethylation reaction, which is a key step in the duplication of virus life, it has become an attractive target for the design of antiviral agents. Wei *et al.* (2007) performed a similarity search in the TCMD using 2′,3′-dihydroxycyclopenten-4′-yl-adenine, the cocrystallized ligand of SAH (PDB entry 1a7a), as the template. The fingerprints of the template molecule and the compounds present in the database were calculated and compared by their Tanimoto similarity coefficient. Seventeen hits were retrieved with high similarity, which were additionally docked

in the binding site of SAH. Unfortunately, they have not been evaluated experimentally.

5.2.2 QSAR

QSAR studies are based on the concept that a property such as biological activity depends on the molecular structure and that it is possible to find a quantitative relationship between this property and the suitable molecular representation (Martinez-Mayorga and Medina-Franco, 2009).

On the one hand, the aim of QSAR studies is the explanation of the property based on the molecular structure (retrospective application). For instance, in the search for anticonvulsant candidates, Gavernet *et al.* (2008) virtually screened an NP database using a combined virtual screening 2D and 3D QSAR methodologies. One of the four selected compounds, 2-(4,6-dimethyl-1-benzofuran-3-yl)acetic acid (**8**), showed anticonvulsant activity in mice at 30 and $100\,mg\,kg^{-1}$ and the absence of neurotoxicity at the tested doses.

On the other hand, the aim of QSAR studies is the development of models that are able to quantitatively predict the property of new molecules (prospective application). Phenylmethylene hydantoin (**9**), isolated from the red sea sponge *Hemimycale arabica*, showed potent *in vitro* antigrowth and anti-invasive properties against PC-3M prostate cancer cells as well as anti-invasive activities in orthotopic xenograft and transgenic mice models. Mudit *et al.* (2009) synthesized derivatives of this NP and evaluated their anti-invasive activities in order to establish a predictive 3D QSAR model. A similar study has been performed by Salum, Dias, and Andricopulo (2009), who employed a fragment-based method to develop QSAR models for a series of synthetic discodermolide analogs. Discodermolide (**10**) has the same mechanism of action similar to the taxanes, but demonstrates superior antitumor efficacy. Their aim was the understanding of drug–receptor interactions on a molecular level.

5.2.3 Neural Networks

On the basis of a simplified concept of the information processing in the human brain, artificial neural networks are powerful data mining tools with a

wide range of applications in drug design. The validated models may be used for establishing QSARs of chemically related structures or for handling compounds having multicategory activities. One of the most important advantages of neural networks is the possibility to discover novel leads without being restricted to only one binding mode.

One may distinguish between supervised and unsupervised learning methods. Concerning supervised learning methods, a series of training objects characterized by their descriptors and their target values of physical, chemical, or biological data is fed into the neural network, which initially contains randomly assigned weights. In order to minimize the output error, the adjustment of the weights is iteratively repeated until the output error has fallen below a predefined threshold (Gasteiger *et al*., 2003). A trained neural network is able to calculate the output value of a new input object, and consequently to distinguish between hits and nonhits. Concerning unsupervised learning, the neural network projects object from a high dimensional space into a lower dimensionality space, usually into a 2D plane, allowing an unbiased view on the data (Gasteiger *et al*., 2003). One such method is the self-organizing neural network, which was introduced by Kohonen (1989). The generation of maps and thereby the visualization of the relationships between the objects is one of the most important advantages of this method.

Schuster *et al*. (2010b) applied two neural network-based virtual screening techniques (novelty detection and counterpropagation neural networks) to explore the NP space for acetylcholinesterase (AChE) inhibitors. This enzyme is involved in accelerating the assembly of β-amyloid into amyloid fibrils that are characteristically found in the brain cells of Alzheimer's patients. Screening the NP database, DIOS revealed virtual hits that were further prioritized by docking. Forsythoside A (**11**) and (+)-sesamolin (**12**) were selected for the experimental evaluation showing dose-dependent and long-lasting inhibitory effects on AChE.

Sesquiterpene lactones possess a variety of biological activities, including the inhibition of the nuclear factor κ-light-chain-enhancer of activated B-cells (NF-κB) and the inhibition of the release of serotonin. Wagner *et al*. (2008) used an artificial neural network to develop a QSAR model that predicts the serotonin release inhibitory activity. In addi-

tion, the comparison of these descriptors with previously published ones used in an NF-κB inhibition model (Wagner *et al*., 2006), provided information on the structural requirements of sesquiterpene lactones for the inhibition of serotonin release in contrast to NF-κB inhibition.

5.2.4 Ligand-Based Pharmacophore Modeling

On the basis of the alignment of the conformational models from known active compounds, a molecular superimposition algorithm arranges the 3D structures of known active compounds in a way that equal chemical functionalities are located in similar positions. Importantly, the user must be aware that the two or more active compounds have the same binding mechanism and act at the same binding site of the same target. Pharmacophore features are then placed on the positions where all compounds share the corresponding chemical functionality. Similar to the structure-based pharmacophore model, it is also possible to include information of inactive ligands by adding exclusion volume spheres or shapes. Before its use as search query, also the ligand-based pharmacophore model should be carefully validated.

For the automated elucidation of common features, the commercial software packages Catalyst/DiscoveryStudio,[13] MOE,[14] and Phase (Dixon *et al*., 2006) employ graph-based distance geometry approaches. For an improved speed and accuracy, the novel programs MOGA (Cottrell, Gillet, and Taylor, 2006) and LigandScout/Espresso (Wolber, Dornhofer, and Langer, 2006) use multiobjective optimization techniques and pattern-matching 3D alignment algorithm, respectively.

Laggner *et al*. (2005) generated a ligand-based pharmacophore model of the σ_1 receptor, which has been reported to mediate a plethora of pharmacological effects, including schizophrenia, depression, ischemia, and cognition disorders. Screening the WDI database, steroidal alkaloids emerged as virtual hits. Tomatidine (**13**, from tomatoes), solasodine (**14**), and solanidine (**15**; both from various *Solanum* species) have been successfully identified as high affinity ligands with K_i values of 81, 207, and 74 n*M*, respectively.

Embelin (**1**)

Celastrol (**2**)

Katsumadain A (**3**)

Hematoxylin (**4**)

3,6-dihydroxyflavone (**5**)

Phloretin (**6**)

Hamamelitannin (**7**)

2-(4,6-dimethyl-1-benzofuran-
3-yl)acetic acid (**8**)

Phenylmethylene hydantoin (**9**)

Discodermolide (**10**)

Forsythoside A (**11**)

(+)-Sesamolin (**12**)

Tomatidine (**13**)

Solasodine (**14**)

Solanidine (**15**)
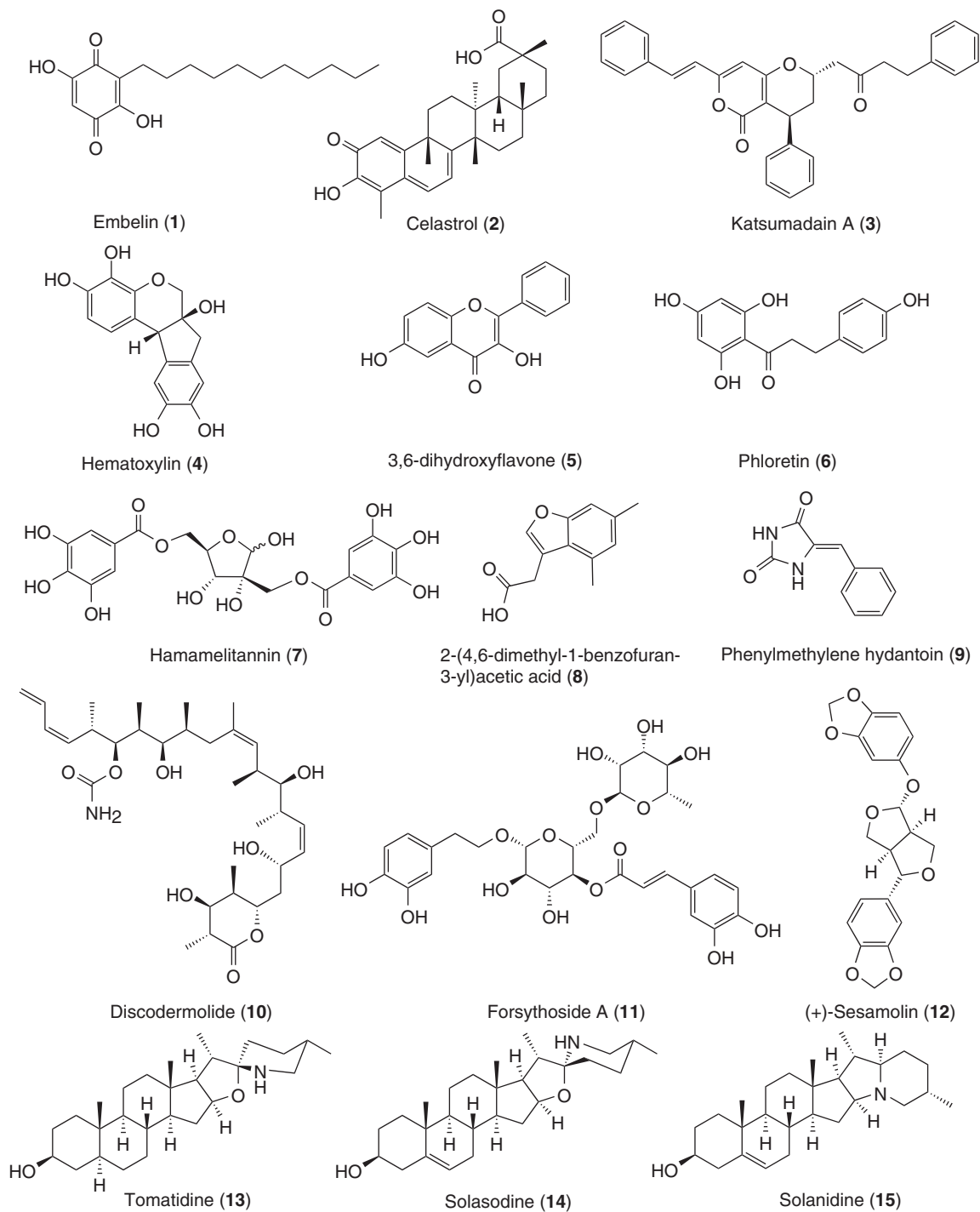
Chart: Chemical structures of bioactive products identified using virtual screening techniques.

5.3 Parallel Screening

The multitarget-oriented biological effects of NPs are of growing interest. The identification of various macromolecular systems, which will possibly be influenced by one ligand, can be facilitated using computational methods. Applying parallel screening, one compound is simultaneously screened against a set of models providing a predicted pharmacological activity profile (Steindl *et al.*, 2006). This profile includes not only desirable activities but also metabolic pathways and side effects. The methodology can be applied using all the virtual screening techniques described earlier, and has therefore the capacity of catalyzing drug discovery profoundly for all those diseases where molecular targets or ligands are well defined.

Especially for NP scientists, this methodology is a highly promising computer-assisted tool to enhance the identification of relevant targets of already laboriously isolated compounds, that is, target fishing (prospective application). For instance, Rollinger *et al.* (2009) isolated 16 secondary metabolites from *R. graveolens* and virtually screened their low energy conformers against a multitude of pharmacophore models. On the basis of the predicted ligand–target interactions, they focused on the three biological proteins, AChE, the human rhinovirus (HRV) coat protein, and the cannabinoid receptor type-2 (CB_2), and experimentally evaluated not only virtual hits but also nonhits. Although the selectivity and the predictive power of the individual pharmacophore models have varied, the obtained *in vitro* results underline the advantages of parallel screening, namely the requirement of a limited number of experiments for the identification of novel targets of the isolated compounds and the further insight into the pharmacological profile of *R. graveolens* as medicinal plant.

In the search for potential inhibitors against several therapeutically important targets belonging to various categories such as cell signal regulators, nitric oxide expression and production, nonsteroidal anti-inflammatory drugs, diabetic complications, and AIDS therapy, Ehrman, Barlow, and Hylands (2007b) used Random Forest to screen a database of Chinese herbal constituents. A large number of herbs were predicted to inhibit multiple targets. In addition, some appeared to contain inhibitors of the same target from different phytochemical classes. To support the predicted herb–target interactions, a literature search was conducted providing evidence.

Besides the discovery of novel targets, the rationalization of complex NP mixtures is a further application scenario of parallel screening (retrospective application). In this context, the traditional use of single medicinal plants or herbal remedies and formulae can be analyzed in terms of affected pathways or synergistic effects.

Huang, Qiao, and Xu (2007) studied the druglike features of the compounds present in Xuefu Zhuyu Decoction, a broad-spectrum recipe of the TCM. As this decoction is widely used in cardiac system diseases, a docking protocol was applied to show the interactions of the constituents with common cardiovascular target enzymes. In addition, the chemical space distribution and absorption, distribution, metabolism, and excretion (ADME) properties of the constituents have been analyzed.

Recently, Ehrman, Barlow, and Hylands (2010) used pharmacophore-assisted docking to screen Chinese herbs for compounds that may be active against four targets involved in inflammation, namely the cyclooxygenases 1 and 2, p38 mitogen-activated protein kinase (p38), c-Jun terminal-NH2 kinase (JNK), and type 4 cAMP-specific phosphodiesterase (PDE4). Interestingly, the results revealed that multitarget inhibitors are likely to be relatively common in Chinese herbs.

6 LIMITATIONS OF MOLECULAR MODELING

Despite the large number of success stories published in literature, there are specific limitations of molecular modeling, which have to be considered.

Concerning structure-based design, the user relies on crystal structures or average NMR structures to elucidate ligand–target interactions. Crystal structures are, however, models themselves including assumptions and interpretations introduced during the derivation of an atomic model from the experimentally observed electron density data (Davis, Teague, and Kleywegt, 2003). Older structure files from the PDB have to be evaluated and treated even more critically. Modelers have consequently to understand the difference between well-defined and less accurate portions of the coordinates. However, electron densities and B-factors in the structure file deposited by the crystallographer help to interpret the information. In addition, there are many software programs and online tools, which help to

better understand X-ray structures (Davis, Teague, and Kleywegt, 2003; Kirchmair *et al*., 2009).

As structure-based methods consider the macromolecule mainly as rigid, the dynamic nature of the target proteins is a further aspect, which can only be addressed to a limited extend. In most of the cases, the proteins still remain functional in their solid form, and many success stories have proved that such a model is extremely useful in early stages of drug discovery. Alternatively, the performance of molecular dynamics simulations of the apo protein system can account for the protein flexibility. In a few cases, there are several conformations of the macromolecule available as crystal structures, which allow the generation and the combination of several models for the virtual screening process.

One of the major challenges in drug discovery is the targeting of pathways mediated by cell surface receptors and large protein–protein interactions that are characterized by relatively flat and featureless binding domains. From the target's side, it is therefore advantageous to select target proteins with at least partly hydrophobic and cavity-shaped binding sites providing distinct overall interactions with the potential ligand.

Concerning similarity or ligand-based approaches, most of the computer programs use a multiconformational overlay to align chemical, steric, or electronic features. It is important for the model building to choose the used ligands in an appropriate way, that is, the ligands bind to the same binding site of the macromolecule and do not bind allosterically. The experimental evaluation of the used ligands should have been performed by a well-defined biological assay and the results should be comparable (preferably from the same laboratory).

Not only the ligands used for the model generation, but also the compounds present in the databases are substances that have been isolated and structurally elucidated. Therefore, the correctness of their structural formula (constitution and stereochemistry) is mandatory for the generation of reliable ligand-based models and the virtual screening procedure of databases.

Even after the validation of the virtual screening procedure, the user has to be aware that *in silico* techniques are only filtering tools that enrich the pool of potential ligands and give indications of ligand–target interactions. In any case, the experimental evaluation (e.g., fluorescence polarization displacement assays or NMR) is mandatory for

the determination of the underlying molecular mechanism.

7   SELECTION PROCESS OF VIRTUAL HITS FOR TESTING

As the isolation and structure elucidation of NPs are time- and cost-consuming procedures and there is only a limited access to commercial NPs, the results of a virtual screening may be too vague for an NP scientist. Thus, before starting a laborious phytochemical work, all indications and methods enabling a more focused procedure for the identification of bioactive NPs should be considered. For a general overview of such integrated or *in combo* approaches, the reader is referred to the article by Rollinger, Stuppner, and Langer (2008).

In general, there are two main strategies of integrated approaches. The first strategy consists of the generation of a sensitive data mining tool, its validation, and the screening of a large 3D multiconformational NP database. The resulting virtual hit list is subsequently evaluated in terms of traditional usage, phenomenological effects, guidance from chemotaxonomy, phylogenetic selection criteria, and information from *in vitro* screening of extracts or enriched fractions assumed to contain the focused metabolites. Further indications such as physicochemical properties, pharmacokinetics, and toxicity contribute to a reasonable choice of the natural material. In some cases, the NP scientist can be confronted with a difficult access to the natural material (domestic organisms are usually easier available) or small quantity of the compound in the source. Therefore, the estimated effort of the access, the isolation procedure, and the structure elucidation should be considered carefully. The chosen natural material or the enriched fractions are then either subjected to a bioguided process or the putative hits are identified by analytical tools such as liquid chromatography-mass spectrometry (LC-MS) or LC-NMR, and isolated.

Applying the second strategy, the selection of the natural material is firstly not guided by a virtual prediction, but a number of extracts is screened *in vitro*, whereas additional selection criteria such as traditional usage and phenomenological effects can also be considered. A small, project-centered database containing the known constituents of the natural material under investigation is established and virtually screened with the generated data mining tool.

This is especially helpful if not all the constituents are covered in large databases. The following procedure consists of the identification of the putative hits in the material as described earlier and the focused isolation process. The second strategy is especially valuable when an intricate pharmacological assay turns a bioguided fractionation into an unrealistic endeavor.

8 FROM THE BIOACTIVE COMPOUND TO POTENTIAL DRUGS

The discovery of a distinct biological activity is only the first step in the drug development process. A sufficient bioavailability and biological half-life, negligible side effects, and a lack of toxicity are additional mandatory requirements of a lead compound.

For an early assessment of drug or lead likeness of NPs, virtual filtering tools might have an invaluable impact. A very popular filter for the assessment of drug likeness is whether a virtual hit fulfills Lipinski's rule of five (molecular weight >500, more than five hydrogen bond donors, ten hydrogen bond acceptors, and the calculated log P is greater than five) (Lipinski *et al*., 1997). A violation of this rule is associated with poor absorption and permeation. Compounds that are intended to be further chemically optimized during drug development (drug likeness) should be smaller (molecular weight 200–350) and less lipophilic (specific derivatizations are possible).

Besides these properties, activities on side effect mediating targets, the so-called antitargets, can also be evaluated. In this context, large-scale *in silico* activity profiling is emerging as an additional tool. Antitarget models address the drug development obstacles such as adsorption, distribution (binding to plasma proteins and drug transporters), metabolism (e.g., cytochrome P450 system), and toxicity (e.g., cardiac potassium channel hERG). For the prediction of side effects, a selection of published antitarget pharmacophore models is given by Schuster and Wolber (2010).

9 FUTURE PERSPECTIVES

For thousands of years, the large pool of NPs has been the exclusive source of medicinal remedies. Although the nineteenth and twentieth centuries have been dominated by the development of synthetic drug substances particularly in industrialized societies, today, naturally derived agents also play an important role in drug discovery, either as single molecular entities or as herbal remedies.

This chapter highlights the relatively young application of *in silico*-guided strategies for the discovery and rationalization of bioactive NPs. In recent years, the use of computational methods in the field of NPs is constantly growing and first success stories have appeared in literature. Even in target classes, which have until recently been regarded as unsuitable for small-molecule inhibitors such as protein–protein interactions, successful studies based on *in silico* techniques have been published (Nikolovska-Coleska *et al*., 2004; Zhang *et al*., 2008).

Concerning the identification of potential inhibitors of drug targets, virtual screening offers valuable hints toward real hits, thereby diminishing experimental time, costs, and efforts and avoiding the environmental waste inherent to large-scale high throughput screenings. The completely general applicability to any target for which structural information, either of the ligand, the target, or both, is available is a further advantage of computational techniques. Advances of analytical methods for the conscientiously structure elucidation of secondary metabolites and target proteins are of immense importance not only for the appropriate use of *in silico* techniques but also for an early impact on drug discovery.

The parallel screening of NPs against a plethora of molecular targets considers their multitarget character. Providing an estimate of their putative pharmacological profile, this approach promises to enhance significantly the identification of relevant targets and antitargets. Furthermore, performing parallel screening on constituents of herbal remedies elucidates particular target ligands and the involved phytochemical classes. Synergistic effects as well as links between the ethnopharmacological use and the information on a molecular level can be elucidated and rationalized. This is especially valuable for the recent and rapid worldwide dissemination of Chinese medicine, which represents perhaps the most comprehensive repository of knowledge available today on the traditional use and synergistic effects of phytochemicals (Ehrman, Barlow, and Hylands, 2010).

Bioactive NPs have often complex molecular structures and are produced in small amounts by the parent organism. However, if there is a great demand for a bioactive NP, scientists try very hard to find an appropriate way for providing large amounts

(total synthesis, semisynthetic route, plant cell fermentation technology, or engineering of bacterial metabolisms) (Marco-Contelles, Rodriguez, and Garcia, 2005; Keasling, 2008). In addition, novel technologies (e.g., biosynthetic enzymes, genetically altered organisms, or combinatorial biosynthesis) hold great promises to generate NP derivatives (Li and Vederas, 2009). Consequently, the success of NPs as drug leads will depend, in future, not on the isolation of large amounts or the synthesis of their derivatives, but mainly on the efficient discovery of their already established bioactivity. *In silico* techniques provide the appropriate tool for this purpose.

10   RELATED ARTICLES

Innovative Strategies in the Search for Bioactive Plant Constituents; High Throughput Screening of Vegetal Natural Substances.

ENDNOTES

1. AnalytiCon. www.ac-discovery.com.
2. Specs Natural Products. www.specs.net.
3. BioFocus DPI Natural Product Database. www.biofocusdpi.com.
4. GreenPharma Natural Compound Library–GPN. www.greenpharma.com.
5. InterBioScreen Natural Compound Library. www.ibscreen.com.
6. TimTec Natural Compound Library. www.timtec.net/natural-compound-library.html.
7. Tulip. vitasml.centro.ru.
8. NCI open database. cactus.nci.nih.gov/ncidb2/download.html.
9. Derwent Chemical Information Systems, Inc.: Aliso Viejo, CA, 2005.
10. Omega. www.eyesopen.com.
11. ROTATE. www.molecular-networks.com.
12. ConFirm. http://accelrys.com.
13. DiscoveryStudio. http://accelrys.com.
14. MOE. www.chemocomp.com.

REFERENCES

Avila, C. M., Romeiro, N. C., Sant'Anna, C. M. R., *et al*. (2009) *Bioorg. Med. Chem. Lett.*, **19**, 6907–6910.

Berman, H. M., Westbrook, J., Feng, Z., *et al*. (2000) *Nucleic Acids Res.*, **28**, 235–242.

Boström, J. (2001) *J. Comput-Aided Mol. Des.*, **15**, 1137–1152.

Buckingham, J. ed. (2007; pp CD-ROM) *Dictionary of natural products on DVD*, CRC Press, Boca Raton, FL.

Chen, J., Nikolovska-Coleska, Z., Wang, G., *et al*. (2006) *Bioorg. Med. Chem. Lett.*, **16**, 5805–5808.

Cottrell, S. J., Gillet, V. J. and Taylor, R. (2006) *J. Comput-Aided Mol. Des.*, **20**, 735–749.

Csermely, P., Agoston, V. and Pongor, S. (2005) *Trends Pharmacol. Sci.*, **26**, 178–182.

Dai, Y., Qiao, L., Chan, K. W., *et al*. (2009) *Cancer Res.*, **69**, 4776–4783.

Danquah, M., Li, F., Duke, C. B. 3rd, *et al*. (2009) *Pharm. Res.*, **26**, 2081–2092.

Davis, A. M., Teague, S. J. and Kleywegt, G. J. (2003) *Angew. Chem. Int. Ed. Engl.*, **42**, 2718–2736.

Dixon, S. L., Smondyrev, A. M., Knoll, E. H., *et al*. (2006) *J. Comput-Aided Mol. Des.*, **20**, 647–671.

Doman, T. N., McGovern, S. L., Witherbee, B. J., *et al*. (2002) *J. Med. Chem.*, **45**, 2213–2221.

Dunkel, M., Füllbeck, M., Neumann, S., *et al*. (2006) *Nucleic Acids Res.*, **34**, D678–D683.

Ehrman, T. M., Barlow, D. J. and Hylands, P. J. (2007a) *J. Chem. Inf. Model.*, **47**, 254–263.

Ehrman, T. M., Barlow, D. J. and Hylands, P. J. (2007b) *J. Chem. Inf. Model*, **47**, 264–278.

Ehrman, T. M., Barlow, D. J. and Hylands, P. J. (2010) *Bioorg. Med. Chem.*, **18**, 2204–2218.

Ertl, P., Roggo, S. and Schuffenhauer, A. (2008) *J. Chem. Inf. Model.*, **48**, 68–74.

Ewing, T. J., Makino, S., Skillman, A. G., *et al*. (2001) *J. Comput-Aided Mol. Des.*, **15**, 411–428.

Fakhrudin, N., Ladurner, A., Atanasov, A. G., *et al*. (2010) *Mol. Pharmacol.*, **77**, 559–566.

Feher, M. and Schmidt, J. M. (2003) *J. Chem. Inf. Comput. Sci.*, **43**, 218–227.

Franke, L., Schwarz, O., Müller-Kuhrt, L., *et al*. (2007) *J. Med. Chem.*, **50**, 2640–2646.

Füllbeck, M., Michalsky, E., Dunkel, M., *et al*. (2006) *Nat. Prod. Rep.*, **23**, 347–356.

Gasteiger, J. (2006) *J. Med. Chem.*, **49**, 6429–6434.

Gasteiger, J., Teckentrup, A., Terfloth, L., *et al*. (2003) *J. Phys. Org. Chem.*, **16**, 232–245.

Gavernet, L., Talevi, A., Castro, E. A., *et al*. (2008) *QSAR Comb. Sci.*, **27**, 1120–1129.

Gerlach, S. (2007) Scientific reflections on traditional medicines in Austria. *Master Thesis, University of Vienna*, 381.

Gouda, H., Terashima, S., Iguchi, K., *et al*. (2009) *Bioorg. Med. Chem.*, **17**, 6270–6278.

Grabowski, K. and Schneider, G. (2007) *Curr. Chem. Biol.*, **1**, 115–127.

Grienke, U., Schmidtke, M., Kirchmair, J., *et al*. (2010) *J. Med. Chem.*, **53**, 778–786.

Grienke, U., Mihaly-Bison, J., Schuster, D., *et al*. (2011) *Bioorg. Med. Chem.*, **19**, 6779–6791.

Haefner, B. (2003) *Drug Discov. Today*, **8**, 536–544.

Halgren, T. A., Murphy, R. B., Friesner, R. A., *et al*. (2004) *J. Med. Chem.*, **47**, 1750–1759.

Henkel, T., Brunne, R. M., Müller, H., *et al*. (1999) *Angew. Chem. Int. Ed.*, **38**, 643–647.

Hieronymus, H., Lamb, J., Ross, K. N., *et al*. (2006) *Cancer Cell*, **10**, 321–330.

Hopkins, A. L. (2008) *Nat. Chem. Biol.*, **4**, 682–690.

Huang, Q., Qiao, X. and Xu, X. (2007) *Bioorg. Med. Chem. Lett.*, **17**, 1779–1783.

Jain, A. N. (2003) *J. Med. Chem.*, **46**, 499–511.

Johnson, M. and Maggoria, G. M. (1990) *Concepts and applications of molecular similarity*, John Wiley & Sons, Inc., New York.

Jones, G., Willett, P., Glen, R. C., *et al*. (1997) *J. Mol. Biol.*, **267**, 727–748.

Keasling, J. D. (2008) *ACS Chem. Biol.*, **3**, 64–76.

Kiran, M. D., Adikesavan, N. V., Cirioni, O., *et al*. (2008) *Mol. Pharmacol.*, **73**, 1578–1586.

Kirchmair, J., Laggner, C., Wolber, G., *et al*. (2005) *J. Chem. Inf. Model.*, **45**, 422–430.

Kirchmair, J., Wolber, G., Laggner, C., *et al*. (2006) *J. Chem. Inf. Model*, **46**, 1848–1861.

Kirchmair, J., Distinto, S., Markt, P., *et al*. (2009) *J. Chem. Inf. Model*, **49**, 678–692.

Koch, M. A., Schuffenhauer, A., Scheck, M., *et al*. (2005) *Proc. Natl. Acad. Sci. USA*, **102**, 17272–17277.

Kohonen, T. (1989) Self-Organization and Associative Memory, Springer, Berlin.

Korcsmaros, T., Szalay, M. S., Bode, C., *et al*. (2007) *Exp. Opin. Drug Discov.*, **2**, 799–808.

Laggner, C., Schieferer, C., Fiechtner, B., *et al*. (2005) *J. Med. Chem.*, **48**, 4754–4764.

Lee, M. L. and Schneider, G. (2001) *J. Comb. Chem.*, **3**, 284–289.

Lee, J.-Y., Jeong, K.-W., Lee, J.-U., *et al*. (2009a) *Bioorg. Med. Chem.*, **17**, 1506–1513.

Lee, J.-Y., Jeong, K.-W., Shin, S., *et al*. (2009b) *Bioorg. Med. Chem.*, **17**, 5408–5413.

Lei, J. and Zhou, J. (2002) *J. Chem. Inf. Computer Sci.*, **42**, 742–748.

Li, J. W.–. H. and Vederas, J. C. (2009) *Science*, **325**, 161–165.

Lin, L.-G., Xie, H., Li, H.-L., *et al*. (2008) *J. Med. Chem.*, **51**, 4419–4429.

Lipinski, C. A., Lombardo, F., Dominy, B. W., *et al*. (1997) *Adv. Drug Deliv. Rev.*, **23**, 3–25.

Maclean, D., Baldwin, J. J., Ivanov, V. T., *et al*. (1999) *Pure Appl. Chem.*, **71**, 2349–2365.

Marco-Contelles, J., Rodriguez, C. and Garcia, A. G. (2005) *Exp. Op. Ther. Pat.*, **15**, 575–587.

Martínez-Mayorga, K. and Medina-Franco, J. L. (2009) Chemoinformatics – applications in food chemistry, in *Advances in Food and Nutrition Research*, ed. S. L. Taylor, Academic Press, Burlington, vol. 58, pp. 33–56.

Miguet, L., Zhang, Z., Barbier, M., *et al*. (2006) *J. Comput-Aided Mol. Des.*, **20**, 67–81.

Morris, G. M., Goodsell, D. S., Halliday, R. S., *et al*. (1998) *J. Comput. Chem.*, **19**, 1639–1662.

Mudit, M., Khanfar, M., Muralidharan, A., *et al*. (2009) *Bioorg. Med. Chem.*, **17**, 1731–1738.

Mullard, A. (2010) *Nat. Rev. Drug Discov.*, **2011**(10), 82–85.

Newman, D. J. (2008) *J. Med. Chem.*, **51**, 2589–2599.

Newman, D. J. and Cragg, G. M. (2012) *J. Nat. Prod.*, **75**, 311–335.

Nikolovska-Coleska, Z., Xu, L., Hu, Z., *et al*. (2004) *J. Med. Chem.*, **47**, 2430–2440.

Pearl, L. H., Prodromou, C. and Workman, P. (2008) *Biochem. J.*, **410**, 439–453.

Pfisterer, P. H., Shen, C., Nikolovska-Coleska, Z., *et al*. (2011) *Bioorg. Med. Chem.*, **19**, 1002–1009.

Rarey, M., Kramer, B., Lengauer, T., *et al*. (1996) *J. Mol. Biol.*, **261**, 470–489.

Reddy, A. S., Pati, S. P., Kumar, P. P., *et al*. (2007) *Curr. Protein Pept. Sci.*, **8**, 329–351.

Rester, U. (2008) *Curr. Opin. Drug Discov. Devel.*, **11**, 559–568.

Rollinger, J. M., Haupt, S., Stuppner, H., *et al*. (2004) *J. Chem. Inf. Computer Sci.*, **44**, 480–488.

Rollinger, J. M., Stuppner, H. and Langer, T. (2008) Virtual screening for the discovery of bioactive natural products, in *Natural Compounds as Drugs*, eds. F. Peterson and R. Amstutz, Birkhäuser Verlag, Basel, pp. 212–249.

Rollinger, J. M., Schuster, D., Danzl, B., *et al*. (2009) *Planta Med.*, **75**, 195–204.

Rosén, J., Gottfries, J., Muresan, S., *et al*. (2009) *J. Med. Chem.*, **52**, 1953–1962.

Rouhi, A. M. (2003) *Chem & Eng News*, **81**, 77–91.

Salum, L. B., Dias, L. C. and Andricopulo, A. D. (2009) *QSAR Comb. Sci.*, **28**, 325–337.

Schuster, D. and Wolber, G. (2010) *Curr. Pharm. Des.*, **16**, 1666–1681.

Schuster, D., Waltenberger, B., Kirchmair, J., *et al*. (2010a) *Mol. Inf.*, **29**, 75–86.

Schuster, D., Kern, L., Hristozov, D. P., *et al*. (2010b) *Comb. Chem. High Throughput Screen.*, **13**, 54–66.

Schuster, D., Markt, P., Grienke, U., *et al*. (2011) *Bioorg. Med. Chem.*, **19**, 7168–7180.

Schwab, C. H. (2010) *Drug Discov. Today Technol.*, **7**, 245–253.

Singla, D., Sharma, A., Kaur, J., *et al*. (2010) *BMC Pharmacol*, **10**, 4.

Stahura, F. L., Godden, J. W., Xue, L., *et al*. (2000) *J. Chem. Inf. Computer Sci.*, **40**, 1245–1252.

Steindl, T. M., Schuster, D., Wolber, G., *et al*. (2006) *J. Comput-Aided Mol. Des.*, **20**, 703–715.

Talevi, A., Gavernet, L. and Bruno-Blanch, L. E. (2009) *Curr. Comp-Aided Drug Design*, **5**, 23–37.

Vallat, B. K., Pillardy, J. and Elber, R. (2008) *Proteins*, **72**, 910–928.

Venkatachalam, C. M., Jiang, X., Oldfield, T., *et al*. (2003) *J. Mol. Graph Model*, **21**, 289–307.

Verdine, G. L. (1996) *Nature*, **384**, 11–13.

Wagner, S., Hofmann, A., Siedle, B., *et al*. (2006) *J. Med. Chem.*, **49**, 2241–2252.

Wagner, S., Arce, R., Murillo, R., *et al*. (2008) *J Med Chem*, **51**, 1324–1332.

Waltenberger, B., Schuster, D., Paramapojn, S., *et al*. (2011) *Phytomedicine*, **18**, 119–133.

Wang, S.-Q., Du, Q.-S., Zhao, K., *et al*. (2007) *Amino Acids*, **33**, 129–135.

Watts, K. S., Dalal, P., Murphy, R. B., *et al*. (2010) *J. Chem. Inf. Model.*, **50**, 534–546.

Wei, H., Zhang, R., Wang, C., *et al*. (2007) *J. Theor. Biol.*, **244**, 692–702.

Wermuth, C.-G., Ganellin, C. R., Lindberg, P., *et al*. (1998) *Pure Appl. Chem.*, **70**, 1129–1143.

Westerheide, S. D., Bosman, J. D., Mbadugha, B. N., *et al*. (2004) *J. Biol. Chem.*, **279**, 56053–56060.

Wolber, G. and Langer, T. (2005) *J. Chem. Inf. Model.*, **45**, 160–169.

Wolber, G., Dornhofer, A. A. and Langer, T. (2006) *J. Comput. Aided Mol. Des.*, **20**, 773–788.

Zhang, T., Hamza, A., Cao, X., *et al*. (2008) *Mol. Cancer Ther.*, **7**, 162–170.

43

Innovative Strategies in the Search for Bioactive Plant Constituents

Emerson F. Queiroz and Jean-Luc Wolfender

School of Pharmaceutical Sciences, University of Geneva, University of Lausanne, Geneva, Switzerland

1 INTRODUCTION

Natural products (NPs) have been a source of inspiration for most of the active ingredients used for drug development throughout history, and nature is still a crucial reservoir of new molecules of potential therapeutic interest. According to Newman and Cragg, an important percentage of small-molecule new chemical entities registered as drugs worldwide during the period of 1981–2010 were NPs or were inspired by NPs. In specific therapeutic areas, such as anti-infective drugs, this percentage reaches 68%, and more than 74% of anticancer drugs are naturally inspired agents (Newman and Cragg, 2012).

One of the historical sources of drugs are the secondary metabolites from higher plants. These organisms are capable of synthesizing hundreds of NPs scaffolds, with incredible structural variations. Higher plants have been used for the treatment of human diseases for centuries. For example, the use of the latex of *Papaver somniferum* L. (Papaveraceae) was described on tablets from Mesopotamia from about 2600 BC (Newman, Cragg, and Snader, 2000). Currently, morphine, codeine, noscapine (narcotine), and papaverine, which were initially isolated from this plant and developed as single chemical drugs, are still used in the clinic (Chin *et al*., 2006). Plant constituents are still a source of inspiration for drug development. In recent 10 years, nine plant

derivatives have been the source of the development of innovative drugs with different therapeutic indications (Figure 1). Some recent examples include the development of apomorphine (Apokyn®) for the treatment of Parkinson's disease (Deleu, Hanssens, and Northway, 2004), nitisinone (Orphadin®) for the treatment of hereditary tyrosinaemia type 1 (Santra and Baumann, 2008), and miglustat (Zavesca®) for the treatment of the Gaucher disease (Abian *et al*., 2011).

According to Saklani and Kutty, 91 plant-derived compounds were in clinical trials in 2008 (Saklani and Kutty, 2008). Among them, for example, several compounds are still heavily studied. This is the case of combretastatin A4, a compound isolated from the South African medicinal tree *Combretum caffrum* Kuntze (Combretaceae). The original compound was derivatized to combretastatin A4 phosphate (Zweifel *et al*., 2011) and ombrabulin, (Sessa *et al*., 2013), and these compounds might be future anticancer drugs (Figure 2). In Australia, ingenol mebutate (ingenol-3-angelate) isolated from *Euphorbia peplus* L. (Euphorbiaceae) is currently in clinical trials for the treatment of actinic keratosis (Figure 2) (Siller *et al*., 2009). These examples indicate a renewal of interest from the industry for lead development from plant sources.

Despite intensive research investigation on plants, it was estimated in 2001 that only 6% of the approximately 300.000 species (some specialists estimate

---
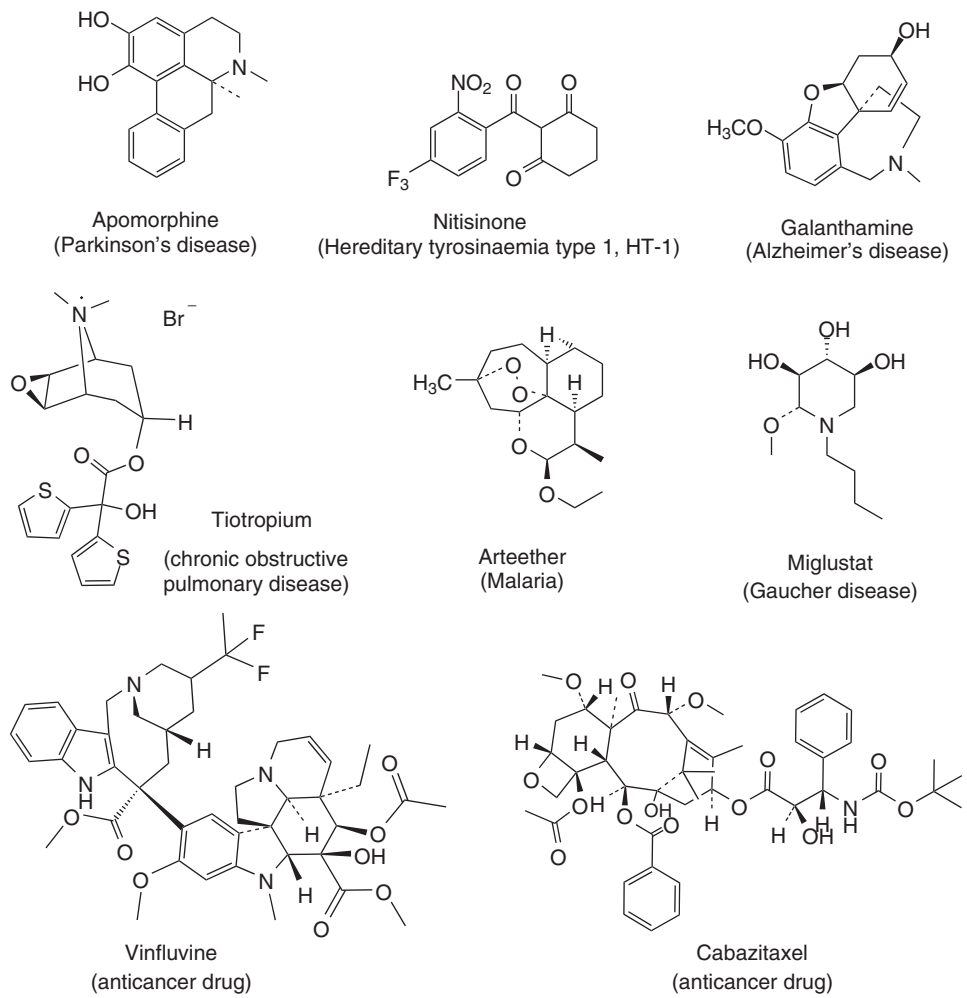
Figure 1 Examples of drugs developed from plants in recent 10 years.

- Apomorphine (Parkinson's disease)
- Nitisinone (Hereditary tyrosinaemia type 1, HT-1)
- Galanthamine (Alzheimer's disease)
- Tiotropium (chronic obstructive pulmonary disease)
- Arteether (Malaria)
- Miglustat (Gaucher disease)
- Vinfluvine (anticancer drug)
- Cabazitaxel (anticancer drug)
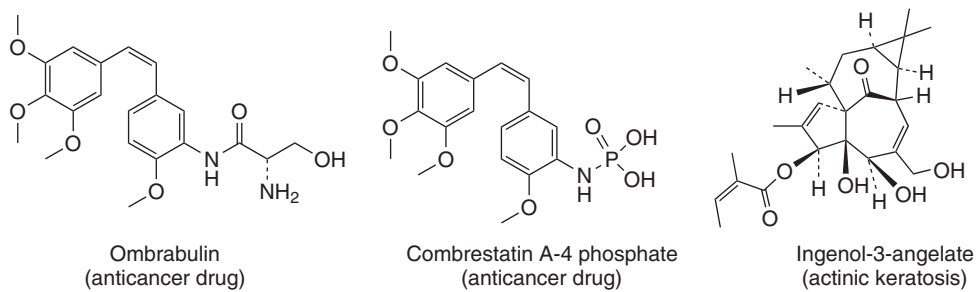
Figure 2 Examples of drug candidates from plants currently under clinical trials.

- Ombrabulin (anticancer drug)
- Combrestatin A-4 phosphate (anticancer drug)
- Ingenol-3-angelate (actinic keratosis)

as 500.000 species) have been studied pharmacologically, and approximately 15% have been the object of phytochemical study (Fabricant and Farnsworth, 2001; Gragg, 2006). In the past decade, NP research in the pharmaceutical industry has declined despite the significance of NPs as a source of therapeutic agents. According to Koehn and Carter, this could be credited to different factors such as the introduction of high throughput screening (HTS) using specific molecular targets, leading some companies to change from NP extract libraries to synthetic compounds libraries more adapted to this type of screening (Koehn and Carter, 2005).

Another reason is the difficulty to access the biological resources of countries of the southern hemisphere often holding rich biodiversity and important information on traditional usage because of new regulations enforced by the Rio Convention on Biodiversity (Kartal, 2007). On the other hand, the traditional bioassay-guided fractionation approach is generally considered by the industry as a slow process that is difficult to perform in the context of intensive HTS campaigns (Kingston and Newman, 2005). In addition, NPs are likely to have complex structures that render their identification difficult and contribute to problems related to supply and manufacturing. Recent scientific developments, however, should help to address these different issues and enhance the interest in NP drug discovery (Koehn and Carter, 2005).

In this chapter, the emphasis will focus on the latest approaches of lead finding from plant sources. Different topics will be discussed, starting from early plant selection to sample preparation, bioassay methods, and preparative chromatographic methods for the isolation of active compounds and dereplication strategies for NPs using hyphenated techniques. The role of these techniques in drug discovery using plants is also highlighted by different application examples.

2   SELECTION OF PLANTS

The collection and correct identification of the plant material is one of the most important points to be considered in phytochemical investigation. The choice of the appropriate plant material is usually based on information from traditional medicine, chemotaxonomic data, field observation, or random collection. Recently, other approaches have been used, such

as reverse pharmacognosy induction of new compounds by stress induction and biotransformation. The plant selection should also consider the legislation in force and, in particular, the exchange of biological material between countries should respect the requirements of biodiversity conventions (Cragg and Newman, 2013; Cragg *et al*., 2012). This should prevent natural resources and traditional knowledge, mainly located in third-world countries, from being exploited by industries of the northern hemisphere without equitable benefit sharing. The plants from tropical and subtropical regions represent an important reservoir of new molecules with potential therapeutic activities waiting to be discovered. Within the framework of these standards, a special place has been reserved for endangered species to ensure their protection.

2.1   Traditional Medicine

The selection of plants based on prior traditional medicinal usage (ethnopharmacology) is the main gateway for the discovery of promising molecules. Medicinal plants have indeed been a resource for early drug discovery. It has been demonstrated that 80% of the 122 compounds isolated from 94 species of plants used worldwide as drugs exhibit biological activities related to the traditional medicinal purposes of the original plant source (Fabricant and Farnsworth, 2001). According to the World Health Organization (WHO), "traditional medicines (including herbal drugs) have always played a key role in world health and continue to be used to treat a vast array of conditions and complaints" (Lu *et al*., 2011). In under-developed countries, these types of medicines often constitute a unique way to handle personal care. According to the WHO's Roll Back Malaria program, in Ghana, Mali, Nigeria, and Zambia, approximately 60% of all febrile cases in children, most likely due to malaria, are treated at home with herbal medicine (WHO, 2005). Thus, plant traditional preparations continue to constitute a valuable basis for the discovery of new active compounds provided that benefit sharing from the related discoveries is correctly handled with the country of origin and the healers holding the medicinal plant usage knowledge.

The advantage of this approach is that the validation of that efficacy is already documented for human usage. Another positive aspect is that traditional preparations with herbal drugs consumed by

population over centuries normally have low toxicity, and their bioactive constituents hold suitable pharmacokinetic properties. The lack of such information in random screening is a common reason of drug development failure (Nogueira *et al.*, 2009).

One of the key issues in ethnopharmacological studies is the correct identification of the disease for which an herbal remedy is used with the help of a physician (Raza, 2006). The traditional healers should also be consulted to translate traditional terminologies into their modern counterparts. The physician could also be an important partner if an ancient text will be the basis of the research. He or she can interpret the signs and symptoms mentioned and suggest the proper use of old traditional remedies in the light of modern medicine (Raza, 2006). Finally, the pharmacological model used to guide the isolation of the active compounds has to incorporate all of the information described earlier.

For example, the Tanzanian plant *Pyrenacantha kaurabassana* Baill (Icacinaceae) was selected for a phytochemical investigation based on its traditional usage for the treatment of HIV-infected patients. The efficacy of the herbal drug prescribed by traditional healers was also observed by physicians working in this region. The bioassay-guided fractionation using a validated HIV inhibitor bioassay led to the isolation of two new antiviral xanthones (Omolo *et al.*, 2012).

2.2 Chemotaxonomy

Chemotaxonomy, or the science of classification of plants as a function of the structures of their chemical constituents, can introduce useful elements. The constituents are often specific to a given botanical family, genus, or species. If an NP has interesting therapeutic properties, it may be possible to find analogous substances in species closely related to the plants from which it has been isolated (Sevenet, 1991). This is interesting because it yields the isolation of a series of structurally related compounds that can be used for further quantitative structure–activity relationship (QSAR) and pharmacomodulation studies. For example, the species-related origin of the given secondary metabolites is illustrated by the fact that acetogenins are only found in plants from the Annonaceae family (McLaughlin, 2008), whereas taxane diterpenoids are mainly observed in the Taxaceae family (Baloglu and Kingston, 1999).

2.3 Random Screening

Random screening also represents an important strategy for the discovery of NPs with interesting bioactivity. Using HTS, extracts, semi-purified fractions, or pure NPs can be screened at a large scale using different targets. For example, random screening has been used throughout the decades by the National Cancer Institute (NCI, USA), initially for the search of anticancer drugs. For this purpose, a huge library of extracts from different origins was created and continues to be developed (McCloud, 2010). In addition to the search for anticancer activity, the library continues to be screened for other purposes, including antiviral and antimicrobial activities (McCloud, 2010). A classic example of this strategy was the discovery of taxol from *Taxus brevifolia* L. (Taxaceae) and camptothecin from *Camptotheca acuminata* Decne. (Nyssaceae), two important anticancer drugs (Wall and Wani, 1996).

2.4 Field Observations

While searching for valuable plants to be investigated, field observations can also be considered. An example of a discovery based on such an approach was the discovery of the molluscicidal properties of saponins by Lemma. He observed that the berries of *Phytolacca dodecandra* L'Herit (Phytolaccaceae) were used as a traditional soap to kill aquatic snails and identified the saponins responsible for this activity. On the basis of the investigation, the berries from *P. dodecandra* were recognized as a readily available molluscicide to control schistosomiasis, a parasitic disease caused by several species of fluke of the genus *Schistosoma*, for which snails represent the intermediate host of the parasite (Lemma, 1970).

2.5 Reverse Pharmacognosy

The classical methods for NP drug discovery are based on the bioguided isolation of active compounds from complex matrices such as plant extracts. Reverse pharmacognosy proposes the opposite; molecules with potential therapeutic activities are selected *in silico* using inverse docking, a technique used to identify potential receptor targets for small

molecules (Do *et al.*, 2005). Once the biological target is chosen, the compound is submitted to a biological assay to confirm the inverse docking results. Finally, the related compounds (possessing related scaffolds) can be isolated in a targeted manner from natural sources reported to contain them.

Such an approach has been successfully used for the study of ε-viniferin, a phenolic compound from *Vitis vinifera* L. (Vitaceae). Inverse docking was used to identify putative binding biological targets for ε-viniferin. Among the 400 screened proteins, two targets were retained. The cyclic nucleotide phosphodiesterase 4 (PDE4) was the most interesting candidate for a future cosmetic application. The experimental binding tests on the six subtypes of PDE revealed a significant selectivity of ε-viniferin for the PDE4 subtype. This selectivity was confirmed by *in vitro* assays of ε-viniferin on the secretion of tumor necrosis factor-alpha (TNF-alpha) and interleukin-8. This approach demonstrated that ε-viniferin possesses anti-inflammatory properties by inhibiting the PDE4 subtype (Do *et al.*, 2005).

Similarly, the biological properties of meranzin, a coumarin derivative isolated from *Limnocitrus littoralis* (Miq.) Swingle (Rutaceae), were screened *in silico*. Among the 400 screened proteins, three targets were selected: COX1 (cyclo-oxygenase 1), COX2, and PPARγ (peroxisome proliferator-activated receptor). Binding tests were performed for these three protein candidates, as well as two negative controls. The inhibition of COX1 and COX2 in binding assays was measured and confirmed the predictions. A cellular assay was performed to prove that PPARγ was activated by meranzin. These results demonstrated that reverse pharmacognosy along with its inverse docking component is a powerful tool to identify the biological properties of natural molecules for plant extracts containing these compounds (Do *et al.*, 2007).

2.6 Stress Induction

Classical drug discovery investigations of plant constituents have essentially explored secondary metabolites present on plant tissues at the moment of plant collection. The dynamic stress response of an organism is usually not considered, although the antimicrobial activity of many phytoalexins has been reported (Hammerschmidt, 1999). In many cases, these compounds are non-detectable in non-stressed plant tissues (Poulev *et al.*, 2003). Using elicitation experiments, it is possible to obtain minor compounds with strong bioactivities (Vasconsuelo and Boland, 2007). For example, it has been demonstrated that the content of biologically active phenolics in the medicinal plant *Echinacea purpurea* L. Moench (Asteraceae) can be significantly increased (up to 10-fold compared with the control) using a foliar application of elicitors (Kuzel *et al.*, 2009). The authors used natural plant stress mediators and their derivatives (acetylsalicylic acid, salicylic acid, and methyl salicylate) as well as a newly introduced biocompatible metal elicitor [titanium(IV) ascorbate] as active components of foliar sprays. Plant metabolite induction and the exploitation of natural elicitation mechanisms in host–plant interactions could be valuable alternatives to enhance the drug-discovery hit rate when screening natural extracts (Poulev *et al.*, 2003; Wolfender and Queiroz, 2012).

3 EXTRACTION METHODS

The classical procedure for the extraction of plant secondary metabolites is generally performed by maceration of the dried material at room temperature with solvents of increasing polarities. An alternative approach consists of the extraction of fresh plant material with a polar solvent for the first extraction step, and subsequent solvent partition allows a finer division into different polarity fractions (Hostettmann, Marston, and Hostettmann, 1998). These approaches are efficient and provide satisfactory partition of the secondary metabolites based on their polarities. However, these procedures require relatively high amounts of solvents and are time consuming.

In recent years, a series of new extraction methods have been applied to plant extraction, including microwave-assisted extraction (MAE) (Xia *et al.*, 2011), ultra-sonication-assisted extraction (UAE) (Klen and Vodopivec, 2011), supercritical fluid extraction (SFE) (Ramandi *et al.*, 2011), and pressurized solid–liquid extraction (PSLE) (Wan *et al.*, 2012; Wijngaard and Brunton, 2009). These different methods have substantial benefits compared with conventional methods in terms of decreasing organic solvent consumption, reproducibility, and throughput. The basic principles of these innovative

extraction methods are discussed in the following reviews (Coelho *et al.*, 2012; Delazar *et al.*, 2012; Mottaleb and Sarker, 2012).

4 STRATEGIES FOR SAMPLE PREPARATION FOR BIOLOGICAL SCREENING

Many ubiquitous naturally occurring compounds may interfere with the isolation and purification of a desired bioactive plant constituent from crude plant extract (e.g., pigments, lipids, and tannins), and they also interfere with the results of bioassays. For example, fatty acids have been observed to give false-positive results in several bioassays (Balunas *et al.*, 2008; Ringbom *et al.*, 2001), and tannins have also been observed to give false results (Wall and Wani, 1996). Thus, it is important to develop an efficient procedure to eliminate those compounds from the crude extracts before biological screening. Some of the most common procedures of sample preparation are summarized in the following sections.

4.1 Pigments

Pigments, such as chlorophylls, may be present at high concentration in leaf extracts. These compounds are not easy to eliminate, but the following methods could be applied. Activated carbon (or charcoal) is known to decolorize solutions through the adsorption of pigments. Extract solutions can be percolated through a charcoal column, or the powder can be mixed with the liquid to be decolorized, left to stand for a period of time and filtered (Lee *et al.*, 1995). One of the limitations of this approach is that adsorption is not selective, and other compounds present may be eliminated. Solid-phase extraction (SPE) on reversed-phase C18 represents an alternative method. The elution of the extract using such columns with methanol containing small amounts of water traps the apolar constituents of extracts, whereas most polar and medium polar secondary metabolites are eluted (Glauser *et al.*, 2008). If the adsorbed lipids, carotenoids, and chlorophylls also have to be studied, they can be recovered by further elution with chloroform (Rodrigues *et al.*, 2008).

4.2 Lipids

To remove the fats and waxes from an extract, the plant material can be percolated with petroleum ether or hexane and allowed to dry before the full extraction process. The material also may be directly extracted with the solvent of choice and defatted later by a liquid–liquid extraction. Acetonitrile has also been reported to be efficient for wax elimination. For example, the solubilization of the chloroform extract of a *Petunia* species in boiling acetonitrile for 1 h followed by cooling at 5°C enables the elimination of the waxes after decantation (Elliger *et al.*, 1990). Another method using acetonitrile precipitation was used to eliminate the waxes from the hexane extract of *Artemisia annua* L. (Asteraceae). The extract was first dissolved in chloroform, and waxes were precipitated with acetonitrile (Zheng, 1994). Lipids can also be removed from an extract by filtering the sample through a reverse-phase SPE cartridge where the lipids are retained as mentioned earlier (Rodrigues *et al.*, 2008).

4.3 Tannins

One convenient method to eliminate tannins consists of filtration on a polyamide support (Collins *et al.*, 1998). Small columns packed with polyamide remove >99% polyphenolic compounds in plant extracts (Collins *et al.*, 1998). This technique has been used with success to eliminate the polymeric polyphenols from *Rhodiola rosea* L. (Crassulaceae) crude extracts (Ioset *et al.*, 2011). Tannins can also be selectively removed with a novel absorbent prepared from bovine skin collagen fibers, which are used as the stationary phase. A systematic study performed on typical medicinal plant constituents as probes suggested that this stationary phase is much more selective than polyamide (Liao and Shi, 2005). This novel absorbent is, however, not yet commercially available.

5 REFINED EXTRACTS

As plant extracts are complex mixtures composed of hundreds of different substances, an active compound can be present in low amounts in this matrix, and

thus, its biological activity may not be detected. A strategy to avoid this problem is to refine the extracts by enriching compounds present in small amounts to more efficiently evaluate the bioactivities of their constituents. An interesting and quick technique is to conduct a preliminary fractionation from a few milligrams of extract using SPE. The selection of the phase chemistry for SPE will depend on the type of extract to be enriched. For example, polar extracts can be pre-fractionated using a reverse stationary phase, whereas apolar extracts can be fractionated with a normal stationary phase. Solid-phase extraction can be efficiently automated and is especially helpful when large numbers of samples have to be regularly screened (Romanik *et al*., 2007).

For the enrichment of alkaloids, nonaqueous silica-based strong cation-exchange (SCX) SPE can be used, taking advantage of the alkaline properties of the analytes for a specific trapping and for the removal of non-alkaloid components. Such a method was used in the study of the alkaloids from *Scopolia tangutica* Maxim (Solanaceae) by liquid chromatography-mass spectrometry (LC-MS) (Long *et al*., 2012).

Very specific retention can be obtained in SPE using unusual stationary phases based on molecularly imprinted polymers (MIPs). This methodology has opened new horizons in the SPE fractionation. Molecular imprinting was first introduced in 1972 by Wulff and Sarhan (Wulff and Sarhan, 1972) and has now found numerous applications. MIPs have potentially the same selective binding capacity of the natural antibodies and enzymes. They are constructed by copolymerization of a functional monomer in the presence of a model molecule to produce a three-dimensional polymer complex. After removal of the model molecule, it is possible to use the polymer matrix to select compounds with a size, shape, and functionality similar to the original model molecule (Lasakova and Jandera, 2009).

The application of SPE on MIPs permits not only pre-concentration and sample cleaning but also selective extraction of the target compound (Lasakova and Jandera, 2009). For example, the selective extraction of the diterpene kirenol with SPE on specially designed MIPs from *Siegesbeckia pubescens* Makino (Asteraceae) herbal extracts, with an extraction yield higher than 80%, was made possible. These results demonstrate that a good imprinting effect and excellent selectivity of MIPs were obtained (Chen, Wang, and Shi, 2012) (Figure 3).

Recently, MIPs were used for the separation of phenolic acids from natural plant extracts using an anion-exchange mechanism. The imprinted anion-exchange polymer achieved high recovery rates by the SPE of phenolic acids from *Salicornia herbacea* Fee. ex Ung.-Sternb. (Chenopodiaceae) (protocatechuic acid: 90.1%, ferulic acid: 95.5%, and caffeic acid: 96.6%). The different phenolic acids were isolated by repeated SPE cycles. The approach could be used to separate other phenolic acids or organic acids from complex samples (Bi, Tian, and Row, 2012).

A particular interest of MIPs is that structural analogs of the original molecule are frequently recognized on the imprinted cavities. Consequently, an identified inhibitor of a receptor or an enzyme can be used as a model to screen unknown structurally related inhibitors. This approach was applied in the study of *Caragana jubata* (Pall.) Poir. (Leguminosae), a plant used in traditional Tibetan medicine. Quercetin, a known protein tyrosine kinase inhibitor, was used as a template to search for epidermal growth factor (EGFR) inhibitors. Using this approach, two new inhibitors, (*E*)-piceatannol (IC$_{50}$ = 4.5 mM) and butein (IC$_{50}$ = 10 mM), were selectively separated from the EtOAc extract and were shown to be the main active extract constituents (Zhu and Xu, 2003).

Thus, SPE MIPs represent an original approach for an efficient and selective extraction of specific compound or compounds classes.

6 HPLC-MICROFRACTIONATION

High performance liquid chromatography (HPLC)-microfractionation is an important technique for biological activity profiling, especially given the development of bioassays using multi-well plates. Using this technique, the biological activities of an individual component of a complex extract can be measured without the need for a tedious isolation process at a large scale. In parallel to HPLC profiling, the coupling of HPLC to different spectroscopic techniques, such as diode array detection (DAD)/ultraviolet (UV), mass spectrometry (MS), and nuclear magnetic resonance (NMR), provides on-line spectroscopic data of the compounds presented in the mixture (Wolfender, Queiroz, and Hostettmann, 2005; 2006; Wolfender and Queiroz, 2012). Thus, the combination of HPLC chemical and
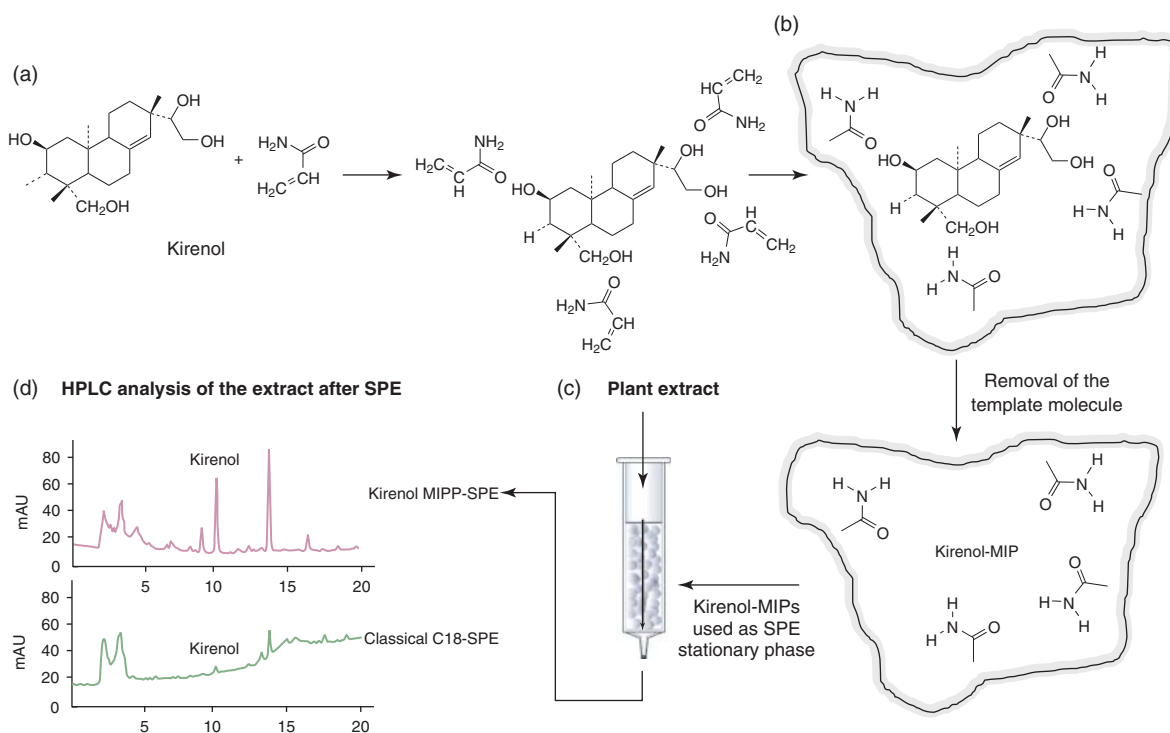
Figure 3 Schematic representation of the preparation and use of the kirenol MIPs. (a) The obtained MIPs were synthesized through the copolymerization of acrylamide and ethylene glycol dimethacrylamide in the presence of kirenol (template) using the non-covalent imprinting approach. (b) The template molecule kirenol is removed from the polymer. (c) The crude extract of *Siegesbeckia pubescens* was enriched in kirenol by a solid-phase extraction procedure on the MIPs. (d) HPLC analysis of the selective extraction of kirenol achieved with 80.9% yield. (*Source*: F.F. Chen, *et al*. (2012). Reproduced by permission of Elsevier.)

biological profilings through microfractionation represents a powerful way to rapidly identify bioactive constituents directly in crude extracts with a minimal amount of biological material.

The identification of the inhibitors of acetylcholinesterase from the ethanolic extract of *Annona glabra* L. (Annonaceae) illustrates the potential of this approach in medicinal plant research (Tsai and Lee, 2010). An analytical-scale sample of the EtOH extract was fractionated by HPLC-DAD-MS into 96-well microplates, which, after evaporation, were assayed against AChE. This strategy afforded the retention time, UV spectrum, and electrospray ionization mass spectrometry (ESIMS) data for each active compound. On the basis of these preliminary data, the targeted isolation of the active compounds using different chromatographic techniques, including centrifugal partition chromatography (CPC) and semi-preparative HPLC, afforded 20 compounds. The most active compounds were palmatine and

pseudopalmatine, with IC_{50} inhibition of acetylcholinesterase values of 0.4 and 1.8 μM, respectively (Tsai and Lee, 2010).

Recently, a similar HPLC-based activity profiling approach was used to screen the GABA A modulators using a two-microelectrode voltage clamp assay on *Xenopus laevis* oocytes (Zaugg *et al*., 2010). The approach was applied in the study of the ethyl acetate extract of *Piper nigrum* L. (Piperaceae) or black pepper fruits. Piperine was identified as the main active compound, together with 12 structurally related less active or inactive piperamides (Figure 4) (Zaugg *et al*., 2010). The strategy has been applied for the study of different plant extracts with success (Li *et al*., 2010; Yang *et al*., 2011; Zaugg *et al*., 2011).

7   NATURAL PRODUCTS LIBRARY

Since the beginning of the 1990s, HTS began to be used by the pharmaceutical industry for the screening
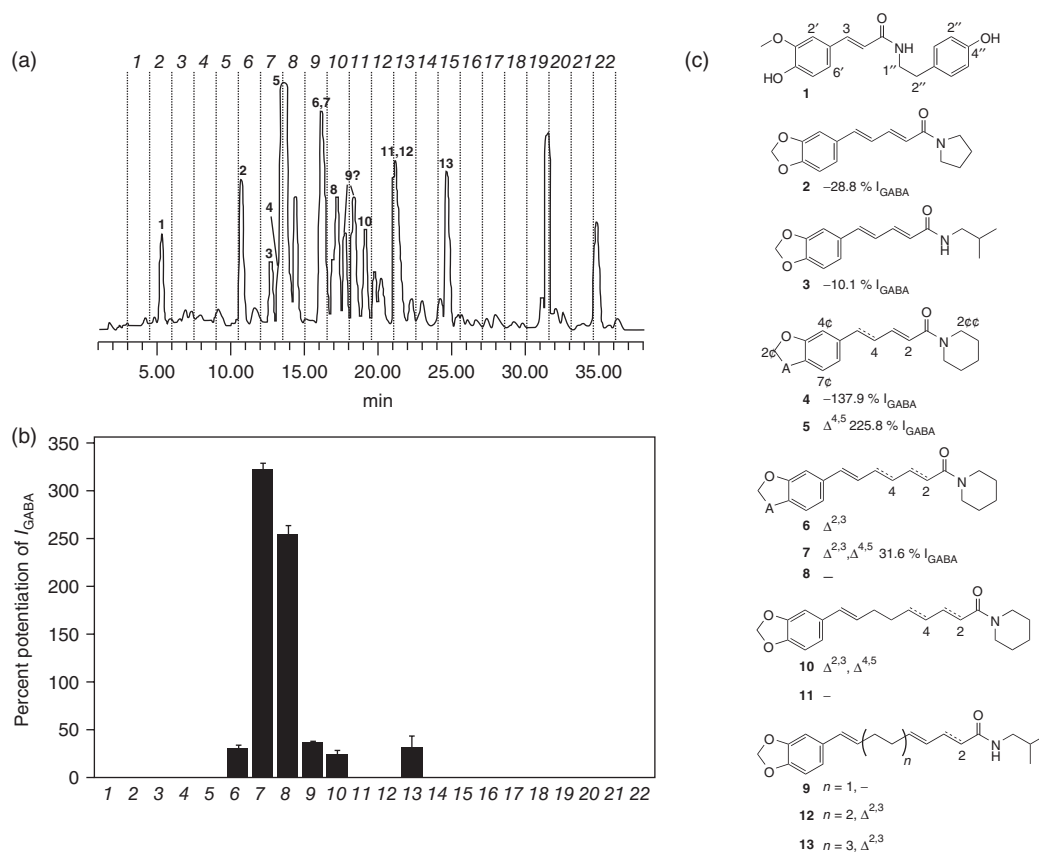
Figure 4 Localization of the active compounds in the crude extract of *Piper nigrum* L. using HPLC-microfractionation. (a) HPLC-UV of the active extract. (b) Biological activity of the collected time-based fractions on the potentiation of the GABA-induced chloride current in *Xenopus* oocytes (IGABA). (c) Structures of piperamides and potentiation (%) of the GABA-induced chloride current (IGABA) in *Xenopus* oocytes by 100 μM 2–5 and 7. (*Source*: Reprinted with permission from Zaugg J, Baburin I, Strommer B, Kim HJ, Hering S, Hamburger M. 2010. HPLC-Based Activity Profiling: Discovery of Piperine as a Positive GABA(A) Receptor Modulator Targeting a Benzodiazepine-Independent Binding Site. J Nat Prod 73(2):185–191. Copyright (2010) American Chemical Society.)

of small-molecule compounds. The majority of the compounds screened using these methodologies are provided by combinatorial chemistry (Eldridge *et al.*, 2002). For many reasons, NPs initially did not work well in the HTS process. Generally, NP drug discovery has been viewed as a time-consuming process. From the screening of crude extracts to the isolation and structure elucidation of the active compounds, months or years are sometimes necessary. In the past decade, with the development of automated processes for separation and identification, it has become possible to build libraries of NPs that are compatible with HTS biological assays.

One way to efficiently create such libraries is to combine advances in high throughput technologies with more conventional chemistry techniques

for the generation of large collections of purified fractions of small-molecule NPs for HTS (Eldridge *et al.*, 2002). The construction process initiates with the fractionation of plant extracts using automatic flash chromatography (Figure 5a and b). The resulting fractions are subjected to an SPE to remove tannins and high molecular weight (MW) components from polar extracts with weight limit filters (Figure 5b). The fractions are further separated using a high throughput parallel four-channel reversed-phase preparative HPLC system, resulting in fractions containing a mixture around of five compounds/well. The fractions are analyzed by a parallel eight-channel analytical liquid chromatography-evaporative light scattering detection-mass spectrometry (LC-ELSD-MS)
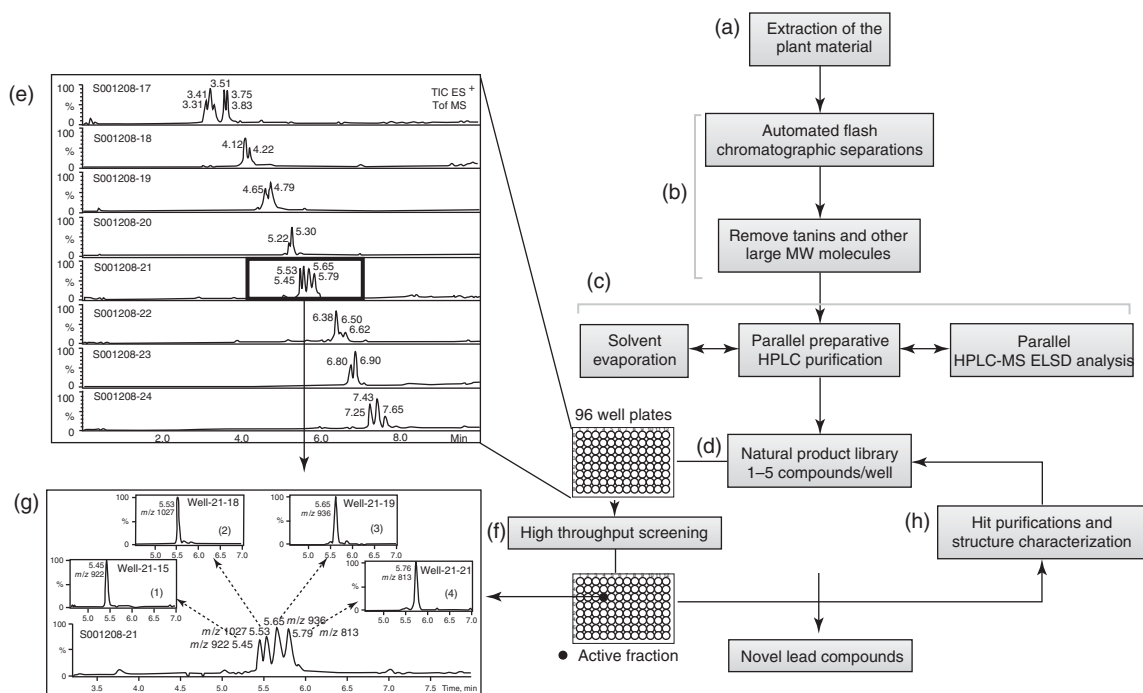
Figure 5 Schematic representation of high throughput methods applied to the process of drug discovery from natural resources (Eldridge *et al.*, 2002). (a) Extraction of the plant material. (b) Pre-purification steps. (c) Semi-preparative HPLC fractionation and parallel LC-ELSD-MS analysis. (d) Collection of the HPLC fractions in 96-well plates (NPs library). (e) Example of typical LC-MS chromatograms (TIC) of some natural product library fractions. (f) 96-well plates containing the NPs submitted to the HTS screening. (g) MS data of the compounds present in the active fraction. (h) Target purification of the active compounds from the original extract for structural elucidation. (*Source*: Reprinted with permission from Eldridge GR, Vervoort HC, Lee CM, Cremin PA, Williams CT, Hart SM, Goering MG, O'Neil-Johnson M, Zeng L. 2002. High-throughput method for the production and analysis of large natural product libraries for drug discovery. Anal Chem 74(16):3963–3971. Copyright (2002) American Chemical Society.)

system (Figure 5c). The fractions containing detectable compounds are collectively called *the library*, from which more focused libraries can be obtained for a future biological screening (Figure 5d and e). The proposed library is stocked in 96- or 384-well plates.

In one study (Eldridge *et al.*, 2002), a library containing up to 36,000 NP fractions was produced. This library has been screened in different drug-discovery programs against diverse biological targets (Figure 5f). The hit rates have been 0.5% or lower depending on the biological assay. Active compounds have been rapidly isolated in small amounts (5–50 μg) from ~100 μg of a preparative HPLC fraction. Figure 5g shows an example of the purification of a bioactive library fraction containing four peaks into a total of four individual components, using a shallow-gradient semi-preparative HPLC separation. The sample loaded onto the column was

~100 μg, and the recovery of each peak was approximately 5–50 μg. For *de novo* structural elucidation, a micro-coil flow probe with a capillary-based microliter-volume flow cell was used (Figure 5h). The methodology was applied with success in the study of the stem bark of *Taxus brevifolia* Nutt. (Taxaceae) (Eldridge *et al.*, 2002). The biological *in vitro* screening with three cancer cell lines demonstrated that the process enabled the discovery of active anticancer compounds that were not detected in the flash fractions from which the library originated (Eldridge *et al.*, 2002).

More recently, a high resolution (HR) and high throughput strategy for NP fractionation system was proposed (Tu *et al.*, 2010). The major consideration in designing this system was to apply high throughput techniques to NP fractionation and provide high quality samples, which was compatible with current modern screening methods.

In this case, the NPs obtained from plant materials were extracted and first pre-fractionated on polyamide SPE cartridges to remove polymeric polyphenols (tannins). The free tannin extracts followed a high throughput automated fractionation, drying, weighing, and reformatting for screening and storage. The fractions were analyzed by ultra high pressure liquid chromatography-mass spectrometry (UHPLC-MS), photo diode array (PDA), and evaporative light scattering detection (ELSD) detectors providing chemical information for the structure elucidation of the compounds present in the active fractions. Parts of the fraction material were submitted to several high throughput cellular screening assays. This processing system was able to fractionate 2600 unique NP samples per year, providing 62.000 fractions at the 0.5–10-mg scale for the creation of libraries that will serve as long-term biological screening resources. The method was determined to be very efficient for the screening of a large number of NP samples (Tu *et al*., 2010).

8 BIOASSAY METHODS

Although a plant extract can contain hundreds or even thousands of chemical substances, it is not rare to find one single compound responsible for the biological activity. Thus, the biological assay must be simple and sensitive to detect the active compound in small concentrations. The assays are also required to be specific for the target involved.

The principal targets for biological tests can be divided into six groups: isolated subcellular systems (enzymes and receptors), animal or human cell cultures, lower organisms (bacteria, fungi, and viruses), invertebrates (insects, crustaceans, and molluscs), vertebrates (zebrafish and African clawed frog), isolated organs of mammalian animals, and whole mammalian animal models. These biological assays have different levels of complexity, cost, and relevance in relation to a possible activity in humans (Segalat, 2007) (Figure 6).

With the development of molecular biology techniques, a series of new targets were discovered,

- Simplicity
- Ease to manipulation
- Cheapness
- Rapidity to execution

1. Enzymatic assays
2. Bacterial and yeast assays
3. Cell line assays
4. Invertebrate models
5. Zebrafish models
6. *Xenopus oocytes*
7. "Human lungs on a chip"
8. Mammalian animal models
9. Humans (clinical trials)

- Complexity
- Cost
- Relevance to humans
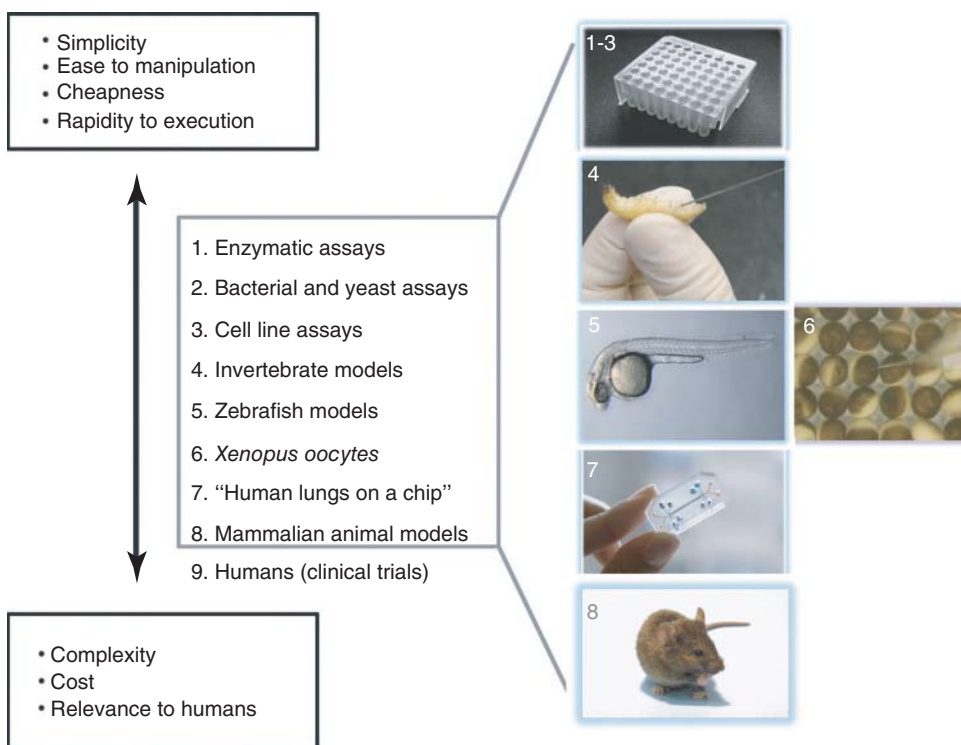
Figure 6 Scheme presenting the main *in vitro* and *in vivo* biological assays used for screening new chemical entities. (*Source*: Reprinted with permission from Segalat L. 2007. Invertebrate animal models of diseases as screening tools in drug discovery. ACS chemical biology 2(4):231–6. Copyright (2007) American Chemical Society.)

enabling the development of *in vitro* tests. Ideally, in NP discovery, the bioassay should be simple and provide, if possible, already relevant *in vivo* information. A few examples that are well suited for NP screening both *in vitro* and *in vivo* are highlighted in the following sections.

8.1 *In Vitro* Bioassays

With the increase of invasive fungal infections and the appearance of multidrug resistant strains, there is an urgent need for new antifungal and antimicrobial drugs, and plants are a valuable source of new antifungals (Ostrosky-Zeichner *et al.*, 2010). Bioassays for the evaluation of antifungal or antibacterial activities are relatively easy to perform. The samples to be assayed are simply placed in contact with the microorganism of interest, and the inhibition of spore growth or death is measured. Several methods for screening antifungal activity are accessible, but because they are not similarly sensitive or not based on the same principle, the results can be significantly influenced by the method chosen.(Hadacek and Greger, 2000) The most used methodologies for assessing antifungal activity are bioautography, (Rahalison *et al.*, 1991) disk diffusion, (Nweze, Mukherjee, and Ghannoum, 2010), agar dilution (Trancassini *et al.*, 1986), and microdilution methods (Liu *et al.*, 2007).

Recently, new strategies have emerged to improve antifungal research (Butts and Krysan, 2012). One approach to increase the sensitivity and specificity of growth-based assays is to involve specific mutants of pathogenic fungi. Another way is to apply nongrowth-based assays with increased sensitivity and/or specificity (Butts and Krysan, 2012). For example, a set of ~4600 gene deletion mutants of *Saccharomyces cerevisiae* were used to study the mechanism of action of the antifungal extracts of *Echinacea purpurea* (L.) Moench and *Echinacea angustifolia* D.C. (Asteraceae) (Mir-Rashed *et al.*, 2010). The analyses revealed an increase of cell death in the yeasts of *S. cerevisiae* and *Cryptococcus neoformans* after *Echinacea* extract treatments. Fluorescence microscopy analysis indicated that *S. cerevisiae* treated with *Echinacea* extracts was significantly more disposed to cell wall damage than non-treated cells. This approach demonstrates the potential of gene deletion arrays to understand the NP antifungal mode of action and provides compelling evidence that the fungal cell walls are a target of *Echinacea* extracts (Mir-Rashed *et al.*, 2010).

In the past decade, spectacular progress has been made in cellular biology and molecular pharmacology, affording the identification of a number of enzymes related to diseases and leading to the development of numerous assays for new enzymatic targets (Copeland, Harpel, and Tummino, 2007). These enzymatic tests are generally very specific and sensitive. Such assays are often relatively easy and require small amounts of material and can be automated for HTS. A substantial number of commercialized drugs used to treat infectious diseases have their mode of action based on enzyme inhibition. For example, the most widely used group of antibiotics act via the inhibition of the β-lactams, enzymes that are involved in bacterial cell wall biosynthesis. Certain new antibiotics, such as retapamulin, the first agent of the new pleuromutilin class of antibacterials, act by a new mode of action involving the inhibition of the protein synthetic apparatus of the bacterial ribosome, a nonprotein-based enzyme (Parish and Parish, 2008). Enzyme inhibition has been the basis of antiviral therapies. For example, AIDS treatment consists of a combination of HIV proteases and HIV reverse-transcriptase inhibitors (Copeland, Harpel, and Tummino, 2007).

A screening for protein kinase C inhibitors revealed the antiviral activity of the very common triterpene betulinic acid based on a bioactivity-guided study of the Chinese plant *Syzygium claviforum* (Roxb.) WALL (Myrtaceae) (Fujioka *et al.*, 1994). On the basis of these results, a series of synthetic derivatives of betulinic acid were tested, and one (Bevirimat®) reached the level of clinical trials for the treatment of HIV (Nguyen *et al.*, 2011). These examples indicate that NP-lead finding based on enzymatic bioassays represents a valuable approach.

More recently, whole-cell screens have been used to search for new drugs against infectious diseases. The reason for this is the failure of high throughput methods to assess the inhibition of enzymes that is crucial in microbial life (Brodin and Christophe, 2011). While more complex than enzymatic assays, whole-cell screening assays can now be automated, and the development of cellular imaging methods has enabled the efficient usage of high content screening technology (HCS). This technology involves the acquisition of images from a large set of cellular samples in microtiter plates coupled

to automated image analysis and data processing. HCS was first applied to various metabolic diseases (Bickle, 2010; Korn and Krausz, 2007; Rausch, 2006) before being extended to the area of infectious diseases (Brodin and Christophe, 2011).

An HCS assay to monitor *Leishmania major* amastigote colonization in human THP-1-differentiated monocytes was recently developed (Siqueira-Neto *et al*., 2010). The assay validation was performed with *Leishmania* promastigote forms. In an effort to find new compounds with leishmanicidal properties, 26.500 structurally diverse chemical compounds were screened. On the basis of inhibition and cellular toxicity criteria, the activity of the remaining 124 compounds was confirmed against the intramacrophagic amastigote form of the parasite. *In vitro* microsomal stability and cytochrome P450 (CYP) inhibition of the two most active compounds from this screening were assessed to obtain preliminary information on their metabolism in the host. Following this strategy, promising antileishmanial compounds that could be potential leads for future drugs were discovered (Siqueira-Neto *et al*., 2010).

In the field of anticancer drug research, plants have always been and remain an invaluable source of potent lead compounds (Harvey and Cree, 2010). In this case, both high throughput cell-based and molecular approaches have been extensively used, and a vast number of NPs or crude extracts have been screened (Harvey and Cree, 2010; McCloud, 2010).

In addition to these classical approaches, during the past decade, a stem-cell-like subset of cancer cells has been identified in many malignancies. These cells, referred to as *cancer stem cells (CSCs)*, are of particular interest because they are believed to be the clonogenic core of the tumor and therefore represent the cell population that drives growth and progression (Vermeulen *et al*., 2012). In this field, new biological assays have been developed to screen agents with CSC-specific toxicity and these open new opportunities for NP research (Gupta *et al*., 2009).

According to the disease, investigations on isolated organs of animals are necessary. Pharmacological models such as the perfused frog heart have been used for the study of cardiac glycosides (Ngombe *et al*., 2010). Other tests are conducting using perfused livers, (Priestap *et al*., 2012) guinea pig heart, (Shi *et al*., 2010), and other organs. However, these assays are complex and difficult to apply to classical bioguided isolation. They are timeconsuming, expensive to maintain, raise ethical issues, and require a considerable amount of the pure compound. The information provided by these tests is often useful but may not predict effects in humans.

To circumvent some of these issues, recently "Organs-on-Chips" technologies have emerged that provide extraordinary opportunities to create cell culture microenvironments of living organs (Huh *et al*., 2012). The principle of this microscale engineering technology relies on the cultivation of living cells within microfluidic devices that have been microengineered to rebuild tissue arrangements observed in living organs (Huh *et al*., 2012). This promising approach will be state of the art in future for the study of physiology in an organ-specific context. On the other hand, this technology could be the basis for the development of specialized *in vitro* disease models.

8.2 *In Vivo* Bioassays

In many instances, *in vivo* studies have to be performed to validate key activities that have been observed *in vitro*. Classical *in vivo* pharmacological studies using mice and rats are efficient, but they are long and require pure isolated compounds in high amounts and requires all ethical issues that have been carefully considered. Furthermore, *in vivo* assays are costly, hampering the processing of a large number of samples. In recent years, novel opportunities for *in vivo* NP discovery have arisen through the recent emergence of novel *in vivo* model systems, such as *Galleria mellonella* and *Danio rerio* (zebrafish) (Crawford *et al*., 2011; Crawford, Esguerra, and deWitte, 2008; Mylonakis *et al*., 2005). The use of these models has increased in popularity owing to the numerous advantages over mammalian models, including ethical, logistical, and budgetary issues.

8.2.1 *Galleria mellonella* Model

Galleria mellonella is a moth of the Pyralidae family. The insect larva of this species is used as a host for human pathogens. This insect has been used as a whole-animal high throughput system for *in vivo* drug evaluation. The model has been used to evaluate the antifungal activity of amphotericin B, flucytosine, and fluconazole following infection with

Cryptococcus neoformans (Mylonakis *et al*., 2005). An evaluation of the virulence of *Candida albicans* mutants' infection of this insect model and mice has shown a good correlation (Brennan *et al*., 2002). Recently, this model was used for assessing the *in vivo* efficacy of compounds against *Staphylococcus aureus* and methicillin-resistant *S. aureus* infections (Desbois and Coote, 2011).

Larvae of *G. mellonella* are cheap to purchase, and the results can be acquired rapidly (48 h). Moreover, a large number of larvae can be inoculated in a short period of time, and the immune response of insect larvae is similar to the innate immune response of mammals (Mylonakis *et al*., 2005). Pathogen death occurs with similar mechanisms in the insects and mammals (Desbois and Coote, 2011). In addition, infections of this larvae model can be performed at the temperature of the human body, 37.8°C, which can be important for the expression of certain microbial virulence factors (Konkel and Tilly, 2000).

8.2.2 Zebrafish Model

Zebrafish have several advantages as a drug discovery model. A number of genetic studies have been performed in zebrafish and have led to the identification of therapeutically relevant genes in several indication areas, including cardiovascular, neurological, gastrointestinal, musculoskeletal, and metabolic disorders (Amsterdam *et al*., 2004; Driever *et al*., 1996; Haffter *et al*., 1996; Patton and Zon, 2001). Zebrafish have a genetic, physiologic, and pharmacologic similarity to humans. Moreover, they have a small size, optical transparency, rapid development, and large numbers of embryos and larvae, which, as a primary system for experimental analysis, can be generated in a short period of the time. The small size of zebrafish embryos and larvae (1–5 mm) is compatible with microtiter plates for screening (primarily 24- and 96-well plates, but even 384-well plates are possible), thus requiring only microgram amounts of each extract, fraction, or pure compound to be evaluated (Crawford, Esguerra, and de Witte, 2008).

Zebrafish have been used to describe the angiogenic activity of the extract of *Angelica sinensis* Oliv. Diels. (Apiaceae) (Lam *et al*., 2008) as well as the anti-angiogenic activity of solenopsin, an alkaloid isolated from the fire ant (*Solenopsis invicta*) (Arbiser *et al*., 2007). To evaluate the efficacy of the zebrafish

angiogenesis assay, transgenic embryos (fli-1:EGFP) have been used for NP discovery. On the basis of this assay, the activity of more than 80 extracts from East African medicinal plants was evaluated and two extracts were observed to inhibit intersegmental vessel (ISV) outgrowth in fli-1:EGFP embryos in a dose-dependent manner. Zebrafish bioassay-guided fractionation identified the active components of these plants as emodin, an inhibitor of the protein kinase CK2, and coleon A lactone, a rare abietane diterpenoid with no previously described bioactivity (Crawford *et al*., 2011). These results suggest that the combination of zebrafish bioassays with simple chromatography methods is an effective strategy for the rapid identification of bioactive NPs. Zebrafish bioassays have strong advantages compared to other *in vivo* models, and zebrafish bioassays can be miniaturized and easily implemented for *in vivo* bioactivity studies (Challal *et al*., 2012).

9 STRATEGY FOR THE ISOLATION OF BIOACTIVE COMPOUNDS

To assess the bioactivity of any NPs, they have to be purified and irrelevant compounds from the original biological matrix have to be discarded. This step is also crucial for the *de novo* structure elucidation of unknown compounds by 1D and 2D NMR and high resolution mass spectrometry (HRMS), whereas new approaches also allow *de novo* identification even in only partially purified mixtures (Robinette *et al*., 2012). Classical isolation strategies involve complex fractionation schemes with various preparative chromatographic methods requiring kilogram amounts of the plant material to obtain milligram amounts of pure NPs (Hostettmann, Marston, and Hostettmann, 1998). Such methodologies are efficient and still important when important amounts of a given NP are required for detailed *in vivo* studies. Nevertheless, these methodologies are relatively time consuming when performed in a nontargeted manner. On the other hand, they can also lead to the isolation of NPs of low interest and with biological properties of limited interest.

On the basis of a combination of recent semi-preparative HPLC-MS methods, microNMR methods, and state-of-the-art bioassays, the entire process can be performed at the microgram level with high throughput. This significantly accelerates

the identification of active compounds and allows the efficient preliminary evaluation of their bioactivities (Wolfender and Queiroz, 2012).

To use these miniaturized methods, a direct association has to be established among the analytical profiling of the crude active extract, the spectroscopic data obtained on-line or at-line, and the information from the bioassays (Wolfender and Queiroz, 2012). For example, a strategy that combines ultra high pressure liquid chromatography-time of flight-mass spectrometry (UHPLC-TOF-MS) profiling followed by semi-preparative HPLC-MS for the isolation of minor key plant biomarkers and their characterization by at-line microflow NMR has been efficient for the isolation and identification novel wound-induced jasmonate derivatives in *Arabidopsis thaliana* Schur. (Brassicaceae) (Glauser *et al*., 2008) and for the isolation of phytotoxic fungal metabolites at the microgram scale (Glauser *et al*., 2009). The possibility of analyzing the metabolites with microNMR in such approaches also gives rise to the possibility to perform absolute quantification even of unknown compounds, and this is key to precisely evaluating bioactive compounds even if only low microgram amounts are available (Crawford *et al*., 2010).

This type of miniaturized approach is illustrated by the microisolation of wound-induced leaf constituents of *A. thaliana* and their subsequent characterization by capillary NMR (CapNMR) (Figure 7) (Glauser *et al*., 2008). The strategy relies on the optimization of the chromatographic analysis using UHPLC-TOF-MS and modeling software. The optimized method was then transferred to semi-preparative LC conditions with MS detection (Figure 7a and b). The semi-preparative step afforded an enriched fraction that contained the target biomarkers (MW at *m/z* 225), which was selected for the final purification (Figure 7c). Modeling software was used to predict the optimal isocratic conditions at ACN/H$_2$O (8 : 92, v/v) for the final separation step. The conditions were directly calculated from data obtained using two generic linear gradients of the crude extract, considering the retention time of all constituents eluting close to the biomarker of interest. The LC-MS semi-preparative chromatogram of this fraction displayed an overall separation of all constituents and a satisfactory resolution of *m/z* 225 isomers (Figure 7d). The retention times for the four isomers were close to those predicted after transfer from the UHPLC isocratic conditions. The MS spectrum demonstrated the absence of co-eluting

ions (Figure 7e). Although the *m/z* 225 peaks were not baseline resolved, the fractionation was precise enough to obtain 96 well/microfractions with more the 80% of each peak for further NMR investigation. The NMR analysis was performed using a microflow CapNMR probe that allowed the elucidation of the different compounds at a microgram scale (Figure 7f).

10 STRATEGIES FOR EARLY IDENTIFICATION OF ACTIVE COMPOUNDS IN PLANT EXTRACTS

The discovery of active metabolites from higher plants relies mainly on the biological screening of crude extracts followed by activity-guided fractionation (Wall and Wani, 1996). In this strategy, bioassays serve as a guide during the whole isolation process. For this reason, the assays should be compatible with the screening of complex extracts and should be easy to implement to analyze the numerous fractions generated during the bioassay-guided fractionation. *In vitro* bioassays offer the convenience of high throughput and are compatible with the assay of small quantity of the material. However, *in vitro* bioassays are not always predictive of clinical efficacy, because in some cases, the active compounds can be pro-drugs that need to be metabolized to express the pharmacological effects. In this case, *in vivo* tests are needed. Salicylic derivatives such as salicin present in the *Salix* species are an example. Akao *et al*. performed a pharmacokinetic and pharmacological study to compare the antipyretic effects of salicin, saligenin, and salicylic acid in rats. The results indicated that the active compound was salicin, an inactive pro-drug, which is gradually transported to the lower part of the intestine, hydrolyzed to saligenin by intestinal bacteria, and converted into salicylic acid after absorption. Thus, the compound produces an antipyretic action without causing gastric injury (Akao *et al*., 2002).

Frequently, the bioassay-guided fractionation approach leads to the isolation of known metabolites. Therefore, the chemical screening of crude extracts constitutes an efficient complementary approach. This procedure enables the recognition of known metabolites at the earliest stage of separation (dereplication), thus avoiding the costly and time-consuming isolation of common constituents (Wolfender *et al*., 2010). On-line spectroscopic information collected directly when profiling an extract

(a) UHPLC-TOF-MS 3% gradient

(b) Semi-preparative HPLC separation

Transfer

(d) Isocratic semi-preparative HPLC-MS chromatogram

(c) Zoom in the region of the targeted compounds (m/z 225)

(e) Purity control by LCMS

(f) Structure elucidation of the isolated compounds
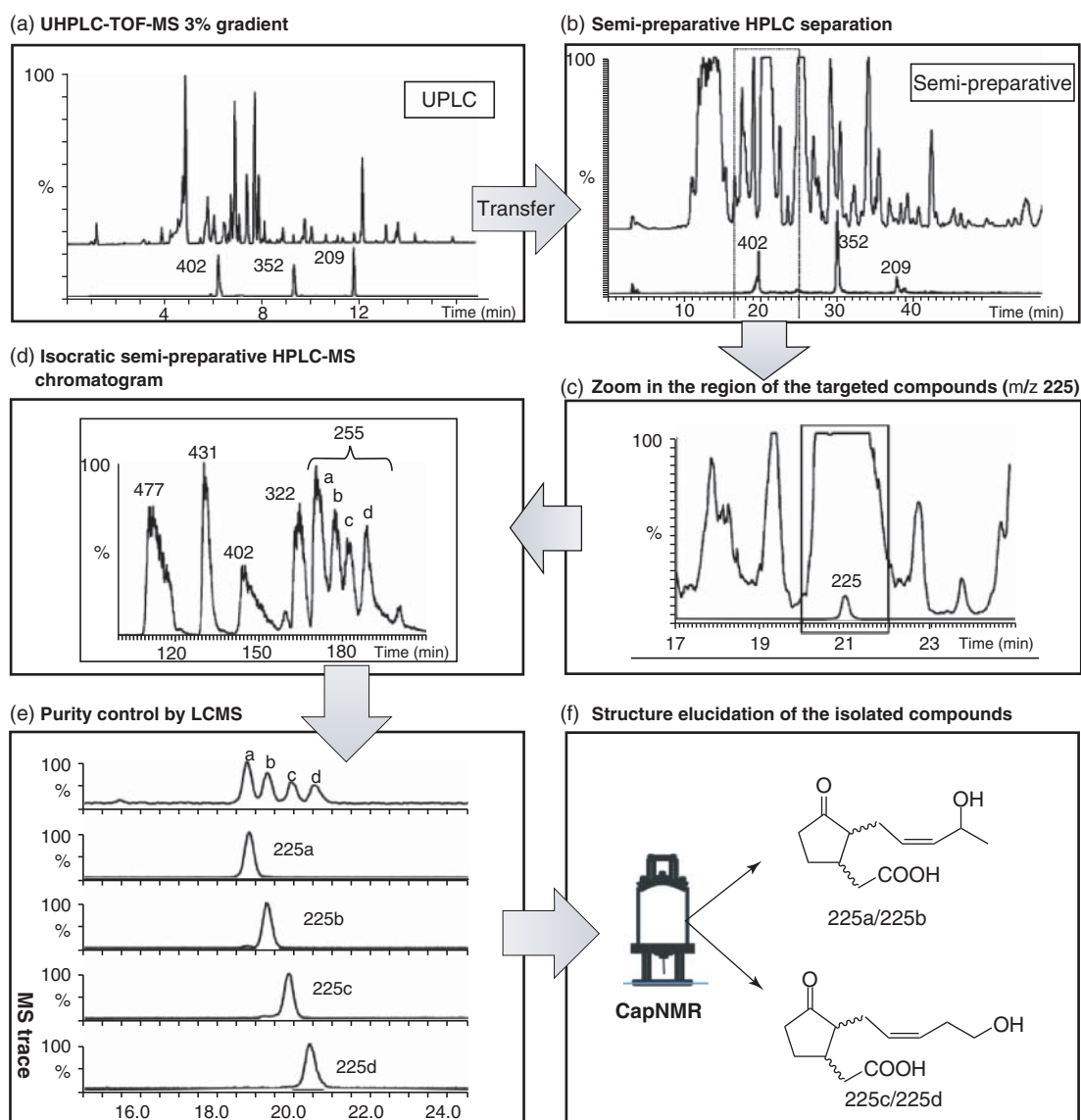
Figure 7 Rational strategy for the isolation of the compounds by transposition of analytical UHPLC-MS to semi-preparative HPLC-MS. (a) UHPLC-TOF-MS profiling (3% gradient) of a crude leaf extract of *Arabidopsis*. (b) Semi-preparative HPLC after gradient transfer from the UHPLC-TOF-MS conditions. (c) Selection of the fraction at the region of RT 20–22 min containing the targeted biomarkers (*m/z* 225 isomers [M − H]$^-$) highlighted by the single ion trace. (d) Separation of the *m/z* 225 isomers using semi-preparative HPLC-MS in isocratic conditions optimized by software modeling. (e) Purity control of the semi-preparative HPLC-MS fractions using UHPLC-TOF-MS. (f) Structure elucidation of the isolated compounds using capillary NMR analysis. (*Source*: G. Glauser, *et al.* (2008). Reproduced by permission of Elsevier.)

also allows for the localization and target isolation of new types of constituents with potential activities. The potential of the chemical screening strategy has been considerably increased by the recent development of hyphenated techniques, which are able to

provide efficient separation of metabolites and, at the same time, valuable on-line and at-line structural information (Wolfender, Marti, and Queiroz, 2010).

For the chemical screening of crude plant extract, HPLC represents a powerful and versatile technique.

This technique has been extensively developed through the years in terms of convenience, speed, choice of stationary column phases, sensitivity, applicability to a broad variety of sample matrices, and ability to hyphenate the chromatographic method to different spectroscopic detectors (Natishan, 2004). Moreover, the method has been adapted to the analysis of a broad range of NPs generally without the need for complex sample preparation. The latest developments in HPLC, including the recent introduction of highly pH-stable phases, sub-2-μm particles (Nguyen *et al*., 2006), and monolith columns, have considerably improved the performances of HPLC systems in terms of resolution, speed, and reproducibility. For the analysis of crude extract from higher plants, two main categories of detectors can be used: simple detectors used for the recording of chromatographic traces for profiling or quantification purposes (e.g., UV, ELSD, and electron capture detector (ECD)) and detectors for hyphenated systems that generate multidimensional data (chromatographic and spectroscopic) for on-line identification and dereplication purposes (e.g., UV-DAD, MS, and NMR). These hyphenated methods include LC-photodiode array detection (LC-PDA)(Huber and George, 1993), LC-mass spectrometry (LC-MS, LC-MS–MS or LC-multiple stage mass spctrometry (MS^n))(Niessen, 2006), and NMR (either directly hyphenated with LC (LC-NMR), (Albert, 2002) used with at-line methods, such LC–SPE–NMR (Jaroszewski, 2005), or after microfractionation with microflow LC-NMR (CapNMR) (Olson *et al*., 2004)). These hyphenated methods should ideally provide the same type of spectroscopic information that could be obtained for pure NPs after isolation. Gathering such structural information before a complete bioactivity-guided isolation study is thus a strategy for efficient targeting in large-scale isolation procedures.

11 IDENTIFICATION OF ACTIVE COMPOUNDS FROM HIGHER PLANTS USING METABOLOMICS

Throughout the years, metabolomics has rapidly developed as an important tool for the study of plant secondary metabolism (Wolfender *et al*., 2013). The ultimate aim of metabolomics is to measure all the metabolites in an organism both qualitatively and quantitatively, which can provide a clear metabolic picture of a living organism under certain conditions (Yuliana *et al*., 2011a). Metabolomics has also proved to be very effective in the research of active compounds from medicinal plants.

In the study of *Galphimia glauca* Cav. (Malpigiaceae), a plant popularly employed in Mexico for the treatment of central nervous system (Cardoso-Taketa *et al*., 2008), a multivariate data analysis (MVDA) of ^{1}H NMR metabolic profiling of crude extracts from wild plant populations, collected from different locations, revealed significant differences in composition. Principal component analysis (PCA) revealed that two populations out of the six studied manifested differences when the main constituents were analyzed (PC1 and PC2). These two PCs permitted the differentiation of the various sample populations, depending on the presence of galphimines. This information consistently correlated with the corresponding HPLC analysis and with neuropharmacological activity (Cardoso-Taketa *et al*., 2008). The strategy applied in this study opened new possibilities for the proper selection of plant populations with suitable metabolic and pharmacologic profiles for the development of standardized herbal medicines.

On the basis of a similar NMR metabolomic approach, combined with a comprehensive extraction method consisting of a continuous flow of solvent mixtures through plant material, the bioactive compounds of *Orthosiphon stamineus* Benth. (Lamiaceae) were efficiently identified. The plant fractions were subjected to MVDA to determine their adenosine A1-binding activities. On the basis of these results, two flavonoids from a large number of chemicals were clearly identified as being responsible for the adenosine A1-binding activities without any further purification steps (Yuliana *et al*., 2011b).

12 CONCLUSION AND PERSPECTIVES

In the drug discovery process, NPs from plant sources still represent an important source of new scaffolds. The development of efficient methods for their rapid identification in crude extracts is crucial to maintain the competiveness of the screening of crude extracts or enriched NP fractions compared with the screening libraries consisting of synthetic compounds. To enlarge the NP diversity of therapeutic

interest, new significant technological advancements provide an opportunity to explore minor constituents that have never been evaluated for their pharmacological potential. Methods that can elicit the production of compounds under stress conditions may also generate new interesting leads. The chemical or biological transformation of NP scaffolds may also enhance the chemical diversity of NPs.

The combination of state-of-the-art analytical and spectroscopic methods and information-rich bioassays until the *in vivo* level opens many new fields for NP research. The screening of well-defined microfractions in 96-well plates offers the possibility of automation and avoids the difficulty of interpreting bioassay results that are obtained with crude extracts and the incompatibilities that are often encountered with such samples.

Microfractions also offer the possibility of constructing semi-purified NP libraries without the need of obtaining fully purified NPs. This significantly speeds up the evaluation of thousands of compounds using an appropriate HTS platform. The introduction of HCS and the development of the new bioassays compatible with the analysis of NP mixture afford new tools for the efficient exploitation of plant chemical diversity. These different methodologies will certainly contribute to the study of tropical plant biodiversity before the large-scale destruction of these natural resources goes too far. Further exploitation of these NPs may be achieved through biosynthetic chemistry or metabolic engineering approaches, which are emerging research fields in drug discovery (Koehn and Carter, 2005).

In addition, the development of metabolomics methods for finding active metabolites provides a good way to target the isolation of biomarkers of interest solely based on the NMR or MS characteristics of these compounds. Metabolomics is presently also being implemented in the analysis of animal and human fluids to investigate the mode of action of herbal drugs with established clinical efficacies for which active ingredients have not been identified (Yuliana *et al*., 2011a). Such an approach is expected to yield information about pro-drugs, synergy, and the effects of drug metabolism, which are other important aspects of NP research on medicinal plants and phytotherapy. The plants used in traditional medicine still represent an invaluable source of NPs that can be efficiently studied based on reverse pharmacological approaches provided that all of the ethical and benefit-sharing issues have been correctly addressed.

For this, the most effective strategy is to perform multidisciplinary studies in which the collaboration among botanists, ethnobotanists, pharmacognosists, phytochemists, biologists, pharmacologists, and physicians is a prerequisite for success.

13 RELATED ARTICLES

Metabolomics.

ACKNOWLEDGMENTS

REFERENCES

Abian, O., Alfonso, P., Velazquez-Campoy, A., *et al*. (2011) *Mol. Pharmaceutics*, **8**(6), 2390–2397.

Akao, T., Yoshino, T., Kobashi, K., *et al*. (2002) *Planta Med.*, **68**(8), 714–718.

Albert, K. (2002) , in *On-line LC-NMR and related techniques*, ed. K. Albert, John Wiley & Sons, Ltd, Chichester.

Amsterdam, A., Nissen, R. M., Sun, Z., *et al*. (2004) *Proc. Natl. Acad. Sci. U. S. A.*, **101**(35), 12792–12797.

Arbiser, J. L., Kau, T., Konar, M., *et al*. (2007) *Blood*, **109**(2), 560–565.

Baloglu, E., and Kingston, D. G. I. (1999) *J. Nat. Prod.*, **62**(10), 1448–1472.

Balunas, M. J., Su, B., Brueggemeier, R. W., *et al*. (2008) *J. Nat. Prod.*, **71**(7), 1161–1166.

Bi, W., Tian, M., and Row, K. H. (2012) *J. Chromatogr. A*, **1232**, 37–42.

Bickle, M. (2010) *Anal. Bioanal. Chem.*, **398**(1), 219–226.

Brennan, M., Thomas, D. Y., Whiteway, M., *et al*. (2002) *FEMS Immunol. Med. Microbiol.*, **34**(2), 153–157.

Brodin, P., and Christophe, T. (2011) *Curr. Opin. Chem. Biol.*, **15**(4), 534–539.

Butts, A., and Krysan, D. J. (2012) *PLoS Pathog*, **8**(9), e1002870.

Cardoso-Taketa, A. T., Pereda-Miranda, R., Choi, Y. H., *et al*. (2008) *Planta Med.*, **74**(10), 1295–1301.

Challal, S., Bohni, N., Buenafe, O. E., *et al.* (2012) *Chimia*, **66**(4), 229–232.

Chen, F. F., Wang, R., and Shi, Y. P. (2012) *Talanta*, **89**, 505–512.

Chin, Y. W., Balunas, M. J., Chai, H. B., *et al.* (2006) *AAPS J.*, **8**(2), E239–E253.

Coelho, J. P., Cristino, A. F., Matos, P. G., *et al.* (2012) *Molecules*, **17**(9), 10550–10573.

Collins, R. A., Ng, T. B., Fong, W. P., *et al.* (1998) *Biochem. Mol. Biol. Int.*, **45**(4), 791–796.

Copeland, R. A., Harpel, M. R., and Tummino, P. J. (2007) *Expert Opin. Ther. Targets*, **11**(7), 967–978.

Cragg, G. M., and Newman, D. J. (2013) *Biochim. Biophys. Acta*, **1830**(6), 3670–3695.

Cragg, G. M., Katz, F., Newman, D. J., *et al.* (2012) *Nat. Prod. Rep.*, **29**(12), 1407–1423.

Crawford, A., Bohni, N., Maes, J., *et al.* (2010) *Planta Med.*, **76**(12), 1331–1332.

Crawford, A. D., Liekens, S., Kamuhabwa, A. R., *et al.* (2011) *PLoS One*, **6**(2), e14694.

Crawford, A. D., Esguerra, C. V., and de Witte, P. A. M. (2008) *Planta Med.*, **74**(6), 624–632.

Delazar, A., Nahar, L., Hamedeyazdan, S., *et al.* (2012) *Methods Mol. Biol.*, **864**, 89–115.

Deleu, D., Hanssens, Y., and Northway, M. G. (2004) *Drugs Aging*, **21**(11), 687–709.

Desbois, A. P., and Coote, P. J. (2011) *J. Antimicrob. Chemother.*, **66**(8), 1785–1790.

Do, Q. T., Renimel, I., Andre, P., *et al.* (2005) *Curr. Drug Discov. Technol.*, **2**(3), 161–167.

Do, Q. T., Lamy, C., Renimel, I., *et al.* (2007) *Planta Med.*, **73**(12), 1235–1240.

Driever, W., Solnica-Krezel, L., Schier, A. F., *et al.* (1996) *Development*, **123**, 37–46.

Eldridge, G. R., Vervoort, H. C., Lee, C. M., *et al.* (2002) *Anal. Chem.*, **74**(16), 3963–3971.

Elliger, C. A., Wong, R. Y., Waiss, A. C., *et al.* (1990) *J. Chem. Soc., Perkin Trans. 1*, **1**(3), 525–531.

Fabricant, D. S., and Farnsworth, N. R. (2001) *Environ. Health Perspect.*, **109**, 69–75.

Fujioka, T., Kashiwada, Y., Kilkuskie, R. E., *et al.* (1994) *J. Nat. Prod.*, **57**(2), 243–247.

Glauser, G., Guillarme, D., Grata, E., *et al.* (2008) *J. Chromatogr. A*, **1180**(1–2), 90–98.

Glauser, G., Gindro, K., Fringeli, J., *et al.* (2009) *J. Agric. Food Chem.*, **57**(4), 1127–1134.

Gragg, G. M. N.D.J. (2006) Natural products sources of drugs: plants, microbes, marine organisms and animals, in *Comprehensive Medicinal Chemistry II*, 1st edn, eds. P. Kennewell, D. Triggle and J. Taylor, Elsevier, Oxford, pp. 355–403.

Gupta, P. B., Onder, T. T., Jiang, G., *et al.* (2009) *Cell*, **138**(4), 645–659.

Hadacek, F., and Greger, H. (2000) *Phytochem. Anal.*, **11**(3), 137–147.

Haffter, P., Granato, M., Brand, M., *et al.* (1996) *Development*, **123**, 1–36.

Hammerschmidt, R. (1999) *Annu. Rev. Phytopathol.*, **37**(1), 285–306.

Harvey, A. L., and Cree, I. A. (2010) *Planta Med.*, **76**(11), 1080–1086.

Hostettmann, K., Marston, A., and Hostettmann, M. (1998) *Preparative Chromatography Techniques: Applications in Natural Product Isolation*, Springer, Berlin/New York.

Huber, L., and George, S. A. (1993) , in *Diode Array Detection in HPLC*, eds. L. Huber and S. A. George, Marcel Dekker, New York, p. 400.

Huh, D., Torisawa, Y. S., Hamilton, G. A., *et al.* (2012) *Lab Chip*, **12**(12), 2156–2164.

Ioset, K. N., Nyberg, N. T., Van Diermen, D., *et al.* (2011) *Phytochem. Anal.*, **22**(2), 158–165.

Jaroszewski, J. W. (2005) *Planta Med.*, **71**(9), 795–802.

Kartal, M. (2007) *Phytother. Res.*, **21**(2), 113–119.

Kingston, D. G. I., and Newman, D. J. (2005) *IDrugs*, **8**(12), 990–992.

Klen, T. J., and Vodopivec, B. M. (2011) *J. Agric. Food Chem.*, **59**(24), 12725–12731.

Koehn, F. E., and Carter, G. T. (2005) *Nat. Rev. Drug Discovery*, **4**(3), 206–220.

Konkel, M. E., and Tilly, K. (2000) *Microbes Infect.*, **2**(2), 157–166.

Korn, K., and Krausz, E. (2007) *Curr. Opin. Chem. Biol.*, **11**(5), 503–510.

Kuzel, S., Vydra, J., Triska, J., *et al.* (2009) *J. Agric. Food Chem.*, **57**(17), 7907–7911.

Lam, H. W., Lin, H. C., Lao, S. C., *et al.* (2008) *J. Cell. Biochem.*, **103**(1), 195–211.

Lasakova, M., and Jandera, P. (2009) *J. Sep. Sci.*, **32**(5–6), 799–812.

Lee, I. S., Ma, X. J., Chai, H. B., *et al.* (1995) *Tetrahedron*, **51**(1), 21–28.

Lemma, A. (1970) *Bull. W. H. O.*, **42**(4), 597–612.

Li, Y. F., Plitzko, I., Zaugg, J., *et al.* (2010) *J. Nat. Prod.*, **73**(4), 768–770.

Liao, X. P., and Shi, B. (2005) *J. Sci. Food Agric.*, **85**(8), 1285–1291.

Liu, M., Seidel, V., Katerere, D. R., *et al.* (2007) *Methods*, **42**(4), 325–329.

Long, Z., Wang, C., Guo, Z., *et al.* (2012) *Analyst*, **137**(6), 1451–1457.

Lu, Y., Hernandez, P., Abegunde, D., *et al.* (2011) *The World Medicines Situation 2011*, World Health Organization, Geneva, Switzerland.

McCloud, T. G. (2010) *Molecules*, **15**(7), 4526–4563.

McLaughlin, J. L. (2008) *J. Nat. Prod.*, **71**(7), 1311–1321.

Mir-Rashed, N., Cruz, I., Jessulat, M., *et al.* (2010) *Med. Mycol.*, **48**(7), 949–958.

Mottaleb, M. A., and Sarker, S. D. (2012) *Methods Mol. Biol.*, **864**, 75–87.

Mylonakis, E., Moreno, R., El Khoury, J. B., *et al.* (2005) *Infect. Immun.*, **73**(7), 3842–3850.

Natishan, T. K. (2004) *J. Liq. Chromatogr. Relat. Technol.*, **27**(7–9), 1237–1316.

Newman, D. J., and Cragg, G. M. (2012) *J. Nat. Prod.*, **75**(3), 311–335.

Newman, D. J., Cragg, G. M., and Snader, K. M. (2000) *Nat. Prod. Rep.*, **17**(3), 215–234.

Ngombe, N. K., Kalenda, D. T., Quetin-Leclercq, J., *et al.* (2010) *Nat. Prod. Commun.*, **5**(3), 369–372.

Nguyen, A. T., Feasley, C. L., Jackson, K. W., *et al*. (2011) *Retrovirology*, **8**.

Nguyen, D. T. T., Guillarme, D., Rudaz, S., *et al*. (2006) *J. Sep. Sci.*, **29**(12), 1836–1848.

Niessen, W. M. A. (2006) *Liquid chromatography – Mass Spectrometry*, Taylor & Francis, Boca Raton.

Nogueira, R. C., Oliveira-Costa, J. F., de Sa, M. S., *et al*. (2009) *Curr. Drug Targets*, **10**(3), 291–298.

Nweze, E. I., Mukherjee, P. K., and Ghannoum, M. A. (2010) *J. Clin. Microbiol.*, **48**(10), 3750–3752.

Olson, D. L., Norcross, J. A., O'Neil-Johnson, M., *et al*. (2004) *Anal. Chem.*, **76**(10), 2966–2974.

Omolo, J. J., Maharaj, V., Naidoo, D., *et al*. (2012) *J. Nat. Prod.*, **75**(10), 1712–1716.

Ostrosky-Zeichner, L., Casadevall, A., Galgiani, J. N., *et al*. (2010) *Nat. Rev. Drug Discov.*, **9**(9), 719–727.

Parish, L. C., and Parish, J. L. (2008) *Drugs Today*, **44**(2), 91–102.

Patton, E. E., and Zon, L. I. (2001) *Nat. Rev. Genet.*, **2**(12), 956–966.

Poulev, A., O'Neal, J. M., Logendra, S., *et al*. (2003) *J. Med. Chem.*, **46**(12), 2542–2547.

Priestap, H. A., Torres, M. C., Rieger, R. A., *et al*. (2012) *Chem. Res. Toxicol.*, **25**(1), 130–139.

Rahalison, L., Hamburger, M., Hostettmann, K., *et al*. (1991) *Phytochem. Anal.*, **2**(5), 199–203.

Ramandi, N. F., Najafi, N. M., Raofie, F., *et al*. (2011) *J. Food Sci.*, **76**(9), C1262–C1266.

Rausch, O. (2006) *Curr. Opin. Chem. Biol.*, **10**(4), 316–320.

Raza, M. (2006) *J. Ethnopharmacol.*, **104**(3), 297–301.

Ringbom, T., Huss, U., Stenholm, A., *et al*. (2001) *J. Nat. Prod.*, **64**(6), 745–749.

Robinette, S. L., Bruschweiler, R., Schroeder, F. C., *et al*. (2012) *Acc. Chem. Res.*, **45**(2), 288–297.

Rodrigues, C. M., Rinaldo, D., Sannomiya, M., *et al*. (2008) *Phytochem. Anal.*, **19**(1), 17–24.

Romanik, G., Gilgenast, E., Przyjazny, A., *et al*. (2007) *J. Biochem. Biophys. Methods*, **70**(2), 253–261.

Saklani, A., and Kutty, S. K. (2008) *Drug Discov. Today*, **13**(3–4), 161–171.

Santra, S., and Baumann, U. (2008) *Expert Opin. Pharmacother.*, **9**(7), 1229–1236.

Segalat, L. (2007) *ACS Chem. Biol.*, **2**(4), 231–236.

Sessa, C., Lorusso, P., Fran, A. T., *et al*. (2013) *Clin Cancer Res*, **19**(17), 4832–4842.

Sevenet, T. (1991) *J. Ethnopharmacol.*, **32**(1–3), 83–90.

Shi, L. S., Liao, Y. R., Su, M. J., *et al*. (2010) *J. Nat. Prod.*, **73**(7), 1214–1222.

Siller, G., Gebauer, K., Welburn, P., *et al*. (2009) *Australas. J. Dermatol.*, **50**(1), 16–22.

Siqueira-Neto, J. L., Song, O. R., Oh, H., *et al*. (2010) *PLoS Negl. Trop. Dis.*, **4**(5), e675.

Trancassini, M., Ghezzi, M. C., Cipriani, P., *et al*. (1986) *Chemioterapia*, **5**(1), 14–17.

Tsai, S. F., and Lee, S. S. (2010) *J. Nat. Prod.*, **73**(10), 1632–1635.

Tu, Y., Jeffries, C., Ruan, H., *et al*. (2010) *J. Nat. Prod.*, **73**(4), 751–754.

Vasconsuelo, A., and Boland, R. (2007) *Plant Sci.*, **172**(5), 861–875.

Vermeulen, L., de Sousa e Melo, F., Richel, D. J., *et al*. (2012) *The Lancet Oncol.*, **13**(2), e83–e89.

Wall, M. E., and Wani, M. C. (1996) *J. Ethnopharmacol.*, **51**(1–3), 239–253.

Wan, J. B., Zhang, Q. W., Hong, S. J., *et al*. (2012) *Molecules*, **17**(5), 5836–5853.

WHO (2005) *WHO Traditional medicine strategy 2002–2005*, World Health Organization, Geneva, Switzerland.

Wijngaard, H., and Brunton, N. (2009) *J. Agric. Food Chem.*, **57**(22), 10625–10631.

Wolfender, J. L., Queiroz, E. F., and Hostettmann, K. (2005) *Magnetic Resonance in Chemistry: MRC*, **43**(9), 697–709.

Wolfender, J. L., Marti, G., and Queiroz, E. F. (2010) *Curr. Org. Chem.*, **14**(16), 1808–1832.

Wolfender, J. L., Queiroz, E. F., and Hostettmann, K. (2006) *Expert Opin. Drug Discovery*, **1**(3), 237–260.

Wolfender, J. L., and Queiroz, E. F. (2012) *Chimia*, **66**(5), 324–329.

Wolfender, J. L., Rudaz, S., Choi, Y. H., *et al*. (2013) *Curr. Med. Chem.*, **20**(8), 1056–1090.

Wulff, G., and Sarhan, A. (1972) *Angew Chem., Int. Ed.*, **11**(4), 341.

Xia, E. Q., Cui, B., Xu, X. R., *et al*. (2011) *Molecules*, **16**(9), 7391–7400.

Yang, X., Baburin, I., Plitzko, I., *et al*. (2011) *Mol. Divers.*, **15**(2), 361–372.

Yuliana, N. D., Khatib, A., Choi, Y. H., *et al*. (2011a) *Phytother. Res.*, **25**(2), 157–169.

Yuliana, N. D., Khatib, A., Verpoorte, R., *et al*. (2011b) *Anal. Chem.*, **83**(17), 6902–6906.

Zaugg, J., Eickmeier, E., Rueda, D. C., *et al*. (2011) *Fitoterapia*, **82**(3), 434–440.

Zaugg, J., Baburin, I., Strommer, B., *et al*. (2010) *J. Nat. Prod.*, **73**(2), 185–191.

Zheng, G. Q. (1994) *Planta Med.*, **60**(1), 54–57.

Zhu, L., and Xu, X. (2003) *J. Chromatogr. A*, **991**(2), 151–158.

Zweifel, M., Jayson, G. C., Reed, N. S., *et al*. (2011) *Ann. Oncol.*, **22**(9), 2036–2041.

High Throughput Screening of Vegetal Natural Substances

Bruno David and Frédéric Ausseil

Institut de Recherche Pierre Fabre, Toulouse, France

1 INTRODUCTION

High throughput screening (HTS) is a technology dedicated to the random identification of active chemical molecules either produced by synthesis or extracted from natural sources. Initially used in the pharmaceutical drug discovery domain, its principle is based on the test of large sets of compounds on selected biological targets in a limited period of time, highlighting, therefore, the importance of throughput in the success of this approach (Entzeroth *et al.*, 2009; Macarrón, 2006; Macarrón *et al.*, 2011; Mayr and Bojanic, 2009). Then, among the numerous compounds tested, the probability to find active ones is high. Typically, several ten or hundred thousand compounds are evaluated on various biological activities. Response rates ranging from 0.05% to 1% are generally observed and lead to the quasi-systematic identification of actives compounds on the considered bioassay.

Despite the development of *in vivo*-based approaches, focused on small organisms such as the worm *Caenorhabditis elegans* or the zebrafish *Danio rerio* (Leung *et al.*, 2011; Delvecchio, Tiefenbach, and Krause, 2011), the majority of the biological activities currently used in HTS approaches are *in vitro* based, such as, for example, enzymatic activities, ligand binding, and cell functions or phenotypes. To reach an adequate throughput, these assays are routinely repeated in large numbers. Automation has been a great and appreciated help to facilitate this point. Thanks to the development of the 96-well microplate format and dedicated laboratory devices such as readers, pipetors, filtration stations, and cell harvesters, automation became essential for HTS. Specific laboratories and fully integrated screening systems were built with inexorably growing capacities (Wunder *et al.*, 2008). Chemical compound libraries should have been adapted to this evolution including compound formatting, registration, storage, and, more importantly, the way they were obtained (Edwards, 2006). From pharmaceutical company corporate compound collections to high throughput parallel or combinatorial synthesis, the evolution was, as for biology, supported by the idea of high efficiency to boost preclinical drug discovery programs. One of the drawbacks of this is the "screen-time window," defining the period of time necessary for the screening of a particular substance collection of compounds on a chosen target.

The intensive use of HTS enabled the identification of many active compounds that have entered in the "hit to lead" process and latter, after exhaustive pharmacological and chemical optimization in the "lead to candidate" one (Macarrón and Hertzberg, 2011). Academia also found an interest in this approach by considering hits as chemical probes to explore living organism structures and functions and

to initiate new drug discovery approaches (Frearson and Collie, 2009; Michael *et al*., 2008). However, "hard to screen" targets and compounds appeared in the screening landscape and complicated the development of high efficiency screening organizations (Meireles and Mustata, 2011). Some of these "black sheeps" were progressively turned down from many discovery initiatives. Natural molecules tend to belong to this last category essentially because the discovery and the validation of an interesting bioactive molecule created by nature are based on different synthesis rules than those into effect for molecules synthesized in laboratories. In the case of bioactive natural products, structural complexity very often prevents its total or hemisynthesis. The structural characterization and synthesis remain a scientific challenge. Second, the access to natural sources is strongly regulated to avoid legal misappropriation (the so-called biopiracy). Collecting in accordance with national and international laws is then of primary importance but often very long. Furthermore, identifying the organisms, extracting in appropriated conditions, and conserving samples in good storage conditions are practical matters that need to be carefully studied and optimized. Thus, the return on investment of natural product-based researches could certainly not be ranked as "AAA" by the top management of pharmaceutical companies even if a large proportion of the drugs on the market have a direct or indirect natural origin (Newman and Cragg, 2012; Bailly, 2009). In these conditions the agreement between the HTS "screen-time window" and the natural compound discovery process could appear as improbable. However, could the exploration of biological and therapeutic potency of the quasi-unlimited biodiversity of living organisms be possible without a tool like HTS? Which strategies can be developed to overcome these difficulties? In this chapter, we will try to answer these questions after a brief description of the discovery of natural product and the HTS technology.

2 VEGETAL NATURAL PRODUCTS FOR HTS

2.1 Interest of Vegetal Natural Substances

The total number of higher or vascular plant species on the Earth is close to 260,000 (Thorne, 2002; Scotland and Wortley, 2003), and each species contains hundreds of potentially active compounds. The plants represent a huge library of potentially bioactive compounds, with over 1 million chemical entities, between 10 and 100 millions according to Verpoorte (2010). Compared to other sources (marine organisms, microorganisms, and insects), plants exhibit a high metabolic prodigality associated to a vast chemodiversity. They are considered an accessible source for the first sampling compared to insects, marine organisms, and microorganisms because they are fixed terrestrial and generally large size organisms. Historically, the success of their use is correlated in practice with easy harvest or agronomic production in fields (Appendino and Pollastro, 2010). When larger quantities and scale-up are needed, microorganisms are by far generally the easiest source.

The plant natural products have many advantages: availability, proof of concept of their interest for medicine since more than thousands of years, high chemical diversity, and big numbers.

2.1.1 *Proof of Concept*

The importance of plants as sources of historical and current effective drugs for human beings is regularly highlighted by the World Health Organization. The mode of action, affinity, and high chemical diversity enabled them to provide a significant portion of the active ingredients in the therapeutic arsenal. Countless examples mark the history of medicine since the use of morphine from *Papaver somniferum*, to artemisinin from *Artemisia annua* through antitumor vinca alkaloids from *Catharanthus roseus*.

Since forever, plants have biosynthesized metabolites to address the macromolecules involved in their metabolism. Secondary metabolites of plants have therefore coevolved to optimize interactions with molecules of life: proteins, enzymes, receptors, ionic channels, and so on. They have the advantage of being made to be active on biological systems. They play a role in the plant where they cross membranes, interfere with metabolic enzymes, or fight against predators and parasites. The products of the chemical factory plant are, therefore, potentially more interesting for biology than compounds created randomly by organic chemists. According to Schreiber (Rouhi, 2003), "Natural products are highly evolved, highly specific, and can be highly effective toward the gene

products with which they coevolved." Meinwald adds in the same reference "Natural products have evolved to interact with something, and that something may not be so different from human proteins."

While the primary metabolites play defined roles in the growth and development of plants, it seems likely that secondary metabolites have evolved specifically with and to modulate the functions of biological macromolecules. The role of many of them is not very clear. It is even possible that the majority of secondary metabolites do not play a biological role. However, the evolution may have favored their preservation as extensive chemical diversity. The presence of an enzymatic complex environment and perpetual changes in the biosynthetic pathways may ensure that, when a natural product has biological activity, it provides an advantage to conserve it during evolution. The fact that many natural products selectively bind to human targets and modulate their activity is not surprising. Most of these human targets have counterparts in plant organisms (Newman *et al.*, 2002). The biological activity of natural products in human cells is a consequence of their coevolution along with similar targets. This makes the plants' natural products extremely effective in the targeting of biological macromolecules (Dixon *et al.*, 2007; Carlson, 2010).

But natural products are probably not useful for all targets. Being preferentially defense substances, they target the vital processes. This is probably the reason why they are present more in cancer chemotherapy than in other therapeutic areas (Dančík *et al.*, 2010).

2.1.2 Number and Chemodiversity

Each year, more and more natural plant substances are described. Advances in analytical techniques and the chemical richness of plants not yet studied explain why, to date, more than 200,000 natural plant molecules have been described (authors estimate), cf. Figure 1.

Vascular plants produce a great skeletal and functional diversity because of a large genome (>100,000 genes) in comparison to bacteria (1000 genes) and fungi (10,000 genes) (Devlin, 1997). This inventory is far to slow down and to be over. Admittedly, herbal substances belong to chemical families known since the nineteenth century, but the number of representatives of these families seems almost infinite and minor chemical structural differences can confer big pharmacological differences. Vegetal molecules address different and larger area of chemical space than synthetic compounds and are amenable to further chemical improvement (Ganesan, 2008) (see **Metabolomics**).

2.2 Access to Plant Samples

The legal acquisition of the samples is a pivotal point even if commercial issues, regulations on protected species, and phytosanitary formalities are also crucial. According to the target 16 of Aichi Biodiversity Targets (http://www.cbd.int/sp/targets), "by 2015, the Nagoya Protocol on Access to genetic resources
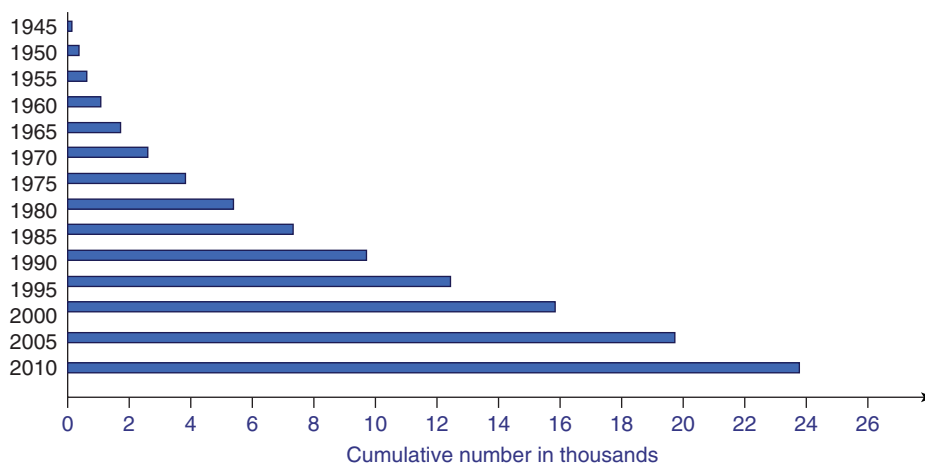
Figure 1 Cumulative estimated number of natural products described (authors' estimate).

and the fair and equitable sharing of benefits arising from their utilization is in force and operational, consistent with national legislation." That is to say that all public or private researchers will be obliged to seek prior informed consent (PIC) and to negotiate a mutually agreed term (MAT) with the competent national authority (Oliva, 2011). This will soon become necessary by the implementation in different countries of the Nagoya Protocol. A plant that will be collected in a personal garden and will be studied in a national research and development center will be considered a national property and should go through the PIC and MAT processes.

Since December 29, 1993, date of application of the Convention on Biological Diversity (CBD), the plants have been under the sovereignty of the states that host them. The CBD is one of the two major agreements signed by the international community at the Earth Summit, held in Rio de Janeiro (Brazil) in June 1992. The other was the United Nations Framework Convention on Climate Change. The objectives of the CBD are to conserve biodiversity, use biological resources sustainably, ensure that we do not use them faster than they can regenerate, and share the benefits arising from the use of genetic resources in a fair and equitable way. According to Article 2 of the CBD, biodiversity is defined as "the variability among living organisms from all sources of terrestrial, marine, and other aquatic ecosystems and ecological complexes. This includes ecosystem diversity, species diversity between and within species." When the recent Nagoya Protocol gets implemented, any plant sample used in the process of R&D will go through a PIC and MAT procedures, specifying the conditions of Access and Benefit Sharing with the authorities of the country source, cf. Figure 2. Until now, the inclusion of biological resources in the scope was subjected to interpretation. The Rio Convention was dealing only with access to genetic resources (i.e., resources carrying functional units of heredity) and not the biological resources (i.e., secondary metabolites). One of the interests of the Nagoya Protocol is to bring clarity on the scope; now research and development on secondary metabolite are clearly "in" according to Articles 2c, 2d, and 2e. Art 2c: "Utilization of genetic resources" means to conduct research and development on the genetic and/or biochemical composition, including through the application of biotechnology as defined in Article 2 of the Convention. Art 2d: "Biotechnology" as defined in Article 2 of the Convention means any technological application that uses biological systems, living organisms, or derivatives thereof, to make or modify products or processes for specific use. Art 2e "Derivative" means a naturally occurring biochemical compound resulting from the genetic expression or metabolism
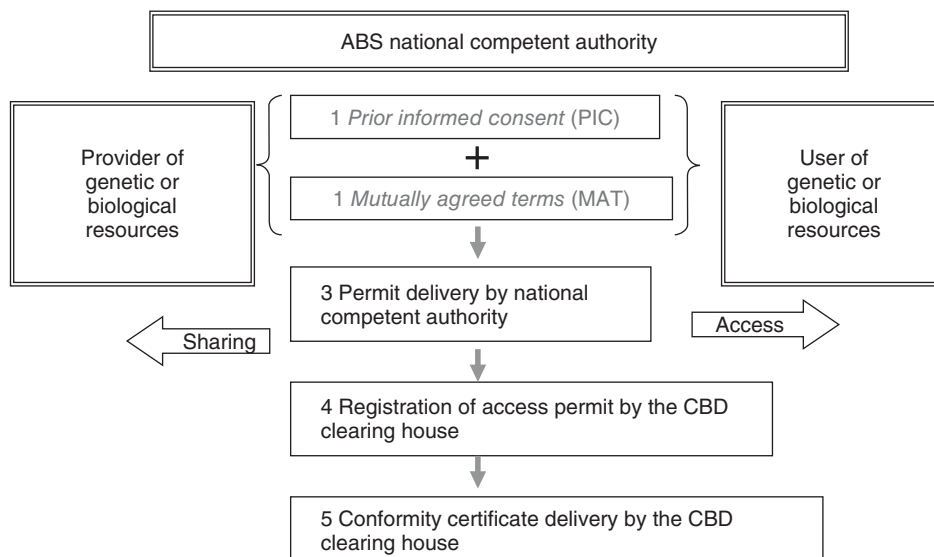
Figure 2 Mechanism and future steps required to access to natural samples.

of biological or genetic resources, even if it does not contain functional units of heredity.

In practice, bioprospecting laws have not yet had the desired effect for many reasons:

- Most developing countries have not established regulations. Less than 30 countries have real relevant legislation to date (Schindel, 2010).
- When biodiversity-rich countries have implemented national regulations, their practical application and interpretation are often difficult, and the rules are not practical. There is a gap between theory and practice (Kingston, 2011).
- Most of the time, access is often very long and requires several years of negotiations.
- There is still a suspicion around bioprospecting in many countries anchored on colonial history and feeling of exploitation of national natural resources.
- Because expectations of Benefit Sharing of biodiversity-rich countries being generally too high, there is a large asymmetry between these expectations and returned opportunities from the pharmaceutical industry and cosmetics. The historical contract between Merck and INBio and the exceptional turnover generated by some rare natural products such as taxoids or camptothecinoids, unfortunately, did not provide an image of realistic possible returns from pharmaceutical manufacturers. The likelihood of such success is extremely low considering that the NCI had to work on 114,000 extracts from 12,000 different species to find only these two molecules (Kingston, 2011).

2.3 Botanical Issues

Regarding plant collection, botanical identification is another key point, which cannot be automated (Miller, 2011). It is important to supply plants that can be identified (presence of flowers or fruit) and harvested again when required. Therefore, rare, very small, and, of course, endangered species will not be collected.

It is necessary to rely on botany specialists to reference plants unambiguously. The advantages are at least three:

- Avoid duplication (a botanical species may have multiple synonyms, these names can be more or less correctly spelled, and so on);
- Unique indexing of a species permits to avoid accessing to a species from different countries; this simplifies the ABS issues,
- Correct names allow fruitful utilization of databases correlating species names and chemical composition. Botanical identification enables dereplication and benefit of the scientific literature. This supposes that the species are correctly identified and described under the valid name (Erkens, 2011). Very often, more or less complex problems of synonymy complicate this task and have to be considered when checking bibliographic references. Classical systematic botany is a science that seems outdated in the middle of the new "omics" techniques, but its importance is paramount. The disappearance of taxonomy botanists and the lack of formation of young scientists are alarming. (see **Selection, Identification, and Collection of Plants** and **Identification of Medicinal Plants Using DNA Barcoding**).

The composition of plant material naturally varies according to seasons, soils, climates, and geographical areas, but the presence of endophytic organisms is increasingly considered to explain the variability in chemical composition (Zhang, Song, and Tan, 2006).

2.4 Plant Libraries Building

Most of the time, the collection needs to be as large and as diverse as possible to get the best sampling. The chemotaxonomy teaches us that if one wishes to obtain an optimal sampling of plant secondary metabolites, it is necessary to get the highest diversity in superior taxa (classes, orders, families, and genera). In some cases, a collection may consist of one particular group of plants, example of Pteridophytes for the interest in protoflavonoids (Pouny *et al*., 2011). It is necessary to treat each organ of a species separately as it contains its own chemical composition (Figure 3).

Herbarium vouchers, sample of leaf for DNA analysis preserved in dry silicagel, and associated data must be systematically linked to the sample collected for screening. This brings security and quality in the process.

Figure 3 Dry ground samples library, Institut de Recherche Pierre Fabre, Toulouse, France. *Source*: Photo supplied by D. Cabrol, Pierre Fabre Laboratories. Reproduced with permission.

Scientific names and families have to be mentioned in one fixed standardized system. Associated data have to be recorded in a data base, cf. Section 2.6.

The Latin binomial name is very useful for structural determination. Use of this scientific name with databases such as the Dictionary of Natural Products allows a good start in structural determination providing structural and substructural information. Screening small amounts of plant material represents only a first step. For *in vivo* confirmations and medicinal chemistry of analogs, kilograms of materials are required. These recollections for further studies can be disconcerting because human pressure would have removed the plants with their biotope. Few months or years after collection, the plant might have disappeared because of the destruction of its habitat reconverted in fields, urban developments, or desert lands.

2.5 Building Natural Chemical Libraries

Ideally, it would be best to test highly purified fractions or even pure compounds, but, in reality, owing to technical issues and costs, it is hardly possible to isolate one by one all the different secondary metabolites. Most of the time, a medium polar solvent such as dichloromethane (CH_2Cl_2) or ethylacetate (AcOEt) is used for extraction. Interfering components are removed by solid phase extraction (SPE) cartridge. This process can be highly automated. Advantages and disadvantages of each strategy are presented in Table 1.

The solvents used in extraction and the postextraction procedures are designed to get rid of the undesired class of compounds presents in plants (Devlin, 1997) (Table 2) (see **Extraction Methodologies: General Introduction** and **New Trends in Extraction of Natural Products: Microwave-Assisted Extraction and Pressurized Liquid Extraction**).

Table 1 Advantages and disadvantages of the different options for building natural chemical libraries.

	Pure compounds	Purified fractions	Purified extracts	Crude extracts
Advantages	– High quality pharmacological results; – Hit comparison possible; – Chemical structure accessible; – Comparable to conventional chemical libraries; – No interference.	– Improves the chances of discovering a minor active component; – Automatable process; – Isolation of bioactive compounds easier; – No interfering compounds.	– Less interferences; – Minority products available; – Automatable process; – Easy to prepare; – Libraries of reasonable size.	– Libraries of reasonable size; – Largest chemodiversity; – Low preparation cost.
Disadvantages	– Burdensome, lengthy processes providing few compounds; – Time and money consuming isolation and dereplication before HTS; – The minor products (often new) are not isolated as a priority.	– High cost of screening related to libraries of large size (numerous fractions); – Redundancy of compounds; – Dereplication before HTS impossible; – Bioguided fractionation needed.	– Complex chemical composition; – Some interfering compounds; – Important bioguided fractionation needed.	– Bad quality pharmacological results (soup effect); – Interfering compounds all present; – Results exploitation difficult; – Isolation process from the very beginning; – Important bioguided fractionation needed.

Table 2 HTS and interest for the different plant metabolites.

Interference/nuisance compounds	Welcomed compounds	Preferred compounds
Lipids	Monofunctional terpenes	Alkaloids
Tannins	Oligosaccharides	Multifunctional terpenes
Proteins	Flavonoids	Multifunctional sterols
Polysaccharides	Quinones	Lignans
Monosides	Amino acids	Macrocycles
Essential oil compounds	Peptides (cyclic and acyclic)	

2.6 Data Management

A dedicated system of data management with inventory management of the sample library (bar code), botanical data, and HTS data is essential. The system needs to record in association to a unique ID: date of collection, ecosystem, country of collection, GPS localization, phenology (sterile, presence of flower, and fruits), biological type (tree, shrub, climber, etc.), short description, characteristic features, species and family names, photographs of the plant, herbarium vouchers, associated pharmacological data on extracts, purified fractions and isolated pure compounds, and so on.

Botanical records are important even if, in theory, isolation of an active substance and structure determination provide sufficient information for a hemisynthesis or a total synthesis, which are virtually always possible, making the reharvesting not necessary. Such data may be, in some cases, of paramount importance to recollect the material in the wild and/or to allow a revision of the botanical determination.

2.7 Summary of Constraints

Several difficulties are significantly hampering the research: never ending or lengthy negotiations of supply contracts (to meet the Rio Convention and Access and Benefit Sharing regulations) (Cordell, 2010), the field survey, identification, and harvesting. Plant samples contain mixtures of several hundred substances, making isolation tedious of the active product through bioguided fractionation or dereplication. The relatively short screen time windows are a point particularly detrimental to natural products, which need iterative bioguided isolation. Mixtures complicate the experiments because many pharmacological substances will interfere more

or less specifically. When the isolation of the active substance has been carried out, its chemical structure needs to be determined, and total synthesis is not always economically accessible (Paterson and Anderson, 2005). Natural products need to be derivatized in order to be protected by patents; patents are also a big issue (Demaine and Fellmeth, 2003). At first glance, everything seems easier for the compounds obtained by combinatorial chemistry, which can furnish very large collections. They are simply and quickly supplied, resupplied at low cost, and the scaling-up is not really a problem. Isolation of plant natural products is a long and fastidious process at an unpredictable cost. It is practically impossible to isolate one by one all the molecules from one plant to screen them. Even if this is feasible, dereplication and structural identification (which cannot be automatic) will require costly investments.

2.8 Interest of Natural Product Screening for Pharmaceutical Research

Drug discovery involving plants and HTS has not been very productive because it adds up problems and specific difficulties. All this takes time compared to conventional medicinal chemistry research.

Ironically, and probably by a causal effect, when the international community relied on benefit sharing to finance conservation and sustainable use of biodiversity, companies started to withdraw from this sector. In the late 1990s with the development of combinatorial chemistry and ultra-HTSs, natural substances have become less attractive for the pharmaceutical industry. Chemical optimization is intrinsically a source of problems when working on natural compounds. This was exacerbated when access started to be difficult and legally uncertain. To this, we must add that, in the cancer field, the key areas

for vegetal natural products (Newman and Cragg, 2012), monoclonal antibodies and synthetic kinase inhibitors, are gaining rapid ground.

The technical and legal difficulties, particularly in terms of biodiversity access, have diverted (Cf. Figure 13) many multinational pharmaceutical companies such as Glaxo, Merck, and Pfizer, from plants-based research programs for the benefit of combinatorial chemistry. However, the usefulness of combinatorial chemistry has not yet been demonstrated by the marketing of numerous new drugs. Only the molecule sorafenib was discovered by combinatorial chemistry (Newman and Cragg, 2012).

3 IMPLEMENTING HTS TO DISCOVER BIOACTIVE COMPOUNDS IN PLANTS

3.1 General Observations on HTS

HTS is at early stage of the drug discovery process following target identification and target-related bioassay development steps (Mayr and Bojanic, 2009). The underlying idea associated to this technology is that a chemical modulator of the target should be discovered rapidly by using a random and systematic evaluation of as many as chemical sources available in time. As ten to hundred thousands of molecules and natural extracts were accessible in pharmaceutical companies for screening, experiment automation became rapidly an issue to guarantee an acceptable throughput in the whole process. Then, HTS has been defined as a technology that "enables a biochemical or cellular event to be reproducibly and rapidly tested against chemical entities many hundreds of thousands of times" (Avery *et al*., 2010).

The output of HTS is called a *hit*, which is the response of one chemical source to one biological event. Hits should be controlled, confirmed, and validated to become significant for drug discovery. Hit structure optimization and testing on more bioassays will conduct to a new entity called *lead*, defining a valuable starting point to generate a drug candidate. Two main advantages of HTS ensured its success in the 1990s. First, little knowledge of the target is sufficient to start the process. This means that target (i.e., enzymes, transmembrane receptor, structural protein folding, and protein–protein interactions) structures are not a prerequisite to identify active molecule.

Second, the random aspect of the technology generates unpredictable, and therefore potentially original, active chemical structures avoiding "me-too" approaches. However, the two main drawbacks of HTS are attrition and interferences. Attrition means that, if relevant hits are relatively easy to obtain, their transformation into leads and further into drug candidate concerns an extremely low percentage of them. Interferences mean that certain hits obtained from a screening are not related to target modulation but to unspecific events occurring within the bioassay. Autogenerated aberrant signals or quenching interferences are very often observed and have a strong impact on technology practicing and experiment quality control (Thorne, Auld, and Inglese, 2010).

3.2 Automation

Testing 100,000 compounds by hand would be dramatically time consuming and source of errors. Automation is a precious ally to solve intensive experiment repetitions in minimal periods of time. However, despite a preexisting environment of automated biology in the medical diagnostic industry, the development of "tried and trusted" screening stations was a long road paved of many system wrecks and software dead-ends. Nevertheless, robust stations emerged and enabled the achievement of well-controlled screening campaigns. These automats are not designed like "one-piece machines" but rather as automated laboratories with modules dedicated to specialized functions physically connected to each other (Figures 4 and 5).

The choice and the arrangement of the modules (i.e., their integration) depend on the screening strategies: oriented toward cell-based assays or ELISA, for example, or filtration with radioactive readouts or GPCR. Natural product screening should be carefully considered at this point to make a wise choice in module integration. Precise and versatile pipetors, tips versus needles, versatile and robust plate washers, multimode readers, potent plate shakers, and centrifugation are options that are particularly of interest for the screening of extracts. Similarly, the type of targets should be also considered. A screening strategy based on a unique target or target type (e.g., GPCRs or kinases) will not be turned to automation similar to an open multitarget screening platform.
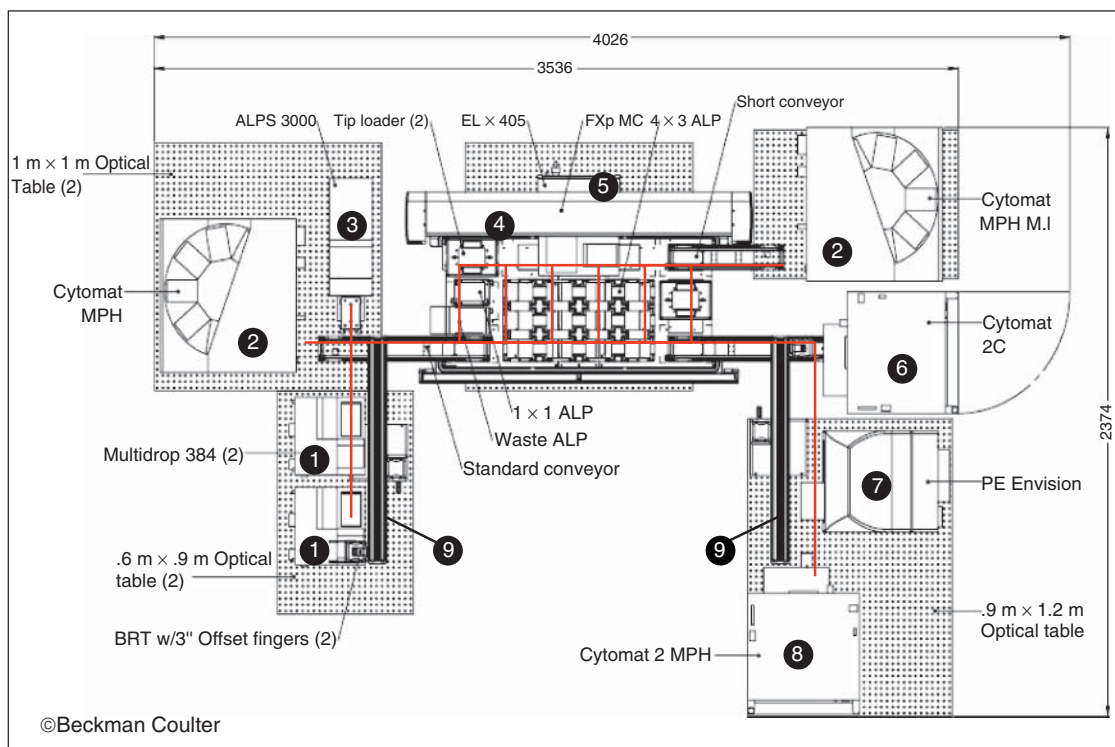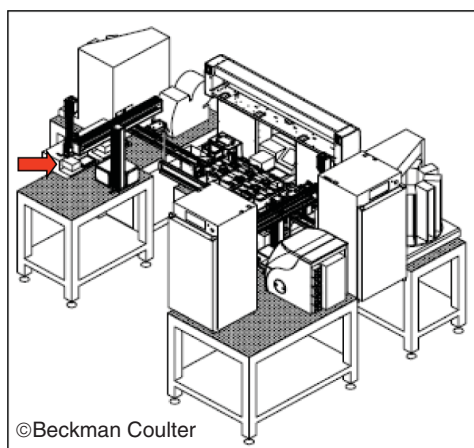
Figure 4 Screening station: (a) top view and metric dimensions of an integrated HTS system containing various modules: (1) reagent injectors, (2) plate and tips storage, (3) plate sealer, (4) pipetor, (5) plate washer, (6) plate and tips storage, (7) plate incubator, (8) multimode reader, (9) plate transporter. Red line: plate displacements. All modules are physically accessed by plates. (b) 3D view of the system. The red arrow shows the view axis in C. (c) System in real situation.

Figure 5 Overview of a fully integrated screening station. Operators are protected from solvent vapors by an enclosure connected to a vapor evacuation system. *Source*: Photo supplied by D. Cabrol, Pierre Fabre Laboratories. Reproduced with permission.

3.3 Assay Formats

Without any doubt, the 96-well plate was a founding element for bio-assay automation with volumes ranging from 50 to 250 μL per well. Furthermore, plate format standardization largely contributed to the development and expansion of laboratory modules integrated or not into screening systems (http://www.slas.org/community/microplate.pdf) (Figure 6).

Figure 6 The 96, 384, and 1536 plate formats. In the 96-well format (right), each well occupies approximately a 50 mm² surface enabling automated and manual experiments. This is important for bioassay development, which can be made in the same format as in HTS. In the 384-well format, well surface is close to 12 mm² and still enables the two types of experiments. In the 1536-well format, well surface is close to 3 mm² and only very simple experiments could be run manually on a small part of the plate.

Moreover, the rarity of some biological reagents, such as certain cell types or purified proteins, led to the intensification of assay miniaturization for high throughput analysis. The 384-well (standard and low volume) plate format enabled the development of assay with total volume between 10 and 80 μL. When 1536-well plate format emerged, it became reasonable to develop bioassay in 10 μL and less (Leister *et al.*, 2011). This race to miniaturization was supported by the last main concern in HTS: cost control. The price to pay to run larger HTS campaign, the cost per well should be tightly controlled to maintain the total cost in acceptable ranges.

3.4 Natural Products as a Chemical Source for HTS: Focused Analysis

HTS could appear as the "descendant" of biology, automation, and miniaturization domains. Making all these highly specialized technologies working together routinely is extremely challenging, and despite the tremendous progress made since the early 1990s, the HTS of large compound collections is never straightforward. Pure molecules are generally easier to screen than natural products because of the relative homogeneity of their physicochemical parameters, such as, for example, color, autofluorescence, solubility, and solidification temperature in DMSO solutions. Natural products can be integrated in the screening process as crude and prefractionated extracts, with various levels of prefractionation ranging from 2–3 to more than 100 fractions, and as purified molecules (Bindseil *et al.*, 2001; Johnson *et al.*, 2011). For a screener, pure molecules extracted from plants can be nearly associated to synthetic compounds, but the different forms of natural extracts represent "challenging" chemical sources. Screening station crashes could happen, and if sacrificing few micrograms of thousands of synthetic compounds is always a great pity, things could turn to a disaster when natural sources, obtained after a long, risky, and expensive process, are considered. Consequently, screening operations should be extensively optimized and secured to warrant the highest level of expertise and quality for these precious samples.

The success of any HTS campaign is also dependent on the quality of the screened libraries (Koehn, 2008). A vast and diverse sampling of the plant biodiversity (class, order, family, and genus) is an expensive investment needing specialists to implement legally a functional library associated with resupply capacities. Resupplying an active sample is a key parameter and can cause delay and project failure for nonfermentable biomasses such as plants (Appendino and Pollastro, 2010).

The first step in the screening process consists in understanding not only what the samples are and contain as valuable and innovative compounds but also as undesired or surely interfering chemical plagues. On this point of view, the "*pas de deux*" between screeners and phytochemists described by Bailly should be micron adjusted to offer an irreproachable ballet (Bailly, 2009).

Unlike synthetic molecules, the exact content of extracts is not defined even if certain types of molecules, that is, primary metabolites, polyphenols, and lipids, for example, are mostly eliminated during a preextraction step (Tu *et al*., 2010). Interference compounds that are most of the time removed by SPE cartridges are presented in Figure 2. As hit-molecule structures and potencies are not available immediately after the screening, but generally after weeks of purification, none adjustment of screening conditions can be done "in the flow." Then, the screening of extracts means generally the screening of the whole library without having any data on hit relevance.

As small active molecules present in living organisms need to cross numerous phospholipidic biological membranes before reaching their target, they generally present a level of lipophilicity compatible with membrane crossing. To extract these molecules, organic solvents, such as methanol, ethyl acetate, or chloroform, are commonly used (Tu *et al*., 2010). At the end of the process, they must be carefully eliminated and extracts dried or dissolved in a storage solvent such as DMSO. All these solvents are not "biofriendly," and enzymes, receptors, and prokaryote and eukaryote cells tolerate only very low levels. Thus, solvent residual concentrations should be considered in the design of screening bioassays. Moreover, as for synthetic molecules, the storage solvent should be systematically tested in each bioassay, and any undesired effect considered as a no-go condition.

To ensure a maximum stability of the extract, storage at very low temperatures is very often performed but could lead to a precipitation of many compounds during thawing (Cheng *et al*., 2003; Engeloch *et al*., 2008). Consequently, even more than synthetic molecules, extracts should always be considered as sources of unsoluble compounds associated to false molecule concentrations and sometimes to pipeting troubles.

Hit rate is one of the most important markers of screening run progress. It is tightly associated to extract content and to the concentration used. When false positive extracts, that is, those influencing the detection step of the assay, such as quenchers or signal autogenerators (e.g., autofluorescent), are eliminated, the remaining active extracts do not represent the real hit rate. Numerous classes of compounds can unspecifically interact with biological systems: they are called the *frequent hitters* (Baell and Holloway, 2010). Polyphenols, tannins, and flavonoids, all widely described as plant secondary metabolites, are known to present this major drawback (Zhu *et al*., 1997). Recently, it was discovered that, under some conditions, synthetic molecules can form self aggregate (size 50–400 nm) that are able to sequester enzymes on their surface (Thorne, Auld, and Inglese, 2010). Such colloid-like aggregation can affect also the activity of natural compounds and are pivotal to explain activity of Traditional Chinese Medicine preparations (Zhuang *et al*., 2008).

Thus, the initial extract hit rate does not predict the success of a screening campaign. Further purification, structural determination, and biological retesting are necessary to confirm the relevance of the results.

Hit confirmation is the final point of a screening operation. Even if multiplets and/or multidose screenings are performed, the secondary and independent testing of natural source bioactivities reinforces the robustness of the primary discovery process. Bioactivity can be reevaluated in a new experiment, but the reproducible preparation of the sample source can be not so simple. A new extraction of the same batch of plant initially active or a new batch from the same species can be tested. If the within-batch reproducibility should be obvious, the inter-batch reproductibilty is not guaranteed. Seasons, age of the organ, sun exposure, soil composition, symbiotic or parasitic fungi, bacteria or insects, and many other parameters can modify secondary metabolite composition

among individuals of the same species. Hit confirmation must be analyzed regarding to these different situations.

3.5 Choice of the Biological Targets for Plant Extract HTS

Both molecular and cellular targets can be addressed through HTS. On one hand, molecular targets have the advantage of being well characterized. Structure–activity relationships are generally performed more easily in this case. On the other hand, it is reductionist to consider that a single target can be sufficient to control a particular disease and the risk of choosing chemically intractable or irrelevant targets is significant and can lead to unsuccessful screening operations. Cellular assays involve more complex biological mechanisms (phenotypes, pathways, and functions), including large populations of targets. However, if it is relevant to the disease, this "black box" can complicate molecule optimization.

Molecular targets include enzymes, extra- or intracellular receptors, and cell structure proteins. Many molecular targets led to the identification of active molecules from plants by HTS. The proteasome, a key player in cancer cell proliferation, is a multiprotein complex responsible for protein homeostasis. Natural products were extensively studied on this target, and both bacteria and plants were shown to be sources of valuables inhibitors such as salinosporamide, extracted from the marine actinomycete *Salinispora tropica*, and pristimerin, extracted from the *Maytenus senegalensis* plant (Feling *et al*., 2003; Tiedemann *et al*., 2009). Genetic integrity represents also a class of targets explored by HTS on plant extracts. The masticadienonic acid, extracted from the *Pistacia lentiscus*, and the acylphloroglucinol derivative mahureones were shown to inhibit DNA polymerase beta in the micromolar range (Boudsocq *et al*., 2005; Massiot *et al*., 2005). Telomerase and topoisomerase enzymes are also targeted by plant-derived molecules (Eitsuka *et al*., 2004; Kawatani *et al*., 2011). Protein processing is also targeted by extracts and molecules obtained through HTS. Parthenolide derivatives, extracted from *Vernonia perrottetii*, and meroterpene derivative dichrostachines were found by screening as tubulin carboxy peptidase and protein farnesyl transferase inhibitors, respectively (Fonrose *et al*., 2007; Long *et al*., 2009).

Cellular targets include cell functions and phenotypes in addition to cell molecular pathway modulation. Many cell viability and proliferation approaches essentially based on two-dimensional cell cultures were adapted to large scale evaluations (Kepp *et al*., 2011). Numerous molecules produced by plants, such as the cytokinin ribosides and ellipticine, were identified there off. (Voller *et al*., 2010; Tian *et al*., 2008). Cell division or apoptosis induction were also explored by HTS leading to the identification of protoapigenone, extracted from the *Equisetum fluviatile* fern, and the proapoptotic eriocalyxin B, extracted from the *Isodon eriocalyx* Chinese herb, respectively (Pouny *et al*., 2011; Han *et al*., 2010). Gene activation through GPCR modulation was also explored by screening a 2,000 ethanol extract collection, leading to the identification of agonists of the human β2-adrenoceptor (Wang *et al*., 2009). Protein homeostasis modulation was also explored by HTS on plant extracts and led to the identification of physalin A, a protein ubiquitination modulator (Ausseil *et al*., 2007; Vandenberghe *et al*., 2008).

The G2 cell cycle check point modulation was also targeted by natural product screening. The psilostachyins A and C, extracted from the common ragweed *Ambrosia artemisiifolia*, were shown to be responsible of the activity without interfering with the microtubule polymerization (Sturgeon *et al*., 2005). Cell heat shock response was also considered to screen pure natural products and extracts (Santagata *et al*., 2012). In this system, a reporter cell line in which expression of enhanced green fluorescent protein is controlled by a thermal stress or exposure to inhibitors of HSP90 or the proteasome was constructed. The screening of 80,000 pure molecules and partially purified natural product extracts enabled the identification of the withaferin A molecule. In the secondary *in vivo* screening using an orthotropic human glioma xenograft model in mice, withaferin A was shown to present a significant effect on mice survival at the 12 mg kg^{-1} dose.

3.6 Bioassay Automation for HTS

Bioassays for HTS must respond to various selection criteria ensuring the feasibility of the campaign (Macarrón *et al*., 2011). These criteria were very

high in the early 1990s, when HTS was an immature technology and led to dramatic restrictions in the choice of assay suitable for automation. The need to develop and use more predictive biological models to fill intelligently the drug discovery pipeline led to a "back to biology" in the 2000s, and to the apparition of the high content screening (HCS) and *in vivo* automated screening.

3.6.1   *Criteria for HTS Assay*

Because it is a massive and risky approach, HTS requires previously that the bioassays fill quality criteria, ensuring the good achievement of screening operations. These criteria cover precision, sensibility, accuracy, robustness, stability, and cost (Macarrón *et al*., 2011). A special focus on robustness, must be made with natural compounds. All these performances markers should be controlled in the development of new assays. To simplify the process, Zhang *et al* introduced the Z' factor, which combine assay precision and sensibility in the following formula (Zhang, Chung, and Oldenburg, 1999).

$$Z' = 1 - 3 \times \frac{\left(\mathrm{SD}_{\mathrm{signal}} + \mathrm{SD}_{\mathrm{background}}\right)}{\left(M_{\mathrm{signal}} - M_{\mathrm{background}}\right)}$$

where SD is the standard deviation, M is the mean, signal is the response for an active assay (100% activity), and background is the response for an inactive assay (0% activity). Because of dimensionless, the Z' factor is largely used to compare a wide diversity of bioassays signals. When a Z' factor more than 0.5 value is calculated from a sufficient dataset, the screening can be considered feasible (regardless many other limitations).

3.6.2   **In Vitro** *Assays*

As shown in Section 3.5, a wide diversity of molecular and cellular targets has been used in HTS of natural products. Another way to analyze HTS assays is supported by the homogenous and heterogeneous phase concept. This approach is particularly

interesting when fully automated screenings are considered. Heterogeneous assays, such as ELISA and filtration-based radioligand binding, were extensively used in the 1990s. Separation steps included in these protocols, that is, plate washing or filtration, were sources of endless robotic problems and dramatically limited the throughput. To overcome these limitations, "mix and read" homogenous phase assays were developed. They are based on physicochemical principles such as the homogenous time resolved-fluorescence energy transfer (Degorce *et al*., 2009) (Figure 7), scintillation proximity (Glickman, Schmid, and Ferrand, 2008), fluorescence polarization (Lea and Simeonov, 2011), enzyme complementation (Lee *et al*., 2007), singlet oxygen transfer (Eglen *et al*., 2008), or fluorescence in microvolume (Lee *et al*., 2003). However, the use of these assays to screen plant extracts is limited by the "mix and read" principle itself as the extract is still present in the medium during the detection step and can generate interferences. The interest of heterogeneous assays was then reconsidered, even if natural compounds and extracts screen have been successful in homogenous procedures such as AlphaScreen, scintillation proximity, fluorescence polarization, enzyme complementation, and HTRF (Tong *et al*., 2010; Jayasuriya *et al*., 2005; Martin *et al*., 2008; Chan *et al*., 2003; Zhao, Yin, and Li, 2010; Dufau *et al*., 2008) (Table 3). In the HTRF case, an optimized kinase assay was used to screen a plant extract collection, and the results were compared with an ELISA-based comparative screening. Authors showed that the HTRF technology produced very reproducible data and less false positive results than ELISA (Dufau *et al*., 2008). Concerning fluorescence polarization, authors demonstrated that the Alexa-Fluor 594 biocytin fluorescent probe has excitation and emission properties compatible with the direct addition and measurement of diluted leaf extracts within the reaction medium (Chan *et al*., 2003). Moreover, Zhao *et al*. screened 352 HPLC-fractionated samples from the ethanol extracts of Chinese medicinal herbs with the enzyme complementation assay in a 96-well format and identified three agonist fractions obtained from *Fructus Aurantii Immaturus*. Secondary cAMP measurements confirmed the agonist effect of the three active fractions.

Distant donor/acceptor —> NO FRET

Distant/acceptor proximity —> FRET
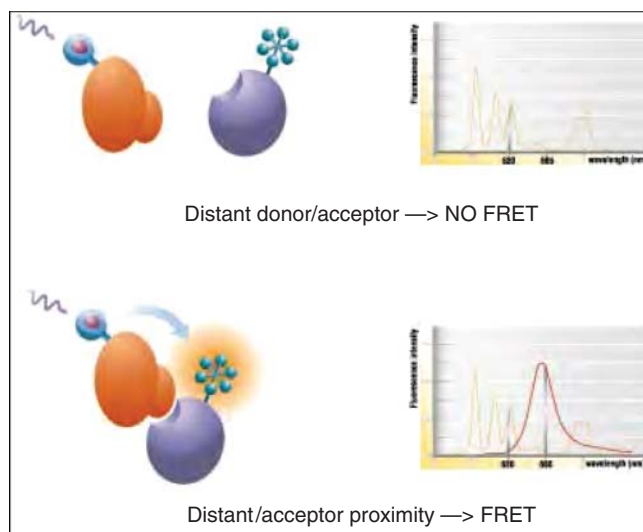
Figure 7 HTRF® technology principle. HTRF is a TR-FRET–based technology that uses the principles of both TRF (time resolved fluorescence) and FRET (fluorescence resonance energy transfer) technologies. TRF takes advantage of the unique photophysical and spectral properties of the rare earth elements called *lanthanides*. Specifically, they have large Stoke's shifts and extremely long emission half-lives (from μsec to msec) when compared to more traditional fluorophores. FRET uses two fluorophores, a donor and an acceptor. Excitation of the lanthanide donor by an energy source triggers a long life energy transfer to the acceptor if they are brought together by a biomolecular interaction. The acceptor in turn emits light at its given wavelength (red spectrum). A portion of the energy captured by the donor during excitation is released through fluorescence emission, in particular at 620 nm (yellow spectrum), whereas the remaining energy is transferred to acceptor. This energy is then released by the acceptor as specific fluorescence at 665 nm. The ratio between emitted fluorescence at 665 and 620 nm is proportional to the targeted biomolecular interaction. *Source*: Reproduced by courtesy of Cisbio Bioassays.

Table 3 Homogenous phase assays used in plant extract screenings.

Technology	Target	Format	Screen	Z′	Read-out	References
HTRF	Hematopoietic cell kinase	384	600 fractions from 150 plants in duplicate	0.85	Ratiometric time resolved fluorescence	Dufau *et al*. (2008)
Scintillation proximity assay	Liver X receptors	96	64,000 natural product extracts	ND	Yttrium silicate protein A –coated SPA beads	Jayasuriya *et al*. (2005)
Fluorescence polarization 1	Biotin–biotin binding proteins	384	Leaf extracts	ND	Fluorescence polariza- tion – Alexa-Fluor 594	Martin *et al*. (2008)
Fluorescence polarization 2	BclXL-Bak BH3 peptide binding	ND	107,423 extracts derived from natural products	ND	Fluorescence polariza- tion – fluorescein	Chan *et al*. (2003)
Enzyme comple- mentation	β2-adrenoceptor	96	352 HPLC-fractionated samples from ethanol extracts	0.68	Ratiometric fluorescence	Zhao, Yin, and Li (2010)

Abbreviations: ND, not determined.

3.6.3　In Vivo *Assays*

Automated *in vivo* screening has emerged in the past few years with the objective of improving early screening predictability using small model organisms. The worm *C. elegans* and the zebrafish

D. rerio were particularly studied (Katiki *et al*., 2011; Crawford *et al*., 2011). In the latter, plant extracts were screened directly on the fish model and *Oxygonum sinuatum* and *Plectranthus barbatus* methanol extracts were found to be sources of emodin and coleon A lactone active molecules.

3.7 Transfer on the HTS System and Screening Launch

The operation connecting a biological protocol noted in a laboratory book to the screening systems is the assay programming and scheduling: the optimized assay is coded in the system software, and then, it calculates the throughput by scheduling plate treatments (Figure 8). Each schedule is a working hypothesis requiring an extensive control by assay developers.
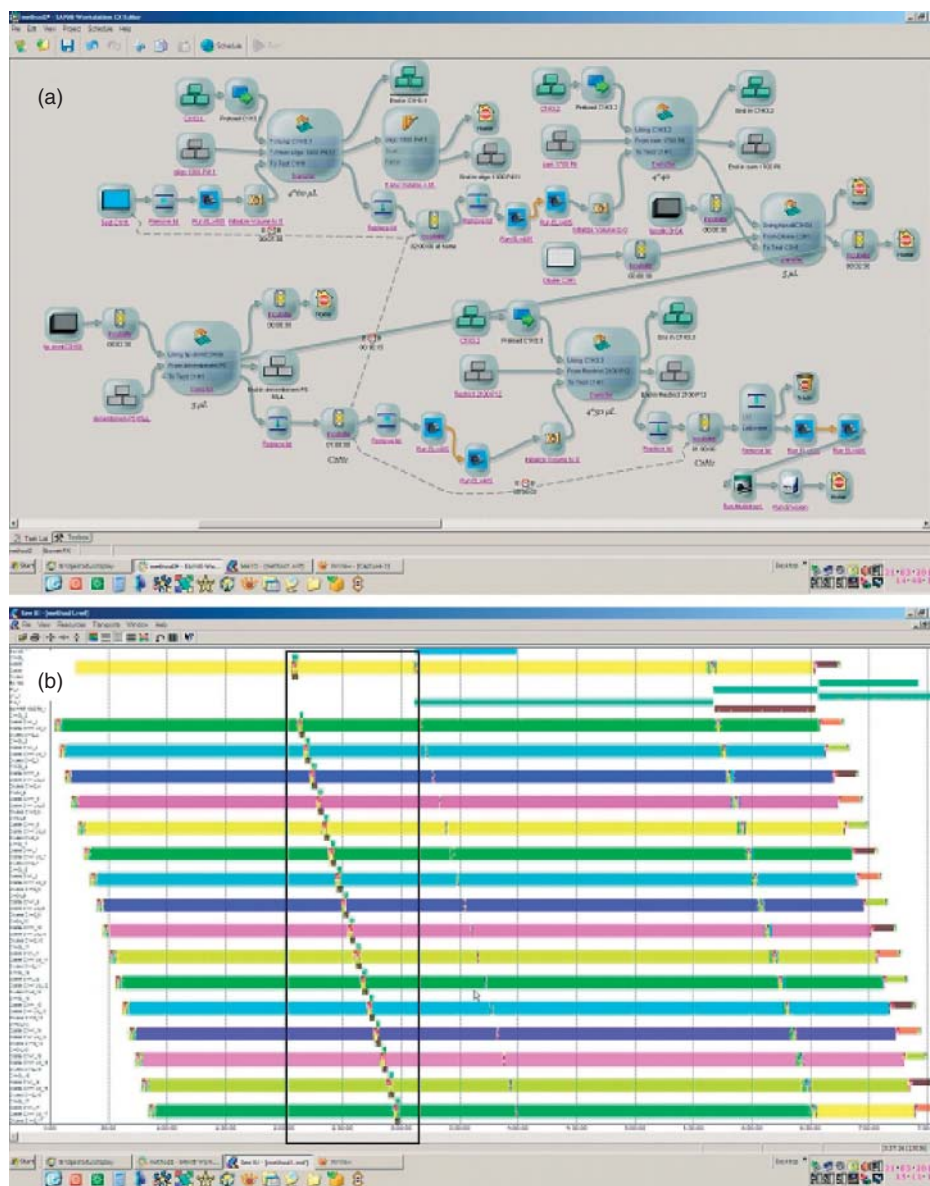
Figure 8 Assay automation and scheduling: (a) automated method developed on the SAMI EX scheduler (Beckman Coulter). Plates, tip boxes, modules, and protocol steps are graphically represented. Drag and drop options enable operators to translate laboratory-book protocols into machine language. (b) When teached to the system, the calculator propose a multiplate experiment. Each colored line represents a plate and time is on the X axis. Plates are operated simultaneously with a slight time shift. Any perturbation in the timing could have negative consequences on several plate processing. In the black frame, this time shift enables a fluid and continuous use of the pipeting resource. Finally, scheduling leads to the screening of seventeen 384-well plates (6528 tests) in less than 8 h whereas 6 h30 min are needed for each plate.

When the assay is fully automated and its performances controlled and validated, the automated screening launch can enter in its final countdown. However, a take-off checklist must be carefully followed. First, reagent amounts should be controlled. Tubing, plates, and reservoirs dead volumes should be integrated in total quantity calculations and an excess of 15–25% is a reasonable precaution. The same operation should be repeated for tips and other consumables. Regulatory use of biohazard and radioactive reagents should also be verified.

The access to chemical libraries enters now in a critical phase. Some systems have the benefit of being physically connected to the screening station, enabling the operator to control the thawing of frozen chemical sources. For the others, this transfer is offline and generally operated manually. Then, a reproducible protocol should be designed to avoid heterogeneity in compoubd libraries utilization.

3.8 Quality Control and Hit Selection

3.8.1 Controlling Screening Runs

When started, automated screening becomes autonomous, but operators have a supervision role in case of trouble. However, human interventions in the ongoing screenings are sources of problems, and, in fact, a few things can be done avoiding any interferences with scheduling calculations (Figure 8b). Several indicators can be followed all along the process such as plate positioning in locators, absence of bubbles in tips and tubing, and respect of predefined operation times.

In HTS, control molecules are inserted into each run. These controls can be located in each screening plate or in dedicated plates plates or both. In the first and more used case, positive and negative controls are positioned in the first and last columns of the plates, respectively. This enables to obtain more reliable data, because each plate in the run contains two control steps; however, this choice has two important consequences. First, rows 1 and 12 in the 96-well format and rows 1, and sometimes 2, in addition to 23 and 24 in the 384-well format are interference prone. Light contamination and temperature and evaporation gradients can be responsible for side effects. Then, these rows represent the

worst positions for controlling. Their use must be carefully validated during the assay automation step. Authors proposed an approach to reduce the systematic row/column effects in HTS (Coma, Herranz, and Martin, 2009). In practice, no single method is the best for every HTS data set, and authors proposed a multistep strategy for HTS data-processing and hit identification (Shun *et al.*, 2011). Second, wells dedicated to controls in the assay plate must be left empty in the plates containing the chemical sources. This means that around 17% of the wells in both the formats are lost for chemical sources.

A benefit of extracts in the screening control appears when several fractions, purified from a mother extract or less purified fraction, are located in the same plate and detected as hits. If the active molecules are not fully separated, there is a link between these fractions, justifying their biological activity.

Automated screening experiments can last from several hours to several days when cells or living organisms are included in the protocol. Maintaining a homogenous biological response all along the experiment is challenging and carefully studied in the assay automation procedure. However, it is difficult to predict and solve all putative defects in small size experiments. Then, the first full scale screening is generally used as a troubleshooting step. Signal drifts are one of these unpredictable difficulties and have dramatic consequences on hit selection. These drifts are time and location dependent. They can be observed within plates (Figure 9a) or all along the run. Screening should be stopped and the cause clearly identified. One of the most important points to keep under control during screening runs is the hit rate. If a good hit rate (0.05–1%) does not warrant a successful screening campaign on extracts (Section 3.4), too many or not enough hits will certainly lead to an overall failure. As extracts contain numerous molecules, it could be anticipated that more active responses will be found compared to pure compound screenings and that the overall run response can be modified (Figure 9b) (Ausseil *et al.*, 2007).

Higher hit rates can also signify that there are too many false positives owing to the lack of specificity of the assay or the target. Another hypothesis is that screening concentrations are too high to ensure relevant responses. *A contrario*, no hit identified among several thousands of substances tested is
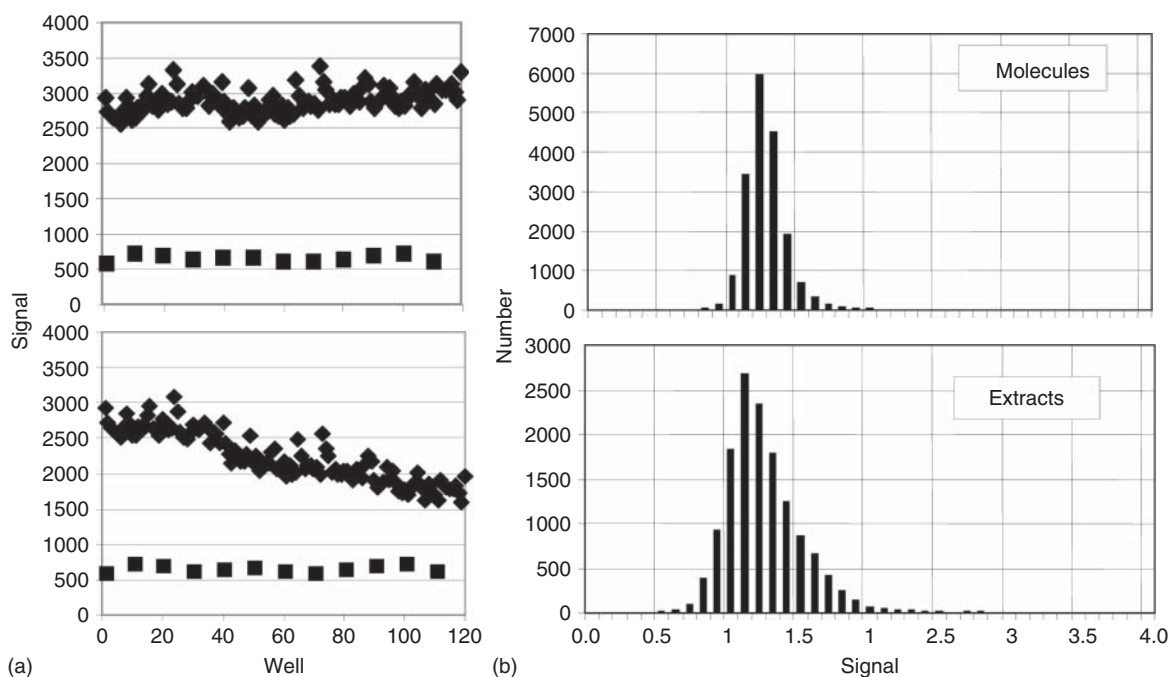
Figure 9 Screening result profiles: (a) within-plate signal drift. Repeats of positive (black diamond) and negative (black square) control wells in a 384-well plate. Top: the two levels of controls are stable, mean, standard deviation, and median calculations of each control are relevant for Z' calculation and hit finding. Down: positive control signals strongly decrease according to their position in the plate. None reliable calculations could be obtained, invalidating the whole process. (b) Number of wells for various signal levels. Top: 18,239 molecules. Bottom: 15,744 plant extracts. Hits are located at high signal levels. Extract distribution is shifted toward higher signal values than molecules. Specific hit selection calculation should be applied to extracts.

disturbing. In these cases, adjustments should be done in the following runs.

3.8.2   Hit Selection

Because it is unrealistic to isolate one by one all the compounds from the plants to evaluate them as pure isolated products, the classical procedure is to screen purified extracts. With pure compounds, a screening concentration can be selected (corresponding, for example, to the IC_{50} or EC_{50} of the reference compounds), with semipurified fractions, molecules concentrations cannot be rationally determined, and the same overall activity can be observed in very different cases (Figure 10).

Moreover, biological responses are often continuously distributed over the signal range. Indeed, hit selection is based on the comparison with a cutoff value separating statistically significant responses

from the bulk. These cutoff values can be more or less stringent depending on the level of significance required (Figure 11). Thus, very active molecules but highly diluted into fractions can be unselected because they do not reach the targeted cutoff. The initial concentration of an active compound might be too minute to be revealed by a HTS screening. Low solubility of the active compound may also prevent an effective detection. Thus, the most attractive fractions in terms of percentage of inhibition are not necessarily corresponding to the most potent molecule. Concentrated poorly active compounds will provide the same overall pharmacological effect than small concentrations of very active ones.

Low hanging fruits have already been collected; therefore, the top 5% result of one HTS campaign has to be considered but will not bring for sure interesting hits. Therefore, we have to carefully study as many active fractions as possible in order to find out nonabundant active molecules, which have until now escaped the researchers.
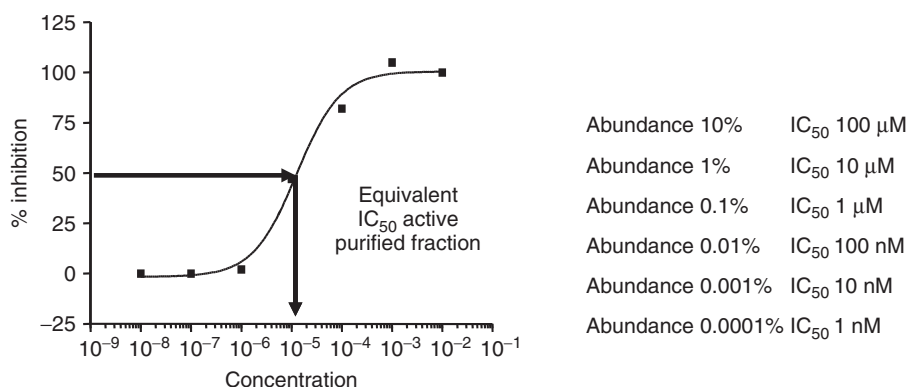
Figure 10 Purified fractions do not allow a real classification of fraction potencies as the overall effect can result from a variety of combination of potency and concentration of a single active component. Activity observed (apparent IC_{50} μM) = abundance (%) × potency (IC_{50} μM).
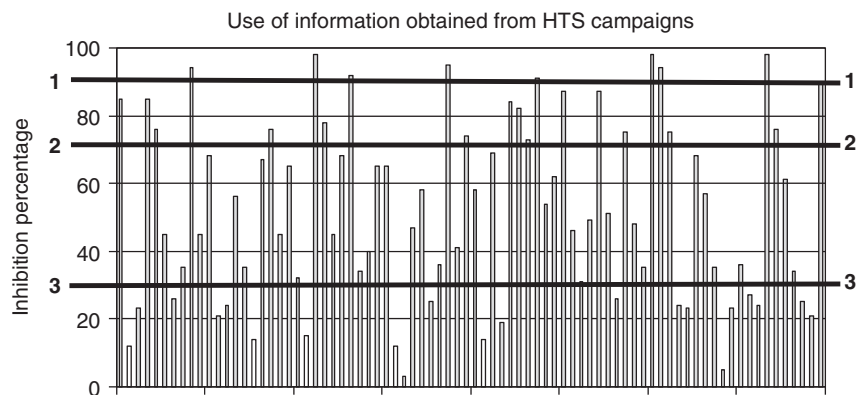
Figure 11 Use of information obtained from HTS campaigns. This picture presents the overall bioactivity results of one plate. The attractive hits are not necessary the most interesting. It is necessary to consider a big number of results (cutoff #2 or 3) to get new bioactive molecule.

3.8.3   *Bioguided Hit Fractionation*

Hit fractionation includes the experiments leading to purified molecules from active extracts. This is generally not included in HTS. The simultaneous handling of fractions produced from tenths to hundreds extracts can results in several thousand assays when various concentrations are tested. Then, automation regains interest. As mentioned in Section 3.4, the first step consists in validating the activity of the plant source available in quantities compatible with molecule purification. In order to reduce process variability, the same assay than the one used in HTS is used to confirm sample activity and all along the

fractionation process. Active molecule concentration might vary during the bioguided fractionation and affect the pharmacological results even if each fraction and subfraction is evaluated at its relative proportion of the initial extract (equivalent method). Thus, it is preferable to test various concentrations of fractions at each step of the process.

4   CONCLUSION

Until the late 1970s, the discovery of drugs from plant origin (extracts and pure compounds) was

mainly based on traditional knowledge and serendipity. This approach relying on tests with small animals and isolated organs allowed access to most of natural active principles still on the markets nowadays. From the nineteenth century to the 1980s, the discovery started rather by clinical or pharmacological observations and then moved to the mechanism and target explanation to the molecular levels. In the 1980s, with the progress in biochemistry and pharmacology, many enzymes and receptors correlated with clinical pathologies were available for extensive screening. Development of biotechnology, molecular biology techniques, and progress in automation and miniaturization made the HTS possible (Figure 12).

Natural products with their huge intrinsic chemodiversity were welcomed as the organic chemists were struggling to produce the thousands of molecules to meet the needs of the robots of this technological revolution. With this new approach, the strategy was reverse compared to preceding decades and started from the molecular target to move to the cells, tissues, isolated organs, the small animals, large animals, and, finally, the patients (Figure 13).

HTS allows a huge number of experiences, thanks to the miniaturization of simple models; but, these experiments are very far away from patient, and therefore success was not as significant as expected in the Big Pharma. Thus, one has to translate enzyme activity toward models of increasing complexity to finally evaluate clinical activity on man. Starting *in vitro* HTS screening is one way, whereas choice of downstream assays that will allow access to proof of concept and to start clinical trials is another. This needs a careful strategic reflection and programming before starting the project.

The efficiency and potentialities of HTS are huge (Figure 12); a classical HTS campaign can explore in less than 2 months the pharmacological potency of an entire world plant diversity collection on one particular target. Therefore, if the expected results are less numerous, it is not that technology and strategy are not correct, but that the low hanging fruits have already been picked. According to the authors' experience, podophyllotoxins, taxoids, vinca alkaloids, and so on are quickly and readily identified but, unfortunately, are already known and on the market for over 50 years. The HTS on natural products, and in particular on plants, can be considered a mature strategy, which already provide its easily accessible results in comparison with other emerging strategies that look more attractive, rational, and, therefore, monopolize research investments (Figure 14). The time is over where plant molecules provided easy hits, by-passed the hit-to lead phase, and the lead optimization to furnish ready to use active principles.
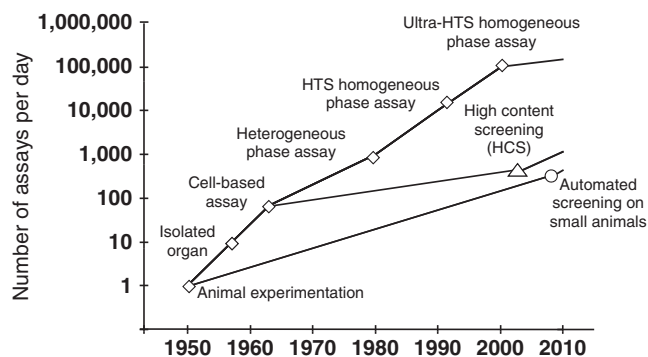
Figure 12 Evolution of screening. From the middle of the twentieth century until now, the number of assays performed by a single technician was dramatically improved by technology and arrival of HTS. Today, the trend is more a quest of quality and diversity of information rather than quantity with high content screening and automated screening on small animals.

Direction of research 1800–1980
Patient → large animal cell → small animal → isolated organ → tissue → molecular target
Direction of research 1980–onwards
Molecular target → cell → tissue→ isolated organ → small animal →large animal → patient

Figure 13 Directions of research before and after development of molecular targets.
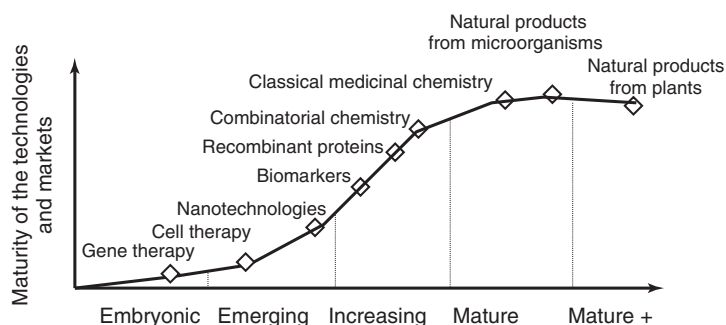
Figure 14   Level of maturity of the strategies and techniques in the pharmaceutical industry, adapted from LEEM (www.leem.org).

Strong reduction or arrest	Still active
Abbott	Dabur
Astellas	Eisai
Bayer	Novartis
Boehringer Ingelheim	Pierre Fabre
Bristol-Myers Squibb	Piramal
Daiichi Sankyo	
Eli Lilly	
GlaxoSmithKline	
Johnson & Johnson	
Kyowa Hakko	
Merck Sharp & Dohme	
Novo Nordisk	
Pfizer	
Roche	
Sanofi	
Servier	
Takeda	

Figure 15   Big Pharma active in plant natural products discovery programs (libraries and HTS). Author's knowledge.

Nowadays, the phytochemical research is stronger in analytics (see **HPLC and Ultra HPLC: Basic Concepts**; **Analysis of Natural Products by Capillary Electrophoresis and Related Techniques**, and **NMR as Analytical Tool for Crude Plant Extracts**) and structural elucidation but needs to work much harder.

The plants are still insufficiently studied as many specific issues regarding unstable legal access, logistic, isolation and so on slowed the academic and pharmaceutical research and managed to divert the majority of Big Pharma (Figure 15).

In the last 10 years, we have observed a tendency to spin-off the HTS plant natural products discovery programs. Microorganisms seem easier to work because of easier legal access (local strains are sufficient); new metabolites pathways can be expressed under various conditions and lead to a huge chemical diversity (Kingston, 2011). Nowadays, research on higher plants is mostly performed in academia laboratory, spin-off and start-up Companies such as: Albany Molecular Research, AnalytiCon, Basilea Pharm, ICSN-CNRS, InterMed Discovery, Merlion, Greenpharma, Phynova, Sequoia Sciences, UniBioScreen, and so on.

The tendency to overlook plant natural products may be turning for several reasons:

- We have to keep in mind the continuous decline in new drug approvals. Neither combinatorial chemistry nor drug design strategies were able to fill up the R&D pipe of pharmaceutical industries (Pammolli, Maggazzini, and Riccaboni, 2012; Scannell *et al.*, 2012);
- Interest for natural products is growing in biodiversity rich countries;

- Since the last decade, academic research has entered this area through HTS;
- Most of the difficulties associated with work on natural products that can technically be overcome (See *In Silico*-**Guided Strategies for the Discovery and Rationalization of Bioactive Natural Products** and **Innovative Strategies in the Search for Bioactive Plant Constituents**). Hyphenated techniques including automated fractionation and dereplication are developing (Marston and Hostettmann, 2009; Johnson *et al.*, 2011) (See **LC and LC-MS: Techniques and Applications** and **On-line and At-line LC-NMR and Related Micro NMR Methods**). This research can be streamlined in order to pinpoint interesting and new compounds in rational and effective processes. Continuing advances in the chemistry separation techniques, HCS, smart screening methods, progress in robotization of bioguided fractionation and dereplication, and high throughput structure elucidation are able to bring a new start in natural product drug discovery programs (Johnson *et al.*, 2011).
- Progress in dereplication using mass spectrometry and nuclear magnetic resonance will make easier constitution of vast libraries of vegetal pure compounds, which can overcome the mixture issues (Avery *et al.*, 2010; Koehn, 2008) (See **LC and LC-MS: Techniques and Applications** and **NMR of Small Molecules**);
- Discovery of new structures and scaffolds will continue and will be associated of "new" or poorly investigated plant samples;
- Most of the plant species are growing in tropical countries but tropics led only to 30% of the plant-derived medicines when Eurasia provided 70% of these medicines. The potential of tropics is therefore huge. We have to admit that restricted distribution of a wide part of tropical plants explains why they are underrepresented as new chemical entities and poorly evaluated even by local healers (Miller, 2011);
- Estimations of the number of new drugs to be discovered from plants range between 500 and 20,000 chemical entities (Miller, 2011). This calculation considers the percentage of plants not yet studied and the number of active molecules provided by the percentage of studied plants;
- The paradigm of the single bullet that led companies to shy away from natural plant substances is

likely to evolve as different national health agencies will accept more and more plant extracts in the future. On October 2006, the US Food and Drug Administration approved a mixture of sinecatechins (Veregen®) extracted from Chinese green tea, *Camellia sinensis*, as a topical treatment of external genital and perianal condyloma (Chen *et al.*, 2008). This FDA approval shows that complex mixtures from plant can be developed as new therapies to meet current health agencies' standards of quality and efficiency. Traditional Chinese Medicine is more and more considered with attention by scientist as proved by an increasing number of scientific publications (Zhu *et al.*, 2010). (see **Quality Assessment of Herbal Drugs and Medicinal Plant Products**). On March 2012, for the very first time in the European Union a marketing authorization was delivered by the Dutch Medecines Evaluation Board to a Chinese traditional herbal medicinal product. The dry extract of *Dioscorea nipponica* rhizomes has been registered in the Netherlands for "the relief of headache and pain and cramps in the muscles of neck, back and legs" (Gilbert, 2012). The guidance of traditional medicine for drug discovery seems appears more and more attractive (Saslis-Lagoudakis *et al.*, 2012);
- It is possible to expand the range of 'druggable' targets with natural products (Bauer, Wurst, and Tan, 2010), but we have to acknowledge that some druggable targets cannot be addressed by natural products;
- Natural products are subject of renewed interest particularly by their conjugation with monoclonal antibodies (Monneret, 2010);
- Most of the Big Pharma, which have stopped higher plant prospecting programs are still using isolated natural vegetal products or even purified fractions in their screening campaigns. When screening campaigns are unfruitful, metabolites from plants are welcomed as the last opportunity;
- The scientific reasoning and exchanges between (phyto)chemists and biologists are back after the period of "Magic of large numbers" (Drews, 2000).
- Natural products are more efficient than the medicinal chemistry molecules to address challenging targets like the protein–protein interactions (Bauer, Wurst, and Tan, 2010). This is a real advantage as the existing drugs act on a very restricted variety of biological targets. Natural products are

also interesting because they turned out to be real chemical probes for biological targets (Carlson, 2010);

- Plant natural molecules, derivatives and inspired compounds are considered important sources for new drug candidates, essential tools for chemical biology and medicinal chemistry (Wang, 2008; Kumar and Waldmann, 2009);
- Beside classical high-throughput screening, which is only oriented toward one target at a time, HTS can discover small molecules that can specifically modulate every function of every gene products in a cell (chemical genomics). In this regard, natural products have proved of immense value (see *In Silico*-**Guided Strategies for the Discovery and Rationalization of Bioactive Natural Products**). This research has revealed numerous "new" targets with a very high binding selectivity with natural products on new biological targets of fundamental or potential interest uncovered by the genome sequencing. In humans, 3 to 10,000 potentially druggable targets are estimated (Penrod, Cowper-Sal-Iari, and Moore, 2011; Sakharkar, Sakharkar, and Pervaiz, 2007), but fewer than 400 distinct molecular targets of drugs are approved by the US Food and Drug Administration (FDA). Maybe one part of the vast druggable genome will not be accessible to the plants molecules but conversely some targets will only be addressed by such molecules;
- Some top managers in Big Pharma, as S. Schwan (CEO of Roche), think that the skeptics are wrong and that "the low-hanging fruits have been harvested. They are enormous opportunities in the pharma industry and we are only now beginning to understand how diseases are working" (Mijuk, 2010). The progress in genetics and molecular biology will provide new perspective and natural products will play an active part;
- Phenotypic screening considered more fruitful than target-centric approaches (Swinney and Anthony, 2011) is an opportunity to the screening of plant chemical diversity and to the renewal of natural products.
- Improving the quality and confidence of target are the more important factors to improve drug discovery (Bunnage, 2011); hope is permitted as the opportunities for addressing new targets are huge.

This research is intrinsically difficult but a broadband screening and technological strategies can be implemented to maximize the chances of success that can be formulated as (Pieters and Vlietinck, 2005):

$$\text{Successes} = \text{samples} \times \text{biodiversity of samples} \\ \times \text{targets} \times \text{pharmacological domains.}$$

This means that the chances of success are proportional to the number of samples, the chemodiversity of these samples, the domains evaluated and the number of targets screened in one domain.

Maybe the expectations have been too high; researchers looking for undescribed natural products associated with new biological activity which represent of course the most attractive but also most difficult research results (Figure 16).

The chemical and pharmacological inventory of vegetal compounds is far to be achieved even if for some researchers, the redundancy of plant natural products makes unlikely the discovery of new chemical scaffolds (Kong *et al.*, 2011). The plant "pharmacological activity friendliness" associated to specific and potent activities are serious advantages.

The screening of natural products will grow again in the pharmaceutical industry if access is facilitated by countries rich in biodiversity and if valuers are willing to share. Without easy access no investment in research and thus no sharing possible. Academic research should not be stifled by to restrictive access regulation (Martinez and Biber-Klemm, 2010). We need to avoid falling into regulation quagmires. This
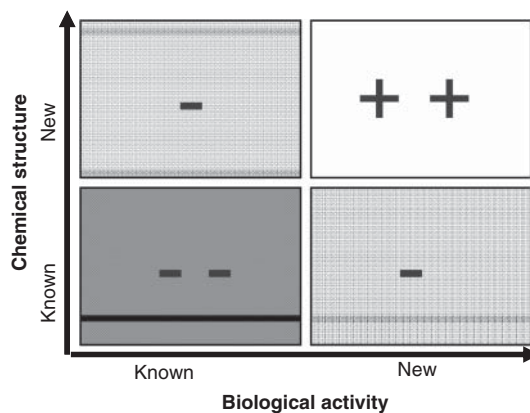
Figure 16 Interest = combination of chemical and biological novelties (Bruhn and Bohlin, 1997). *Source*: J.G. Bruhn and L. Bohlin (1997). Reproduced from Elsevier.

is an opportunity for everybody: supplier countries, academic and private researchers. We are at a key moment where everything can switch in one direction or another. Hopefully bioprospection can give value to biodiversity and contributes to its preservation with the sharing of benefits in a sustainable and responsible way. If accesses are too difficult, paradoxically logging will take over and the ecosystems will be irreversibly devastated. Since decades, the potential loss of active chemodiversity with tropical deforestation and disturbance of biotopes is denounced by scientists (Abelson, 1990) while less than 15% of the vascular plant have been pharmacologically studied (Miller, 2011; Wang, 2008).

Access should be facilitated and technical hurdles must be overcome in order to share benefit, in a "win-win-win" strategy. To find out the new effective treatments which are urgently needed by patients by example in the field of oncology, to fill the development pipe of pharmaceutical industries and to develop nature conservation programs in biodiversity-rich countries.

ACKNOWLEDGMENTS

REFERENCES

Abelson (1990) *Science*, **247**, 513.

Appendino, G. and Pollastro, F. (2010) Plants: revamping the oldest source of medicines with modern science, in Natural Products Chemistry for Drug Discovery. RSC Biomolecular Sciences N°18, eds. A. D. Buss and M. S. Butler, Royal Society of Chemistry, Cambridge, pp. 140–173.

Ausseil, F., Samson, A., Aussagues, Y., *et al.* (2007) *J. Biomol. Screen.*, **12**, 106–116.

Avery, V. M., Camp, D., Carroll, A. R., *et al.* (2010) The identification of bioactive natural products by HTS, in *Comprehensive Natural Products II*, ed. R. Verpoorte, Elsevier, Amsterdam, vol. **3**, pp. 177–203.

Baell, J. B. and Holloway, G. A. (2010) *J. Med. Chem.*, **53**, 2719–2740.

Bailly, C. (2009) *Biochem. Pharmacol.*, **77**, 1447–1457.

Bauer, R. A., Wurst, J. M. and Tan, D. S. (2010) *Curr. Opin. Chem. Biol.*, **14**, 308–314.

Bindseil, K. U., Jakupovic, J., Wolf, D., *et al.* (2001) *Drug Discov. Today*, **6**, 840–847.

Boudsocq, F., Benaim, P., Canitrot, Y., *et al.* (2005) *Mol. Pharmacol.*, **67**, 1485–1492.

Bruhn, J. G. and Bohlin, L. (1997) *Drug Discov. Today*, 243–246.

Bunnage, M. E. (2011) *Nat. Chem. Biol.*, **7**, 335–339.

Carlson, E. E. (2010) *Acs Chem. Biol.*, **5**, 639–653.

Chan, S. L., Lee, M. C., Tan, K. O., *et al.* (2003) *J. Biol. Chem.*, **278**, 20453–20456.

Chen, S. T., Dou, J. H., Temple, R., *et al.* (2008) *Nat. Biotechnol.*, **26**, 1077–1083.

Cheng, X. H., Hochlowski, J., Tang, H., *et al.* (2003) *J. Biomol. Screen.*, **8**, 292–304.

Coma, I., Herranz, J. and Martin, J. (2009) *Methods Mol. Biol.*, **565**, 69–106.

Cordell, G. A. (2010) The convention on biological diversity and its impact on natural product research, in Natural Products Chemistry for Drug Discovery. RSC Biomolecular Sciences N°18, eds. A. D. Buss and M. S. Butler, Royal Society of Chemistry, Cambridge, pp. 81–139.

Crawford, A. D., Liekens, S., Kamuhabwa, A. R., *et al.* (2011) *PLoS One*, **6**, e14694.

Dančík, V., Seiler, K. P., Young, D. W., *et al.* (2010) *J. Am. Chem. Soc.*, **132**, 9259–9261.

Degorce, F., Card, A., Soh, S., *et al.* (2009) *Curr. Chem. Genom.*, **3**, 22–32.

Delvecchio, C., Tiefenbach, J. and Krause, H. M. (2011) *Assay Drug Dev. Technol.*, **9**, 354–361.

Demaine, L. J. and Fellmeth, A. X. (2003) *Science*, **300**, 1375–1376.

Devlin, J. P. (1997) Microcollection of plants for biochemical profiling, in *High Throughput Screening. The Discovery of Bioactive Substances*, ed. J. P. Devlin, Marcel Dekker, New York, pp. 49–76.

Dixon, N., Wong, L. S., Geerlings, T. H., *et al.* (2007) *Nat. Prod. Rep.*, **24**, 1197–1432.

Drews, J. (2000) *Science*, **287**, 1960–1964.

Dufau, I., Lazzari, A., Samson, A., *et al.* (2008) *Assay Drug Dev. Technol.*, **6**, 673–682.

Edwards, P. J. (2006) *IDrugs*, **9**, 347–353.

Eglen, R. M., Reisine, T., Roby, P., *et al.* (2008) *Curr. Chem. Genom.*, **1**, 2–10.

Eitsuka, T., Nakagawa, K., Igarashi, M., *et al.* (2004) *Cancer Lett.*, **212**, 15–20.

Engeloch, C., Schopfer, U., Muckenschnabel, I., *et al.* (2008) *J. Biomol. Screen.*, **13**, 999–2008.

Entzeroth, M., Flotow, H., Condron, P., *et al.* (2009) *Current Protocols in Pharmacology. Overview of High-Throughput Screening*, John Wiley & Sons, Ltd, Chichester, Chapter 9: Unit 9.4.

Erkens, R. H. J. (2011) *Nat. Prod. Rep.*, **28**, 11–14.

Feling, R. H., Buchanan, G. O., Mincer, T. J., *et al.* (2003) *Angew. Chem. Int. Ed. Engl.*, **42**, 355–357.

Fonrose, X., Ausseil, F., Soleilhac, E., *et al.* (2007) *Cancer Res.*, **67**, 3371–3378.

Frearson, J. A. and Collie, I. T. (2009) *Drug Discov. Today*, **14**, 1150–1158.

Ganesan, A. (2008) *Curr. Opin. Chem. Biol.*, **12**, 306–317.

Gilbert, N. (2012) Chinese herbal medicine breaks into EU market. *Nature News Blog*, 19 April.

Glickman, J. F., Schmid, A. and Ferrand, S. (2008) *Assay Drug Dev. Technol.*, **6**, 433–455.

Han, Q. B., Yu, T., Lai, F., *et al*. (2010) *Talanta*, **82**, 1521–1527.

Jayasuriya, H., Herath, K. B., Ondeyka, J. G., *et al*. (2005) *J. Nat. Prod.*, **68**, 1247–1252.

Johnson, T. A., Sohn, J., Inman, W. D., *et al*. (2011) *J. Nat. Prod.*, **74**, 2545–2555.

Katiki, L. M., Ferreira, J. F., Zajac, A. M., *et al*. (2011) *Vet. Parasitol.*, **182**, 264–268.

Kawatani, M., Takayama, H., Muroi, M., *et al*. (2011) *Chem. Biol.*, **18**, 743–751.

Kepp, O., Galluzzi, L., Lipinski, M., *et al*. (2011) *Nat. Rev. Drug Discov.*, **10**, 221–237.

Kingston, D. G. I. (2011) *J. Nat. Prod.*, **74**, 496–511.

Koehn, F. E. (2008) High impact technologies for natural products screening, in *Progress in Drug Research*, eds. F. Petersen and R. Amstutz, Birkhäuser, Basel, vol. **65**, pp. 176–210.

Kong, D. X., Guo, M. Y., Xiao, Z. H., *et al*. (2011) *Chem. Biodiv.*, **8**, 1968–1977.

Kumar, K. and Waldmann, H. (2009) *Angew. Chem. Int. Ed.*, **48**, 3224–3242.

Lea, W. A. and Simeonov, A. (2011) *Expert Opin. Drug Discov.*, **6**, 17–32.

Lee, J. Y., Miraglia, S., Yan, X., *et al*. (2003) *J. Biomol. Screen.*, **8**, 81–88.

Lee, H. K., Brown, S. J., Rosen, H., *et al*. (2007) *Mol. Pharmacol.*, **72**, 868–875.

Leister, K. P., Huang, R., Goodwin, B. L., *et al*. (2011) *Curr. Chem. Genom.*, **5**, 21–29.

Leung, C. K., Deonarine, A., Strange, K., *et al*. (2011) *J. Vis. Exp.*, **51**pii: 2745.

Long, C., Marcourt, L., Raux, R., *et al*. (2009) *J. Nat. Prod.*, **72**, 1804–1815.

Macarrón, R. (2006) *Drug Discov. Today*, **11**, 277–279.

Macarrón, R. and Hertzberg, R. P. (2011) *Mol. Biotechnol.*, **47**, 270–285.

Macarrón, R., Banks, M. N., Bojanic, D., *et al*. (2011) *Nat. Rev. Drug Discov.*, **10**, 188–195.

Marston, A. and Hostettmann, K. (2009) *Planta Med.*, **75**, 672–682.

Martin, H., Murray, C., Christeller, J., *et al*. (2008) *Anal. Biochem.*, **381**, 107–112.

Martinez, S. I. and Biber-Klemm, S. (2010) *Curr. Opin. Environ. Sustain.*, **2**, 1–7.

Massiot, G., Long, C., David, B., *et al*. (2005) *J. Nat. Prod.*, **68**, 979–984.

Mayr, L. M. and Bojanic, D. (2009) *Curr. Opin. Pharmacol.*, **9**, 580–588.

Meireles, L. M. and Mustata, G. (2011) *Curr. Top. Med. Chem.*, **11**, 248–257.

Michael, S., Auld, D., Klumpp, C., *et al*. (2008) *Assay Drug Dev. Technol.*, **6**, 637–657.

Mijuk, G. (2010) A healthy forecast for Pharma. *The Wall Street Journal*. July 22.

Miller, J. S. (2011) *Econ. Bot.*, **65**, 396–407.

Monneret, C. (2010) *Ann. Pharm. Fr.*, **68**, 218–232.

Newman, D. J. and Cragg, G. M. (2012) *J. Nat. Prod.*, **75**, 311–335.

Newman, D. J., Cragg, G. M., Holbeck, S., *et al*. (2002) *Curr. Cancer Drug Targets*, **2**, 279–308.

Oliva, M. J. (2011) *Planta Med.*, **77**, 1221–1227.

Pammolli, F., Maggazzini, L. and Riccaboni, M. (2012) *Nat. Rev. Drug Discov.*, **10**, 428–438.

Paterson, I. and Anderson, E. (2005) *Science*, **310**, 451–453.

Penrod, N. M., Cowper-Sal-Iari, R. and Moore, J. H. (2011) *Trends Pharmacol. Sci.*, **32**, 623–630.

Pieters, L. and Vlietinck, A. J. (2005) *J. Ethnopharm.*, **100**, 57–60.

Pouny, I., Etiévant, C., Marcourt, L., *et al*. (2011) *Planta Med.*, **77**, 461–466.

Rouhi, A. M. (2003) *Chem. Eng. News*, **81**, 77–91.

Sakharkar, M. K., Sakharkar, K. R. and Pervaiz, S. (2007) *Intern. J. Biochem. Cell Biol.*, **39**, 1156–1164.

Santagata, S., Xu, Y. M., Wijeratne, E. M., *et al*. (2012) *ACS Chem. Biol.*, **7**, 340–349.

Saslis-Lagoudakis, C. H., Savolainenb, V., Williamson, E. M., *et al*. (2012) *Proc. Natl. Acad. Sci. U. S. A.*, **109**, 15835–15840.

Scannell, J. W., Blanckley, A., Boldon, H., *et al*. (2012) *Nat. Rev. Drug Discov.*, **11**, 191–200.

Schindel, D. E. (2010) *Nature*, **467**, 779–781.

Scotland, R. W. and Wortley, A. H. (2003) *Taxon*, **52**, 101–104.

Shun, T. Y., Lazo, J. S., Sharlow, E. R., *et al*. (2011) *J. Biomol. Screen.*, **16**, 1–14.

Sturgeon, C. M., Craig, K., Brown, C., *et al*. (2005) *Planta Med.*, **71**, 938–943.

Swinney, D. C. and Anthony, J. (2011) *Nat. Rev. Drug Discov.*, **10**, 507–519.

Thorne, R. F. (2002) *Taxon*, **51**, 511–512.

Thorne, N., Auld, D. S. and Inglese, J. (2010) *Curr. Opin. Chem. Biol.*, **14**, 315–324.

Tian, E., Landowski, T. H., Stephens, O. W., *et al*. (2008) *Mol. Cancer Ther.*, **7**, 500–509.

Tiedemann, R. E., Schmidt, J., Keats, J. J., *et al*. (2009) *Blood*, **113**, 4027–4037.

Tong, X. G., Wu, G. S., Huang, C. G., *et al*. (2010) *J. Nat. Prod.*, **73**, 1160–1163.

Tu, Y., Jeffries, C., Ruan, H., *et al*. (2010) *J. Nat. Prod.*, **73**, 751–754.

Vandenberghe, I., Créancier, L., Vispé, S., *et al*. (2008) *Biochem. Pharmacol.*, **76**, 453–462.

Verpoorte, R. (2010) Overview and introduction, in *Comprehensive Natural Products II*, ed. R. Verpoorte, Elsevier, Amsterdam, vol. **3** Development & Modification of Bioactivity, pp. 1–4.

Voller, J., Zatloukal, M., Lenobel, R., *et al*. (2010) *Phytochemistry*, **11–12**, 1350–1359.

Wang, Y. (2008) *Phytochem. Rev.*, **7**, 395–406.

Wang, H., Li, S. Y., Zhao, C. K., *et al*. (2009) *J. Zhejiang Univ. Sci. B*, **10**, 243–250.

Wunder, F., Kalthof, B., Müller, T., *et al*. (2008) *Comb. Chem. High Throughput Screen.*, **11**, 495–504.

Zhang, J. H., Chung, T. D. and Oldenburg, K. R. (1999) *J. Biomol. Screen.*, **4**, 67–73.

Zhang, H. W., Song, Y. C. and Tan, R. X. (2006) *Nat. Prod. Rep.*, **23**, 753–771.

Zhao, C. K., Yin, Q. and Li, S. Y. (2010) *Acta Pharmacol. Sin.*, **31**, 1618–1624.

Zhu, M., Phillipson, J. D., Greengrass, P. M., *et al*. (1997) *Phytochemistry*, **44**, 441–447.

Zhu, Y. H., Zhang, Z. Y., Zhang, M., *et al*. (2010) *Comb. Chem. High Throughput Screen.*, **13**, 837–848.

Zhuang, Y., Yan, J. J., Zhu, W., *et al*. (2008) *J. Ethnopharm.*, **117**, 378–384.

Mycotoxins Contamination in Food: Alternative Plant Preservatives, Legislation, and Detection Methods

Monica R. Zuzarte[1], Ana P. Martins[2], Maria J. Gonçalves[1,2] and Lígia R. Salgueiro[1,2]

[1]*Center for Neuroscience and Cell Biology, Department of Zoology, University of Coimbra, Coimbra, Portugal* and [2]*Center of Pharmaceutical Studies, Faculty of Pharmacy, University of Coimbra, Coimbra, Portugal*

1 INTRODUCTION

Fungi that contaminate crops, foods, commodities, and raw materials are the source of highly toxic compounds, generally known as *mycotoxins*. These contaminations are a severe problem that depends on production and storage practices and varies with the type of food. Mycotoxin production is influenced by abiotic factors (temperature, water availability, and gas composition) and biotic factors (nature of the substrate, moisture content, and insect infestation) (Dowd, 2003; Magan *et al*., 2003). Moreover, mycotoxin production may occur before any detectable evidence of fungal growth (Pirbalouti *et al*., 2011), causing severe health concerns.

Fungal toxins of greatest concern are produced by species of the genera *Aspergillus*, *Fusarium*, and *Penicillium* and frequently occur in major field crops, which they continue to contaminate during storage (Reddy and Bhoola, 2010). The Food and Agriculture Organization of the United Nations (FAO) has estimated that up to 25% of the world's food

crop, as well as an even higher percentage of feedstuff, is significantly contaminated by mycotoxins that pose a risk to human safety (Magda and Castillo, 2002). When ingested, mycotoxins are known to cause general toxic effects, immune system suppression, mutations, cancer, teratogenic effects, and infertility (El-Azab *et al*., 2010). More than 400 different mycotoxins have been identified to date; however, as many fungi are not well studied and as not all species have been identified, some researchers have suggested that there could potentially be as many as 300 000 distinct mycotoxins. The major types of mycotoxins include aflatoxins, ochratoxins, fumonisins, patulin, trichothecenes [including T-2 and HT-2 toxins, deoxynivalenol (vomitoxin), and nivalenol], and zearalenone (Commission SFSTP, 2007).

The economic consequences of mycotoxin contamination are far-reaching and profound. Firstly, crops that are found to be contaminated with large amounts of mycotoxins often have to be destroyed. Alternatively, contaminated crops that are diverted into animal feed can cause severe health problems in the

animals that consume it, including reduced growth rates, illness, and death.

Moreover, the meat and milk from these animals intended for human consumption may contain toxic residues and biotransformation products. For example, aflatoxins in cattle feed can be metabolized by cows into aflatoxin M_1, which is then secreted in milk (Van Egmond, 1989), whereas ochratoxins in pig feed can accumulate in porcine tissues (Rutqvist *et al.*, 1978).

Mycotoxicoses (caused by mycotoxins) are examples of "poisoning by natural means" analogous to the pathologies caused by exposure to pesticides or heavy metal residues. The symptoms of mycotoxicoses depend on several factors, including the type of mycotoxin, amount and duration of the exposure, and the age, health, and sex of the exposed individual. Other synergistic effects involving genetics and dietary status are relevant being the severity of mycotoxin poisoning compounded by vitamin deficiency, caloric deprivation, alcohol abuse, and infectious disease status. Mycotoxicoses can enhance vulnerability to microbial diseases, worsen the effects of malnutrition, and interact synergistically with other toxins (Bennett and Klich, 2003). Several outbreaks of mycotoxicoses have been reported after the consumption of contaminated food and feed (Reddy and Raghavender, 2007), including the mycotoxins aflatoxin, fumonisins, and ochratoxin A, the most common mycotoxin hazards for humans and animals because of both chronic and acute toxicological manifestations (Dönmez-Altuntaş *et al.*, 2003; Reddy *et al.*, 2010). Other mycotoxins, such as ergot alkaloids (indol alkaloids: ergometrine, ergotamine, ergocristine, etc.), produced by *Claviceps purpurea* and other related species on several members of the Poaceae (Gramineae) family, were responsible for many poisonings (ergotism) in the past. Sporadic reports of ergot poisoning still appear in the current literature.

As stated earlier, because the prevention of contamination by mycotoxins is not possible, the food industry has established an internal regulatory system for the monitoring of contamination levels in foods and feeds. In addition, government regulatory agencies survey for the occurrence of mycotoxins in foods and feeds and establish regulatory limits based on epidemiological data and extrapolations from animal models, being estimations of an appropriate safe dose usually stated as a tolerable daily intake (Kuiper-Goodman, 1994; 1998; Smith

et al., 1995). These regulations, along with consumer demands, have led to the exploitation of alternative food processing and preservation techniques.

Humans have been preserving food to protect it from spoilage for thousands of years. Different techniques including curing with salt, pickling, dehydrating, pasteurizing, and freezing have been applied to maintain the rich nutrition of foodstuffs. Currently, synthetic fungicides have also been used in order to control contaminations. However, the continued applications of fungicides may disrupt the equilibrium of ecosystems, leading to dramatic disease outbreaks, development of fungicide-resistant strains, toxicity to nontarget organisms, and environmental concerns (Lee *et al.*, 2009).

Furthermore, growing concerns on the presence of chemical residues in the food chain (Lee *et al.*, 2008), together with revocations in registration of some of the most effective fungicides, have highlighted the need for developing alternative control strategies or innovative crop protection and postharvest methods with reduced use of conventional fungicides, or without synthetic chemicals at all (Kim, Choi, and Lee, 2003). The use of natural antimicrobial compounds as food preservatives could help avoid the excessive physical processing of food to ensure microbial safety, which often alters organoleptic properties (Juneja, Dwivedi, and Yan, 2012). Plant-derived fungicides are becoming a subject of increasing interest, as it becomes evident that these substances have enormous potential to improve future agrochemical technology. As secondary plant metabolites are biodegradable to nontoxic products, they are potentially useful in integrated pest management programs and could allow for the development of a new class of safer disease-control substances.

Most studies on the use of plant metabolites as food and crop preservatives and their modes of action have been performed on bacteria. There exists far less information on fungi strains, namely mycotoxicogenic ones. This review aims to update the knowledge concerning mycotoxin contamination in food by focusing on plant metabolites as alternative preservatives, the current legislation on mycotoxins, and the main analytical methods for mycotoxin detection.

2   MYCOTOXINS AND MYCOTOXICOSES

Some fungi that grow on animal hosts are responsible for several diseases collectively called *mycoses* that

can vary from mild to life-threatening infections. These fungi are classified as primary pathogens, affecting healthy individuals, or opportunistic pathogens that take advantage of immunocompromised hosts and are more common (Kwon-Chung and Bennett, 1992). Moreover, dietary, respiratory, dermal, and other exposures to toxic fungal metabolites, known as *mycotoxins*, are also responsible for several pathologies collectively called *mycotoxicoses* (Bennett and Klich, 2003). Their symptoms are analogous to the pathologies caused by exposure to pesticides or heavy metal residues and depend on various factors, namely the type of mycotoxin, the amount and duration of the exposure, age, health, and sex of the exposed individual as well as synergistic effects involving genetics, dietary status, and interactions with other toxic insults (Bennett and Klich, 2003). These fungi often grow on edible plants, thus contaminating food and feed. Some contaminations of agricultural food commodities seriously impact human and animal health and reduce the commercial value of crops (Gnonlonfin *et al*., 2013). Mycotoxins are classified as the most important chronic and noninfectious foodborne risk factors (Bhat, Rai, and Karim, 2010). It is known that their ingestion causes general toxic effects, namely immune system

suppression, mutations, cancer, teratogenic effects, and infertility (El-Azab *et al*., 2010).

Fungal infection can result in mycotoxin contamination during growing, harvesting, storage, transport, and processing with mycotoxin accumulation in crops increasing worldwide owing to climatic changes and agricultural practices such as the use of plant varieties being more susceptible to mycotoxin accumulation (Chulze, 2010). Phytopathogenic fungi infect living plants in the field and/or greenhouse, whereas saprophytic fungi colonize plant products postharvest (Bhat, Rai, and Karim, 2010). While only a small number of plant pathogenic fungal species are known to produce mycotoxins, most spoilage fungi secrete a range of toxic metabolites (Berthiller *et al*., 2013). More than 400 different mycotoxins have been identified and classified based on their structural similarities and most important toxic effects (Commission SFSTP, 2007). The most important toxigenic fungi belong to several *Aspergillus*, *Fusarium*, *Penicillium*, and *Alternaria* species (Berthiller *et al*., 2013) and the major types of mycotoxins include aflatoxins, ochratoxins, fumonisins, patulin, trichothecenes (including T-2 and HT-2 toxins, deoxynivalenol (vomitoxin), and nivalenol), and zearalenone (Table 1; Commission SFSTP, 2007).

Table 1 Fungal species producing mycotoxins.

Micro-organisms	Mycotoxins
Alternaria spp. *A. alternata*	Alternariol (AOH), alternariol monomethyl ether (AME), tenuazonic acid, altenuene, altertoxin
Aspergillus spp. *A. flavus*, *A. parasiticus*, *A. versicolor*, *A. nomius*, *A. oryzae*, *A. pseudotamarii*, *A. parvisclerotigenus*, *A. bombycis*, *A. rambellii*, *A. tamari*, *A. pseudotamarii* *A. ochraceus*, *A. niger*	Aflatoxins B1, B2, G1, and G2 Ochratoxin A
Aspergillus spp., *Penicillium* spp., *Byssochlamys* spp.	Patulin
Emericella spp. *E. nidulans*, *E. venezuelensis*	Sterigmatocystin
Fusarium spp. *F. verticillioides*, *F. proliferatum*, *F. subglutinans* *F. graminearum*, *F. culmorum* *F. cerealis*, *F. poae*, *F. culmorum*, *F. graminearum* *F. sporotrichioides*, *F. poae*, *F. equiseti*, *F. acuminatum* *F. graminearum*, *F. semitectum*, *F. equiseti*, *F. crookwellense*, *F. culmorum*	Fumonisins Deoxynivalenol Nivalenol T-2 and HT-2 toxins Zearalenone
Penicillium spp. *P. verrucosum*, *P. viridicatum*, *P. variable*, *P. cyclopium*	Ochratoxin A

Source: Adapted from Commission SFSTP (2007).

Aflatoxins, ochratoxins, and fumonisins are the most common, economically important, and health-relevant mycotoxins for humans and domesticated animals because of both chronic and acute toxicological manifestations (Wu *et al.*, 2011). A brief characterization of these mycotoxins is presented in the following sections.

2.1 Aflatoxins

Aflatoxins are secondary polyketide-derived, toxic, and carcinogenic metabolites primarily produced by *Aspergillus* species, namely *A. flavus*, *A. parasiticus*, and *A. nomius*, which colonize a wide variety of food commodities, including corns, oilseeds (cotton, peanut, rapeseed, coconut, and sunflowers), spices, and tree nuts as well as milk, meat, and dried fruits (Strosnider *et al.*, 2006; Abyaneh, 2013; Magnussen and Parsi, 2013).

There are several different types of aflatoxins, nearly all of which are capable of causing illness in human beings and animals. Aflatoxins accumulate in nucleic acids and upon years of consumption, cause chronic illness. They are known to be hepatotoxic and carcinogenic (Groopman and Kensler, 2005; Magnussen and Parsi, 2013) and have immunosuppressive properties, increasing susceptibility to infections and reducing the protection conferred by vaccination (Venturini *et al.*, 1996; Quadri *et al.*, 2013).

The four major aflatoxins commonly isolated from food and feed are aflatoxins B1 [1], B2 [2], G1 [3], and G2 [4], based on their fluorescence under UV light (blue or green) and relative chromatographic mobility during thin-layer chromatography (Bennett and Klich, 2003). Aflatoxin B1 is the most abundant and the most toxic and has been classified by the World Health Organization as a group A carcinogen because of its proven contribution to the pathogenesis of hepatocellular carcinoma (López *et al.*, 2002).

2.2 Ochratoxins

Ochratoxins consist of an isocoumarin moiety and a phenylalanine moiety linked by an amide bond (Størmer, 1992). Different types of ochratoxins occur naturally, namely ochratoxin A (chlorinated OTA), ochratoxin B (dechlorinated OTA), and ochratoxin C

(ethylated OTA) (Reddy and Bhoola, 2010). Ochratoxins are produced by several species of *Aspergillus* (*A. alliaceus*, *A. auricomus*, *A. carbonarius*, *A. glaucus*, *A. melleus*, *A. niger*) and *Penicillium* (Bennett and Klich, 2003; Commission SFSTP, 2007).

OTA [5] is the most potent member of this group and has been found in barley, oats, rye, wheat, coffee beans, maize, and other plants as well as in meat, wine, dairy, and baked products, such as swine sausages and breads (Abouzied *et al.*, 2002; Reddy and Bhoola, 2010). OTA has a long half-life in humans and is, thus, easily detected in serum. It has been reported to contribute to endemic nephrotoxicity and carcinogenicity in humans and animals and is classified as a possible human carcinogen (group 2B) (Creppy, 1999). Moreover, OTA has been reported to be embryotoxic, teratogenic, and immunotoxic (Kuiper-Goodman and Scott, 1989). The primary effects of OTA are associated with the enzymes involved in phenylalanine metabolism by competing with phenylalanine–transfer ribonucleic acid (tRNA) ligase (Marquardt and Frohlich, 1992). Other mechanisms of action include the formation of deoxyribonucleic acid (DNA) adducts, apoptosis, interference with the cytoskeleton, lipid peroxidation, and inhibition of mitochondrial respiration (Størmer, 1992).

2.3 Fumonisins

Fumonisins are hydrophilic mycotoxins that are thought to be synthesized by condensation of the amino acid alanine into an acetate-derived precursor (Bennett and Klich, 2003; Sweeney and Dobson, 2006). These toxins are produced primarily by *Fusarium verticillioides* and *F. proliferatum*, although a few other *Fusarium* species also produce them. Fumonisins are the most common mycotoxins in maize but can affect also wheat, barley, rice, oats, and rye in the field (Chelkowiski 1989; Scott, 1989). Although *Fusarium* species are considered predominantly field fungi, fumonisin production can occur postharvest when storage conditions are inadequate (Marin *et al.*, 2004). There are at least 28 different forms of fumonisins (A-series, B-series, C-series, and P-series), fumonisin B_1 [6] being the most common and the most economically significant.

Fumonisins are able to interfere with animal sphingolipid metabolism and are known to cause a

fatal disease in horses known as *leukoencephalo-malacia* (Kellerman *et al*., 1990) and pulmonary edema and hydrothorax in swine (Harrison *et al*., 1990). In addition, hepatotoxic and carcinogenic effects have been reported in rats (Gelderblom *et al*., 1991, 1996). Human health effects of these toxins are uncertain but are suspected risk factors for esophageal (Marasas, 2001) and liver (Ueno *et al*., 1997) cancers and, therefore, have been classified by the International Agency for Research on Cancer as group 2B (probably carcinogenic) (Rheeder, Marasas, and Vismer, 2002). Neural tube defects (Missmer *et al*., 2006) and cardiovascular problems (Fincham *et al*., 1992) have also been correlated with these toxins in populations consuming relatively large amounts of food made with contaminated maize (Voss, Smith, and Haschek, 2007).
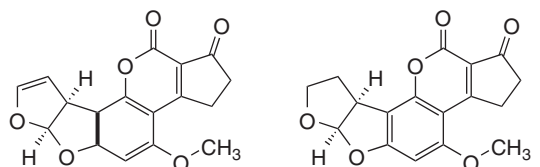
1. Aflatoxin B1
2. Aflatoxin B2
3. Aflatoxin G1
4. Aflatoxin G2
5. Ochratoxin A
6. Fumonisin B1
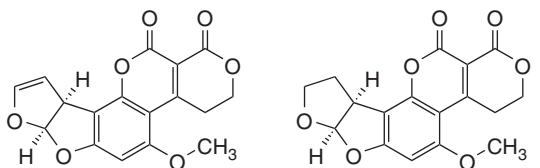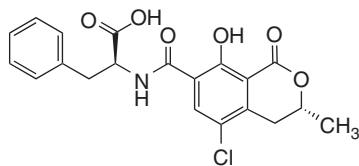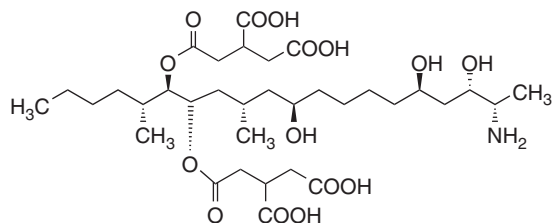
3 ANTIFUNGAL ACTIVITY OF SECONDARY PLANT METABOLITES

Research on plant-derived fungicides is gaining interest, as it becomes evident that these products have enormous potential to improve future agrochemical technology. Although several studies have pointed out the use of plant metabolites to control postharvest fruit, seed, and vegetable diseases (Reddy *et al.*, 1998; Chebli *et al.*, 2004; Tzortzakis and Economakis, 2007; Lopez-Reyes *et al.*, 2010) as well as other phytopathogens (Conte *et al.*, 2007; Camele *et al.*, 2010), many of these products have not yet been exploited commercially. As secondary plant metabolites are biodegradable to nontoxic products, they are potentially useful in integrated pest management programs and could allow for the development of a new class of safer disease-control substances. However, quality attributes such as flavor, odor, color, texture, and nutritional value must be considered together with the guarantee of food safety.

Standard antifungal assays currently used for the evaluation of classical antifungal drugs can also be used as experimental systems to study the antifungal activity of natural compounds such as secondary plant metabolites. However, certain modifications must be considered, taking into account chemical composition, solubility in water, and volatility of these compounds (Lahlou, 2004). The most common tests used to assess the antifungal activity include the disk diffusion (Rammanee and Hongpattarakere, 2011), agar dilution (Viuda-Martos *et al.*, 2007), broth dilution (Rammanee and Hongpattarakere, 2011; Liu *et al.*, 2012), vapor-phase activity test (Inouye *et al.*, 2001), poisoned food technique (Nene and Thapilyal, 2000), and killing curves determination (Samie and Nefefe, 2012). Although several methodologies are available, the lack of standardization in several criteria makes it difficult to compare the results that are obtained by different laboratories. Hadacek and Greger (2000) suggested the standardized broth microdilution methodology to meet the demands of all researchers involved in antifungal susceptibility testing of filamentous fungi. The main variables in these kinds of assays are the plant material used, the methods used to extract the active compounds, the role of solvents, the type of strains used and their growing conditions, the type of culture medium and pH value, incubation time and temperature, and the test assay chosen

(Carson and Riley, 1995; Inouye *et al.*, 2001). Furthermore, several characteristics of antifungals, such as solubility, stability, modes of action, and partial inhibition ability, may also interfere in the methodology adapted (Espinel-Ingroff, Canton and Peman, 2009). Moreover, environmental factors such as temperature, water availability, and gas composition also affect the efficacy of plant metabolites on fungal growth and mycotoxin production (Magan *et al.*, 2003; Cairns-Fuller, 2004; Aldred, Cairns-Fuller, and Magan, 2008). *In vivo* tests regarding compound efficiency, toxicity, and applicability are also required in order to decide whether a future commercial exploitation of the compound can be envisaged.

In this chapter, we reviewed the most relevant reports published during the recent two decades on the antifungal activity of secondary plant metabolites against microorganisms frequently involved in food and storage product contamination by mycotoxins. The data were organized based on a potential relation between chemical composition of plant extracts and their bioactivity. To achieve this goal, plant extracts were divided into two main groups: volatile and nonvolatile compounds. A brief characterization of these compounds, their antifungal activity, and the mode of action is presented in the following sections.

3.1 Volatile Compounds

Essential oils are complex mixtures of several volatile compounds, mainly terpenoids and phenylpropanoids/benzenoids that constitute about 1% of the secondary plant metabolites (Dudareva *et al.*, 2006). Essential oils can be extracted from different parts of aromatic plants by hydrodistillation, steam distillation, and dry distillation, or by cold expression exclusively used for the production of *Citrus* peel oils (Mukherjee, 2002; Bakkali *et al.*, 2008). These products are of particular interest, as they can be employed to storeroom air, and perhaps better than fungicides commonly used for this purpose, to protect stored fruits, crops, or other types of plant materials such as tubers or bulbs. In fact, one of the main advantages of essential oils is their bioactivity in the vapor phase, a characteristic that makes them attractive as a fumigant for stored product protection (Tripati and Dubey, 2004).

For decades, essential oils have been widely used to enhance the flavor and aroma of foods, with many also recognized as valuable food preservatives because of their antimicrobial properties (Hyldgaard, Mygind, and Meyer, 2012). Several studies have confirmed the fungicidal activity of essential oils against toxigenic fungi. For example, Bluma, Amaiden, and Etcheverry (2008) evaluated the effect of 41 essential oils on the growth of *Aspergillus* spp., with *Syzygium aromaticum* (L.) Merrill & Perry (clove), *Hedeoma multiflora* Benth (mountain thyme), and *Mentha pulegium* L. (poleo) oils showing the best results. Also, the effectiveness of *Cinnamomum* sp. (cinnamon) and *Syzygium aromaticum* (L.) Merrill & Perry (clove) leaf oil to control *Penicillium verrucosum*, *Aspergillus ochraceus*, and *Fusarium* spp. and their mycotoxins was shown and a dependency on the environmental conditions identified (Cairns and Magan, 2003; Hope, Jestoi, and Magan, 2003).

More recently, Sumalan, Alexa, and Poiana (2013) reported on the inhibitory effect of several essential oils, namely *Melissa officinalis* L. (lemon balm), *Salvia officinalis* L. (garden sage), *Coriandrum sativum* L. (coriander), *Thymus vulgaris* L. (thyme), *Mentha × piperita* L. (peppermint), and *Cinnamomum zeylanicum* L. (cinnamon) against *Fusarium* mycotoxin production and correlated the results with the antioxidant properties of the oils. The highest inhibitory effect was recorded after 5 days of treatment, with thyme oil found to be the most fungicidal, followed by lemon balm, whereas cinnamon oil showed the best results regarding the inhibition of fumonisin production.

In general, essential oils with high levels of eugenol (allspice, clove bud and leaf, bay, and cinnamomum leaf), cinnamamic aldehyde (cinnamon bark and cassia), citral (lemon myrtle and lime), thymol (thyme), or carvacrol (thyme and origanum) tend to be strong antifungals (Lis-Balchin, Deans, and Eaglesham, 1998; Kumar *et al.*, 2008; Zeng *et al.*, 2011; Sumalan, Alexa, and Poiana, 2013).

Some studies have suggested that the essential oils are more active than the individual components because of the synergistic or cumulative effects between these compounds. For instance, Rammanee and Hongpattarakere (2011) showed that the main components of kaffir (*Citrus hystrix* DC) and acid lime [*Citrus aurantifolia* (Christm.) Swingle] oils were 5 to 6 times less fungicidal than the essential oils, indicating the synergistic activity of many active compounds present in the oils.

As essential oils are a blend of several compounds, they do not act on specific targets in the microbial

cells, and therefore, no resistance or adaptation to the oils has been reported (Carson, Mee, and Riley, 2002). Moreover, several essential oils are classified as generally recognized as safe (GRAS) and have low risk for developing resistance to pathogenic microorganisms.

3.1.1 Terpenes

These compounds result from the condensation of pentacarbonate units (isoprene, C_5H_8) and are classified according to the number of isoprene units in their structure. Monoterpenes (2 units of isoprene) and sesquiterpenes (3 units of isoprene) are the main compounds found in essential oils, although diterpenes (4 units of isoprene) may also occur. In higher plants, terpenes are biosynthesized through the mevalonate-dependent (MVA) and mevalonate-independent or 2C-metil-D-erythritol 4-phosphate/1-deoxi-D-xilu-lose-5-phosphate (MEP/DOXP) pathways. The former takes place in the cytoplasm, whereas the latter occurs in the chloroplasts (Bouwmeester, 2006).

Terpenes, primarily monoterpenic phenols, such as thymol and carvacrol, are considered to be important antifungal compounds. Their activity is related to the presence of hydroxyl groups and delocalized electrons, which explains the lower activity of derivatives, such as carvacrol methyl ether and p-cymene, in comparison to carvacrol (Ultee, Bennik, and Moezelaar, 2002; Ben Arfa et al., 2006). The relevance of thymol and carvacrol is evident in the strong antifungal activities reported for the oils of Thymus spp. (Pina-Vaz et al., 2004; Figueiredo et al., 2008; Zuzarte et al., 2013), Origanum spp. (Soković et al., 2002; López-Malo, Alzamoura, and Palou, 2005; Vale-Silva et al., 2012), and Lavandula multifida L. (Zuzarte et al., 2011). The antifungal effect of these compounds may be due to several mechanisms of action, such as their ability to modify microbial cell permeability to interfere with membrane function, including electron transport, nutrient uptake, protein synthesis, and enzyme activity, as well as their capacity to interact with membrane proteins, causing deformation in structure and functionality (Veres et al., 2003; Bakkali et al., 2008). More recently, Rao et al. (2010) proposed a mechanism of action through disruption of Ca^{2+} and H^+ homeostasis and up-/downregulation of gene transcription similar to Ca^{2+} stress and starvation, while Ahmad et al. (2011) demonstrated disruption of membrane integrity and impairment of ergosterol biosynthesis in Candida spp.

3.1.2 Phenylpropanoids

Several essential oils with antifungal properties are also characterized by the presence of phenyl-propanoids that are molecules with one or more C_6–C_3 fragments with a benzene ring. Phenyl-propanoids are synthesized via the shikimic acid pathway, their main precursors being cinnamic acid and p-hydroxycinnamic acid, originating from the aromatic amino acids phenylalanine and tyrosine, respectively (Baser and Demirci, 2007).

Some studies report the use of Azadirachta indica (L.) Adelb. (neem) oil (e.g., Bankole, 1997; Mossini, de Oliveira, and Kemmelmeier, 2004; Mossini and Kemmelmeier, 2008; Mossini, Arrotéia, and Kemmelmeier, 2009) and Syzygium aromaticum (L.) Merrill & Perry (clove) oil (e.g., Hope, Jestoi and Magan, 2003) to control mycotoxigenic fungi and mycotoxins, with eugenol, isoeugenol, vanillin, saf-role, and cinnamaldehyde being the most thoroughly studied phenylpropanoids. Eugenol, a major compound of clove, cinnamon, and nutmeg essential oils, is one of the most active phenylpropanoids, exhibiting a significant antifungal activity related to the inhibition of fungal growth and mycotoxin production by Aspergillus spp., Penicillium spp., and Fusarium spp. (Hitokoto et al., 1980; Bilgrami, Sinha, and Sinha, 1992; Jayashree and Subramanyam 1999; Cairns and Magan, 2003; Velluti et al., 2003; Kong et al., 2004; Chami et al., 2005; Hope et al., 2005).

Reddy et al. (2007) showed that in rice treated with eugenol ($2.4 \, mg \, g^{-1}$ of grains), A. flavus growth was inhibited entirely and production of aflatoxin B1 prevented. The mechanism of action of this compound has been well investigated on bacteria but studies on fungi strains are scarce. Jayashree and Subramanyam (1999) related the mechanism of action of eugenol to its antioxidant activity owing to the reduction in the activity of enzymes involved in oxidative stress, namely superoxide dismutase, xanthine oxidase, glutathione peroxidase, and two microsomal reductases. Moreover, glucose-6-phosphate dehydrogenase activity was also lower in the presence of eugenol.

Alterations of cell membranes and cell wall structures of proliferating Saccharomyces cerevisiae cells have also been reported (Bennis et al., 2004); however, other studies have shown a stimulation of

aflatoxin production in the presence of this compound (Karapinar 1990; Mahmoud, 1994). These conflicting results may be explained by differences in the time at which aflatoxin production was measured, suggesting that when given sufficient time in culture, most inhibitors are eventually overcome (Holmes, Boston and Payne, 2008).

Some of the most relevant data regarding the antifungal activity of volatile extracts – namely essential oils against fungi affecting food and storage products – are summarized in Table 2. The plant species and part of plants used for essential oil extraction are referred, and in cases in which the oils were analyzed, their main compounds are also listed. The effect of the metabolites on several mycotoxicogenic strains and their mechanism of action (when evaluated) are pointed out as well.

3.2   Nonvolatile Compounds

Nonvolatile compounds are also natural bioactive agents. These products are obtained with solvent extraction methodologies such as maceration, percolation, countercurrent extraction, and supercritical fluids extraction (Samuelsson, 1999). In this group, some of the most relevant compounds with antifungal properties include phenolic compounds, alkaloids, and saponins, as summarized in the following sections.

3.2.1   *Phenolic Compounds*

Phenolic compounds constitute one of the largest and most heterogeneous and ubiquitous groups of higher plant phytochemicals. These compounds exhibit at least one aromatic ring with one or more hydroxyl groups attached. They may vary from simple molecules with a single aromatic ring to highly complex polymeric substances with several rings. The plant shikimate pathway is the entry to the biosynthesis of a large variety of phenolic compounds. Phenols may also have aromatic rings derived by acetate condensation. Several classes of phenolic compounds have been categorized based on their basic skeleton, such as phenolic acids, flavonoids, coumarins, quinones, tannins, stilbenes, and lignans (Evans, 2009).

Phenolic compounds play relevant roles in plant pigmentation, growth, reproduction, protection against ultraviolet light, and resistance to pathogens (Lattanzio, Lattanzio, and Cardinali, 2006). Their antimicrobial properties are mostly related to their ability to penetrate biological membranes (Davidson and Naidu, 2000). Several studies have pointed out the antifungal potential of flavonoids against *Aspergillus* spp. and *Penicillium* spp. (Krämer *et al.*, 1984). For example, the flavonoids cyanidin, peonidin, delphinidin, pelargonidin, kaempferol, and luteolin inhibited aflatoxin production by *A. flavus* (Norton, 1999), whereas tannins from peanut shells of some genotypes showed inhibitory activity against aflatoxin production in *A. parasiticus* (Azaizeh *et al.*, 1990). Another interesting compound is resveratrol, which has demonstrated a particularly wide spectrum of mycotoxin control, though it is a relatively expensive product (Fanelli *et al.*, 2003).

The antioxidant properties of phenolic compounds may be important for their mode of action, but it seems likely that additional factors (bioavailability and uptake, intracellular mobility, and interactions with specific enzymes, metabolite pools, or signaling pathways) are required as well, as the strength of the oxygen-quenching activity is not positively correlated with the inhibition of toxin biosynthesis (Holmes, Boston and Payne, 2008).

3.2.2   *Alkaloids*

Alkaloids are a diverse group of compounds with an alkaline behavior, which present at least one nitrogen atom. Most of these compounds have the nitrogen atom in a heterocyclic ring (typical alkaloids) and are synthesized from amino acid precursors. These compounds can be found in approximately 20% of plant species and play a defensive role against herbivores and pathogens (Ziegler and Facchini, 2008). They are highly valued and widely exploited as pharmaceuticals, stimulants, narcotics, and poisons (Evans, 2009). Regarding food preservation, caffeine is a well-studied inhibitor of aflatoxin production (Holmes, Boston and Payne, 2008). The mechanisms of action suggested for this compound include interference with glucose uptake and inhibition of cyclic adenosine monophosphate (cAMP) phosphodiesterase activity (Holmes, Boston and Payne, 2008). Also *Piper nigrum* L. spices have

Table 2 Antifungal activity of essential oils against strains involved in the contamination of food and storage products.

Plant species	Plant parts	Main compounds	Fungal species	Antifungal activity	Mechanisms of action	References
Azadirachta indica L. (neem) MELIACEAE	Seeds		*Penicillium verrucosum* (isolates K11 and K13) *P. brevicompactum* (isolate K20)	Inhibition of mycelial growth by 47.24% (K11), 82.6% (K13), and 80.1% (K20) Inhibition of sporulation in K13 (75.6%) and K20 (82.58%)	—	Mossini, Arrotéia, and Kemmelmeier (2009)
Citrus aurantifolia (Christm.) Swingle (acid lime) RUTACEAE	Epicarps	Limonene (69.1%), *p*-cymene (12.8%), Essential oil more active than main compounds	*Aspergillus flavus* *A. parasiticus* *A. fumigatus* *A. niger* *Penicillium* sp.	Growth inhibition (MIC and MLC $= 0.56$–2.25 mg mL^{-1}) Aflatoxin inhibition (2.25 mg mL^{-1})	Destruction of plasma and nucleus membrane, loss of cytoplasm, vacuole fusion, and detachment of fibrillar layer in *A. flavus*	Rammanee and Hongpattarakere (2011)
Citrus hystrix DC. (kaffir lime) RUTACEAE		Citronellol (10.7%), limonene (7.3%), linalool (5.8%); essential oil more active than main compounds	*A. flavus* *A. parasiticus* *A. fumigatus* *A. niger* *Penicillium* sp.	Growth inhibition (MIC and MLC $= 0.56$–1.13 mg mL^{-1}) Aflatoxin inhibition (2.25 mg mL^{-1})		
Gnaphalium affine D. Don (Jersey cudweed) ASTERACEAE	Leaves and flowers	Eugenol (18.2%), linalool (10.6%), *trans*-caryophyllene (8.9%)	*Aspergillus niger*	MIC $= 0.20 \pm 0.18$ µg mL^{-1} MLC $= 0.39 \pm 0.14$ µg mL^{-1}		Zeng *et al.* (2011)
			Penicillium citrinum	MIC $= 0.20 \pm 0.15$ µg mL^{-1} MLC $= 0.78 \pm 0.20$ µg mL^{-1}		
			Aspergillus flavus	MIC $= 0.20 \pm 0.20$ µg mL^{-1} MLC $= 0.39 \pm 0.16$ µg mL^{-1}		
Laurus nobilis L. (bay) LAURACEAE			*Fusarium culmorum*	Controlled growth (50–200 ppm)		Hope, Jestoi, and Magan (2003)

Table 2 *(Continued)*

Plant species	Plant parts	Main compounds	Fungal species	Antifungal activity	Mechanisms of action	References
Origanum vulgare L. (oregano) LAMIACEAE	Flowers		*Aspergillus niger* *A. flavus*	Total growth inhibition ($4\,\mu L\,18\,mL^{-1}$ culture medium)		Viuda-Martos *et al.* (2007)
Thymus spathulifolius Hausskn. and Velen. (thyme) LAMIACEAE	Leaves, stem, and flowers	Thymol (36.5%), carvacrol (29.8%), *p*-cymene (10.0%), and γ-terpinene (6.3%)	*Alternaria* spp. *Aspergillus flavus* *Fusarium* spp. *Penicillium* spp.	Growth inhibition ($10\,\mu L\,L^{-1}$)	—	Sokmen *et al.* (2004)
Thymus vulgaris L. (thyme) LAMIACEAE	Leaves, stem, and flowers		*Aspergillus flavus* *A. niger* *A. terreus* *Fusarium oxysporum* *Alternaria alternata*	Broad fungitoxic effect; Complete inhibition of mycelial growth of *A. flavus* ($0.7\,\mu L\,mL^{-1}$). Inhibition of aflatoxin B_1 production ($0.6\,\mu L\,mL^{-1}$)		Kumar *et al.* (2008)
Syzygium aromaticum (L.) Merrill & Perry (clove) MYRTACEAE	Fruits		*Aspergillus niger* *A. flavus*	Total growth inhibition ($6\,\mu L\,18\,mL^{-1}$ culture medium)		Viuda-Martos *et al.* (2007)
Syzygium aromaticum (L.) Merrill & Perry (clove) MYRTACEAE			*Fusarium culmorum*	Controlled growth (50–200 ppm) Inhibition of nivalenol and deoxynivalenol production (100 ppm *in vitro* and 500-ppm concentration on wheat grain)		Hope, Jestoi, and Magan (2003)

MIC –minimal inhibitory concentration; MLC –minimal lethal concentration.

been shown to inhibit aflatoxin production by *A. flavus* (Mabrouk and El-Shayeb, 1992) and the alkaloid piperoctadecalidine identified in the fruits of *Piper longum* L. was most active against aflatoxin production (Lee, Mahoney and Campbell, 2002). However, the mode of action of these alkaloids against aflatoxin production is unclear.

3.2.3 Saponins

Saponins are naturally occurring glycosides with high molecular weight and high polarity that often occur as complex mixtures with slight differences in the nature of the sugars or in the structure of the aglycone. Two types of saponins are recognized based on the structure of the aglycone: the steroidal (commonly tetracyclic triterpenoids) and the pentacyclictriterpenoid types. Both of these have a glycosidic linkage at C-3 and a common biogenetic origin via mevalonic acid and isoprenoid units (Evans, 2009). Saponins interact with the sterols and fatty acids on microbial membranes (Davidson and Naidu, 2000), this antimicrobial activity being related to their aglycone structure. Although saponins are used primarily as emulsifiers and flavoring agents in food and beverages (Juneja, Dwivedi and Yan, 2012), they have the potential to be used also as food preservatives. For example, *Quillaja saponaria* Molina (quillaia) and *Yuca schidigera* Roezl ex Ortgies, two species rich in saponins, are listed as GRAS and, therefore, allowed in food and beverages by Food and Drug Administration (FDA) guidelines. However, very few studies have shown the antifungal potential of these compounds. For example, Jadhav, Sharma, and Salunke (1981) reported their antifungal activity against *Aspergillus* spp.

Some of the most relevant data regarding the antifungal activity of nonvolatile compounds against mycotoxicogenic fungi are summarized in Table 3. If the compounds were isolated from plants, the *taxon* is referred. In addition, their antifungal activity on several mycotoxicogenic strains and the mechanism of action (if evaluated) are cited.

4 FOOD SAFETY LEGISLATION

Contamination of foodstuff with mycotoxin-producing mold species is extremely common

because they can grow on a wide range of substrates under an extremely broad range of environmental conditions, and contamination may occur at any moment during production, harvesting, processing, transportation, and storage (CAST, 2003). Mycotoxins occur with varying severity in agricultural products across the world, but most commonly in hot and humid weather and in inadequate storage facilities. A wide variety of foods may be affected, such as corn, peanuts, cotton seeds, rice, pistachio, almonds, chestnuts, Brazil nuts, and pumpkin seeds, as well as other oily seeds, such as sunflower and coconut. The estimate usually given is that one quarter of the world's crops are contaminated to some extent with mycotoxins (FAO, 2004).

According to Magnussen and Parsi (2013), the incidence of hepatocellular carcinoma (HCC) in some areas of the world constitutes a major public health concern. For instance, in Asia and Africa, HCC accounts for nearly 70% of cancer deaths, whereas in China, it is the third leading cause of cancer mortality, accounting for at least 250 000 deaths per year and with an incidence rate in some countries approaching 100 cases per 100 000 individuals per year. Human exposure to mycotoxins may result from consumption of plant-derived foods that are contaminated with toxins; from ingestion of their metabolites in animal products, such as milk, meat, and eggs; or other types of exposure, such as exposure to air and dust containing toxins (Jarvis, 2002; CAST, 2003).

Kuiper-Goodman (1998), a leading figure in the risk assessment field, ranks mycotoxins as the most important chronic dietary risk factor, higher than synthetic contaminants, plant toxins, food additives, or pesticide residues. Since the discovery of aflatoxins and the recognition of their potential to cause serious effects on humans and animals, many countries have established regulations on mycotoxins in food and feed in recent decades to safeguard health concerns as well as economic interests of producers and traders.

Setting mycotoxin regulations is a complex activity that involves many factors and interested parties, and over 77 countries worldwide have regulations for these toxins (FAO, 2004). The first limits for mycotoxins were set in the late 1960s for the aflatoxins (FAO, 2004), but for substances of this type, there is no threshold below which no harmful effect is observed. In addition, despite advances in scientific and technical knowledge about production and storage techniques, it is impossible to eliminate entirely the occurrence of mycotoxins in the

Table 3 Antifungal activity of nonvolatile secondary plant metabolites against strains involved in food and storage product contamination.

Non-volatile compounds	Bioactive compounds	Plant species	Plant part	Fungal species	Antifungal activity	Mechanism of action	References
Phenolic compounds	Phenolic acid	Caffeic acid / Vanillic acid	*Olea europaea* L. (olive tree) OLEACEAE	Olive cake	*Aspergillus flavus* / *A. parasiticus*	Complete growth inhibition and aflatoxin production (0.2 mg mL^{-1})	—	Aziz *et al.* (1998)
		p-Hydroxybenzoic acid / Protocatechuic acid / Syringic acid / *p*-Coumaric acid				Complete growth inhibition (0.3 mg mL^{-1})	—	Aziz *et al.* (1998)
		Gallic acid	*Juglans regia* L. "Tulare" (walnut) JUGLANDACEAE	Seed coat	*Aspergillus flavus*	Inhibition of aflatoxin biosynthesis	—	Mahoney and Molyneux (2004)
Phenolic compounds	Flavones	Polymethoxylated flavones	*Citrus reticulata* Blanco (mandarin – three varieties) / *Citrus sinensis* (L.) Osbeck (sweet orange – three varieties) RUTACEAE	Peels	*Aspergillus niger*	Growth inhibition (MIC = mg mL^{-1})	—	Liu *et al.* (2012)
	Flavonols	Quercetin	*Olea europaea* L. (olive tree) OLEACEAE	Olive cake	*Aspergillus flavus* / *A. parasiticus*	Complete growth inhibition (0.3 mg mL^{-1})	—	Aziz *et al.* (1998)
Phenolic compounds	Biflavonoids	6,6″-Bigenkwanin	*Ouratea spectabilis* (Mart. ex Engl.) Engl. OCHNACEAE	—	*Aspergillus flavus*	Inhibition of aflatoxin B1 by 88% and B2 by 81.4% (10 µg mL^{-1})	—	Gonçalez, Felicio, and Pinto (2001)
		Amentoflavone	*O. multiflora* (Mart. ex Engl.) Engl. OCHNACEAE	—		Inhibition of aflatoxin B1 by 90.2% and B2 by 91.2% (10 µg mL^{-1})	—	
		7,7″-Dimethoxy-agastisflavone	*O. parviflora* (Mart. ex Engl.) Engl. OCHNACEAE	Leaves		Inhibition of aflatoxin B1 by 85.4% and B2 by 87.7% (10 µg mL^{-1})	—	
		6,6″-Tetradimeth-oxybigenkwanin	Obtained by methylation of 6,6″-bigenkwanin	—		Inhibition of aflatoxin B1 by 88.2% and B2 by 94% (10 µg mL^{-1})	—	

Phenolic compounds	Isoflavonoid derivatives	Glyceollin	—	Glycine max (L.) Merr. (soybean) FABACEAE	—	Decrease in aflatoxin B1 levels by 70% (6.25 µg mL^{-1}) and 95% (62.5 µg mL^{-1}); 11.5% decrease in maximum fungal mass (62.5 µg mL^{-1})	Aspergillus flavus	Song and Karr (1993)
Phenolic compounds	Coumarins	Bergapten	—	Ammi majus L. (bishop's flower) APIACEAE	—	Growth inhibition to 21% (5 mM), Inhibition of aflatoxin release up to 14% (0.1 mM)	Aspergillus flavus	Mabrouk and El-Shayeb (1992)
Phenolic compounds	Coumarins	Psoralene	—	Ficus sycomorus L. (sycamore fig) MORACEAE	—	Growth inhibition to 17% (5 mM) Inhibition of aflatoxin release up to 27% (0.1 mM)		
		Xanthotoxin	—	Ammi majus L. (Bishop's flower) APIACEAE	—	Growth inhibition to 26% (5 mM) Complete inhibition of aflatoxin release (5 mM)	Aspergillus flavus	Mabrouk and El-Shayeb (1992)
Phenolic compounds	Chromones	Khellin	—	Ammi visnaga (L.) Lam. (bisnaga) APIACEAE	—	Growth inhibition to 42% (5 mM) Total inhibition of aflatoxin release (5 mM)	Aspergillus flavus	Mabrouk and El-Shayeb (1992)
		Visnagin	—		—	Growth inhibition to 36% (5 mM) and inhibition of aflatoxin release by more than 75% (0.1 mM)		
Phenolic compounds	Curcuminoids	Curcumin	—	—	—	Inhibition of mycelial growth of A. alliaceus isolate 791 [0.1% (w/v)] Decrease in ochratoxin A production by 70% [0.01% (w/v)]	Aspergillus auricomus A. sclerotiorum A. alliaceus (2 isolates)	Lee et al. (2007)

Table 3 (Continued)

Nonvolatile compounds	Bioactive compounds	Plant species	Plant part	Fungal species	Antifungal activity	Mechanism of action	References
Stilbenes	Resveratrol	—	—	*Aspergillus ochraceus*	Growth inhibition (EC50 = 60 – 190 ppm and OTA inhibition (EC50 = 10 – 130 ppm), dependent on T°C and water availability	—	Cairns-Fuller (2004)
Phenolic compounds Stilbenes	Resveratrol	—	—	*Penicillium verrucosum* *Aspergillus westerdijkiae* (=*A. ochraceus*)	Reduction in fungi and OTA contamination on natural wheat grain by >60% (200 $\mu g\,g^{-1}$, at 0.85–0.995a_w and 15/25°C, 28-day storage)	—	Aldred, Cairns-Fuller, and Magan (2008)
Lignans	Sesamin	*Sesamum indicum* L. (sesame oil) PEDALIACEAE	—	—	Inhibited of OTA by 60% and OTB production by 45% [0.01% (w/v)]	Effect on chloroperoxidase and polyketide synthase activity	Lee *et al.* (2007)
Alkaloids	Berberine	—	—	*Aspergillus flavus*	Mycelial growth inhibition (1–4 mM mL^{-1})	Inhibition of α-amylase activity	Tintu *et al.* (2012)
	Piperine Piperlongumine	*Piper longum* L. (Long pepper) PIPERACEAE	Dried infructescences	*Aspergillus auricomus* *A. sclerotiorum* *A. alliaceus* (2 isolates)	Inhibition of ochratoxin A production [0.001% (w/v)] Inhibition of citrinin production	Effect on polyketide synthesis step of ochratoxin A production	Lee *et al.* (2007)
Saponins	CAY-1	*Capsicum frutescens* L. (cayenne pepper) SOLANACEAE	—	*Aspergillus japonicus* *Penicillium sclerotiorum*, *P. thomii*	Lethal to the germinating (incubated for 8 h) but not the nongerminating (0-h incubation) conidia (>7.5 μM)	—	Lucca *et al.* (2008)

EC50 – half maximal effective concentration; MIC – minimal inhibitory concentration.

food supply. It is, therefore, advisable to set limits "as low as reasonably achievable" (ALARA). This accounts for the variability in regulations across countries on the type of mycotoxin, matrix (type of food), as well as the maximum allowed level. In the European Union, the legislation in force aims to protect consumers by setting maximum levels for mycotoxins in food and feed to ensure that they are not harmful to humans or animals and by keeping mycotoxin levels as low as reasonably achievable following recommended good agricultural, storage, and processing practices (EFSA, 2012). Maximum levels for mycotoxins and certain other contaminants in food are established in Regulation (EC) 1881/2006, and subsequent amendments (for instance, Regulation (EC) 1126/2007, Regulation (EU) 165/2010, and by the last one, Regulation (EU) 1058/2012), and are listed in Table 4. The consolidated version can be consulted in the following link: http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:2006R1881:20121203:EN:PDF.

The analysis of the notifications in the Rapid Alert System for Food and Feed of the European Union (RASFF) shows that mycotoxins are the contaminants with the highest number of notifications, followed by food additives, heavy metals, residues of veterinary medicinal products, migration of organic compounds from food contact materials, and pesticide residues (ASAE, 2006). For instance, in 2010, there were 679 notifications related to mycotoxins (640 to aflatoxins, 25 to ochratoxin A, 9 to aflatoxins and ochratoxin A, 2 to deoxynivalenol, and 3 to fumonisins), and in 2011, the number of mycotoxin notifications decreased only moderately (RASFF, 2011; 2012).

In order to determine the level of food contamination, an analysis should be carried out. In order to perform an analysis, it is necessary first of all to have a representative sample of the food, mainly when, as in this particular case, the distribution of the mycotoxins in the product can be very heterogeneous. This is why, beyond the legislative maximum limits on specific mycotoxins in certain types of foods, the European Union has established requirements for sampling and requires the use of validated methods of analysis.

The basic sets of directives regarding these issues are the following:

- Directive 85/591/EEC that lays down the framework for sampling and analysis methods on community level (European Council, 1985), repealed by Regulation (EC) No. 882/2004 of the European Parliament and of the Council of 29 April 2004 on official controls performed to ensure the verification of compliance with feed and food law, animal health, and animal welfare rules.
- Directives 89/591/EEC and 93/99/EEC that define general principles for the official control of foodstuffs (European Council, 1989, 1993), repealed by Regulation (EC) No. 882/2004 of the European Parliament and of the Council of 29 April 2004 on official controls performed to ensure the verification of compliance with feed and food law, animal health, and animal welfare rules.
- Directive 98/53/EEC that defines sampling methods and the statistical requirements (method performance) for analysis methods (European Commission, 1998), repealed by Commission Regulation (EC) No. 401/2006 laying down the methods of sampling and analysis for the official control of the levels of mycotoxins in foodstuffs – annex I of this regulation establishes the methods of sampling for official control of the levels of mycotoxins in foodstuffs and annex II establishes criteria for sample preparation and methods of analysis used for official control of the levels of mycotoxins in foodstuffs.

Sampling a batch could be the greatest source of variability associated with the mycotoxin analysis procedure, as well as the most crucial step in obtaining reliable results (Köppen *et al.*, 2010). Shephard (2008) refers a trial in which 2.27-kg samples were withdrawn from a lot of nuts and a 100-g subsample of ground material was extracted and analyzed. The results obtained showed that the sampling error contributed 92.7% to the total variability, the sample preparation contributed 7.2%, and the actual sample analysis contributed merely 0.1%.

For this reason, the European Commission has published the Commission Regulation (EC) 401/2006, providing sampling plans for groups of food commodities and accounting for the heterogeneous distribution of mycotoxins. Different sampling plans were also established in other countries, for instance, for aflatoxins in peanuts in the United States (FDA, 2002). Generally, it is possible to recommend that the most effective way to reduce the overall variability of results is to increase the size of the laboratory

Table 4 Maximum levels of mycotoxins allowed in foodstuffs in the European Union.

	Foodstuffs	Maximum levels ($\mu g\,kg^{-1}$)		
2.1	**Aflatoxins**	**B$_1$**	**Sum of B$_1$, B$_2$, G$_1$, and G$_2$**	**M$_1$**
2.1.1	Groundnuts (peanuts) and other oilseeds to be subjected to sorting or other physical treatment, before human consumption or use as an ingredient in foodstuffs, with the exception of – groundnuts (peanuts) and other oilseeds for crushing for refined vegetable oil production	8.0	15.0	—
2.1.2	Almonds, pistachios, and apricot kernels to be subjected to sorting or other physical treatment, before human consumption or use as an ingredient in foodstuffs	12.0	15.0	—
2.1.3	Hazelnuts and Brazil nuts to be subjected to sorting or other physical treatment, before human consumption or use as an ingredient in foodstuffs	8.0	15.0	—
2.1.4	Tree nuts, other than the tree nuts listed in 2.1.2 and 2.1.3 to be subjected to sorting or other physical treatment, before human consumption or use as an ingredient in foodstuffs	5.0	10.0	
2.1.5	Groundnuts (peanuts) and other oilseeds and processed products thereof intended for direct human consumption or use as an ingredient in foodstuffs, with the exception of – crude vegetable oils destined for refining – refined vegetable oils	2.0	4.0	—
2.1.6	Almonds, pistachios, and apricot kernels intended for direct human consumption or use as an ingredient in foodstuffs	8.0	10.0	—
2.1.7	Hazelnuts and Brazil nuts intended for direct human consumption or use as an ingredient in foodstuffs	5.0	10.0	—
2.1.8	Tree nuts, other than the tree nuts listed in 2.1.6 and 2.1.7, and processed products thereof, intended for direct human consumption or use as an ingredient in foodstuffs	2.0	4.0	—
2.1.9	Dried fruit, other than dried figs, to be subjected to sorting or other physical treatment, before human consumption or use as an ingredient in foodstuffs	5.0	10.0	—
2.1.10	Dried fruit, other than dried figs, and processed products thereof, intended for direct human consumption or use as an ingredient in foodstuffs	2.0	4.0	—
2.1.11	All cereals and all products derived from cereals, including processed cereal products, with the exception of foodstuffs listed in 2.1.12, 2.1.15, and 2.1.17	2.0	4.0	—
2.1.12	Maize and rice to be subjected to sorting or other physical treatment before human consumption or use as an ingredient in foodstuffs	5.0	10.0	—
2.1.13	Raw milk, heat-treated milk, and milk for the manufacture of milk-based products	—	—	0.050
2.1.14	Following species of spices: *Capsicum* spp. (dried fruits thereof, whole or ground, including chillies, chilli powder, cayenne, and paprika) *Piper* spp. (fruits thereof, including white and black pepper) *Myristica fragrans* Houtt (nutmeg) *Zingiber officinale* Roscoe (ginger) *Curcuma longa* L. (turmeric) Mixtures of spices containing one or more of the above-mentioned spices	5.0	10.0	—
2.1.15	Processed cereal-based foods and baby foods for infants and young children	0.10	—	—
2.1.16	Infant formulae and follow-on formulae, including infant milk and follow-on milk	—	—	0.025

Table 4   (*Continued*)

	Foodstuffs	Maximum levels ($\mu g\,kg^{-1}$)		
2.1.17	Dietary foods for special medical purposes intended specifically for infants	0.10	—	0.025
2.1.18	Dried figs	6.0	10.0	—
2.2	**Ochratoxin A**			
2.2.1	Unprocessed cereals		5.0	—
2.2.2	All products derived from unprocessed cereals, including processed cereal products and cereals intended for direct human consumption with the exception of foodstuffs listed in 2.2.9, 2.2.10, and 2.2.13		3.0	—
2.2.3	Dried vine fruit (currants, raisins, and sultanas)	10.0		
2.2.4	Roasted coffee beans and ground roasted coffee, excluding soluble coffee	5.0		
2.2.5	Soluble coffee (instant coffee)	10.0		
2.2.6	Wine (including sparkling wine, excluding liqueur wine and wine with an alcoholic strength of not less than 15% vol) and fruit wine	2.0		
2.2.7	Aromatized wine, aromatized wine-based drinks, and aromatized wine-product cocktails	2.0		
2.2.8	Grape juice, concentrated grape juice as reconstituted, grape nectar, grape must and concentrated grape must as reconstituted, intended for direct human consumption	2.0		
2.2.9	Processed cereal-based foods and baby foods for infants and young children	0.50		
2.2.10	Dietary foods for special medical purposes intended specifically for infants	0.50		
2.2.11	Spices, including dried spices *Piper* spp. (fruits thereof, including white and black pepper) *Myristica fragrans* Houtt (nutmeg) *Zingiber officinale* Roscoe (ginger) *Curcuma longa* L. (turmeric) *Capsicum* spp. (dried fruits thereof, whole or ground, including chillies, chilli powder, cayenne, and paprika) Mixtures of spices containing one of the above-mentioned spices	$15\,\mu g\,kg^{-1}$ $30\,\mu g\,kg^{-1}$ until 31.12.2014 $15\,\mu g\,kg^{-1}$ as from 1.1.2015 $15\,\mu g\,kg^{-1}$		
2.2.12	Liquorice (*Glycyrrhiza glabra* L., *Glycyrrhiza inflata* Batalin, and other species)			
2.2.11.1	Liquorice root, ingredient for herbal infusion	$20\,\mu g\,kg^{-1}$		
2.2.11.2	Liquorice extract, for use in food in particular beverages and confectionary	$80\,\mu g\,kg^{-1}$		
2.2.13	Wheat gluten not sold directly to the consumer	8.0		
2.3	**Patulin**			
2.3.1	Fruit juices, concentrated fruit juices as reconstituted and fruit nectars	50		
2.3.2	Spirit drinks, cider, and other fermented drinks derived from apples or containing apple juice	50		
2.3.3	Solid apple products, including apple compote, apple puree intended for direct consumption with the exception of foodstuffs listed in 2.3.4 and 2.3.5	25		
2.3.4	Apple juice and solid apple products, including apple compote and apple puree, for infants and young children and labeled and sold as such	10.0		
2.3.5	Baby foods other than processed cereal-based foods for infants and young children	10.0		
2.4	**Deoxynivalenol**			
2.4.1	Unprocessed cereals other than durum wheat, oats, and maize	1250		

Table 4   (*Continued*)

	Foodstuffs	Maximum levels ($\mu g\,kg^{-1}$)
2.4.2	Unprocessed durum wheat and oats	1750
2.4.3	Unprocessed maize, with the exception of unprocessed maize intended to be processed by wet milling	1750
2.4.4	Cereals intended for direct human consumption, cereal flour, bran and germ as end products marketed for direct human consumption, with the exception of foodstuffs listed in 2.4.7, 2.4.8, and 2.4.9	750
2.4.5	Pasta (dry)	750
2.4.6	Bread (including small bakery wares), pastries, biscuits, cereal snacks, and breakfast cereals	500
2.4.7	Processed cereal-based foods and baby foods for infants and young children	500
2.4.8	Milling fractions of maize with particle size >500 μm falling within CN code 1103 13 or 1103 20 40 and other maize milling products with particle size >50 μm not used for direct human consumption falling within CN code 1904 10 10	750
2.4.9	Milling fractions of maize with particle size ≤500 μm falling within CN code 1102 20 and other maize milling products with particle size ≤500 μm not used for direct human consumption falling within CN code 1904 10 10	1250
2.5	**Zearalenone**	
2.5.1	Unprocessed cereals other than maize	100
2.5.2	Unprocessed maize with the exception of unprocessed maize intended to be processed by wet milling	350
2.5.3	Cereals intended for direct human consumption, cereal flour, bran, and germ as end products marketed for direct human consumption, with the exception of foodstuffs listed in 2.5.6, 2.5.7, 2.5.8, 2.5.9, and 2.5.10	75
2.5.4	Refined maize oil	400
2.5.5	Bread (including small bakery wares), pastries, biscuits, cereal snacks, and breakfast cereals, excluding maize snacks and maize-based breakfast cereals	50
2.5.6	Maize intended for direct human consumption, maize-based snacks and breakfast cereals	100
2.5.7	Processed cereal-based foods (excluding processed maize-based foods) and baby foods for infants and young children	20
2.5.8	Processed maize-based foods for infants and young children	20
2.5.9	Milling fractions of maize with particle size >500 μm falling within CN code 1103 13 or 1103 20 40 and other maize milling products with particle size >500 μm not used for direct human consumption falling within CN code 1904 10 10	200
2.5.10	Milling fractions of maize with particle size ≤500 μm falling within CN code 1102 20 and other maize milling products with particle size ≤500 μm not used for direct human consumption falling within CN code 1904 10 10	300
2.6	**Fumonisins**	**Sum of B1 and B2**
2.6.1	Unprocessed maize, with the exception of unprocessed maize intended to be processed by wet milling	4000
2.6.2	Maize intended for direct human consumption, maize-based foods for direct human consumption, with the exception of foodstuffs listed in 2.6.3 and 2.6.4	1000
2.6.3	Maize-based breakfast cereals and snacks	800

Table 4 (*Continued*)

	Foodstuffs	Maximum levels ($\mu g\,kg^{-1}$)
2.6.4	Processed maize-based foods and baby foods for infants and young children	200
2.6.5	Milling fractions of maize with particle size >500 μm falling within CN code 1103 13 or 1103 20 40 and other maize milling products with particle size >500 μm not used for direct human consumption falling within CN code 1904 10 10	1400
2.6.6	Milling fractions of maize with particle size ≤500 μm falling within CN code 1102 20 and other maize milling products with particle size ≤500 μm not used for direct human consumption falling within CN code 1904 10 10	2000
2.7	**T-2 and HT-2 toxins**	**Sum of T-2 and HT-2 toxins**[a]
2.7.1	Unprocessed cereals and cereal products	

[a] The European Union has called for more data concerning the occurrence, behavior, and toxicity of T-2 and HT-2 toxins, and legal limits for these toxins had been envisaged to be set by July 2008 (European Commission, 2006). The European Food Safety Authority (EFSA) has published an opinion on the toxicity of T-2 and HT-2 toxins (EFSA, 2011) but limits for these toxins have presently not been established.

sample and ensure proper milling and homogenization (Whitaker, 2006). In order to reduce the effect of sampling in qualitative and quantitative analysis, the European Pharmacopoeia has also established a monograph (*2.8.20 Herbal Drugs*: *Sampling and Sample Preparation*), which was introduced in the 6th edition, in order to ensure that the composition of the sample that is used is representative of the batch of material that is being examined (EDQM, 2007). All these types of regulations attempt to avoid the misclassification of a batch, as either false positive (batches with a concentration equal or inferior to the legal limit) or false negative (batches with a concentration superior to the legal limit) classifications can have associated economical or health consequences (Whitaker, 2006).

5 ANALYTICAL METHODS USED TO DETECT MYCOTOXIN CONTAMINATION IN FOOD

The chemical diversity of mycotoxins and their varying concentrations across a wide range of agricultural commodities, foods, and biological samples pose a great challenge to analytical chemists. Each group of compounds and each substrate has different chemical and physical properties; therefore, the methods for separation of toxins from substrates must be developed on a case-by-case basis, requiring specific extraction, cleanup, separation, and detection methods. Consequently, most methods target only individual mycotoxins or, at best, a group of mycotoxins

that are closely related. Moreover, mycotoxins are often produced in trace concentrations, so the sensitivity of the detection systems is essential. Traditional methods have relied upon solvents for cleanup steps and various chromatographic techniques for quantification and require lengthy extraction procedures, expensive chemical cleanup, and use of hazardous materials with production of contaminated wastes. More recently, immunogenic assays that can be applied to samples with little or no cleanup have been developed (Chu, 1991; Chu, 2000).

Since the discovery of mycotoxins, several methodologies for their determination have been developed. These methods were based mainly on thin-layer chromatography (TLC), which was among the first methods used; nowadays, however, the most routinely used methods are those based on enzyme-linked immunosorbent assay (ELISA) or high performance liquid chromatography (HPLC). Gas chromatography (GC) with electron capture detection (ECD) or mass spectrometric (MS) detection is also used in mycotoxin analysis, for example, for trichothecene or patulin determination (Bennett and Klich, 2003). These methods are usually based on labor-intensive sample preparation protocols, followed by traditional chromatographic separation (mostly liquid chromatography – LC).

TLC provides a cheaper alternative to LC-based methods and plays an important role, in developing countries in particular, for surveillance purposes and the control of regulatory limits (Gilbert and Anklam, 2002). Modern sample cleanup techniques,

such as immunoaffinity columns (IAC) or solid-phase extraction (SPE) methods, help simplify protocols and improve selectivity and, therefore, performance characteristics. Alternative techniques, such as, for instance, biosensor-based techniques with surface plasmon resonance detection or hyperspectral imaging, are beginning to be used (Van der Gaag *et al.*, 2003; Elmasry *et al.*, 2012).

According to Stroka and Anklam (2002), certain cultural methods, such as blue fluorescence, particularly in the presence of an enhancer in the medium such as *p*-cyclodextrin; yellow pigmentation, particularly on the undersides of colonies; color change of the yellow pigment to plum-red on exposure of the culture to ammonium hydroxide vapor; and cyclodextrin-enhanced blue fluorescence combined with ammonium hydroxide vapor-induced color change, can also be used. These types of methods offer several advantages (they are inexpensive and do not require expensive equipments and skilled technicians) and also present various drawbacks (positive assessments are not specific with respect to mycotoxin type, the amount of mycotoxin present will not necessarily correlate with the infestation level, and they are generally considered less sensitive than analytical methods). For this reason, this review focuses primarily on the most-used analytical methods for mycotoxin identification and quantification.

5.1 Extraction

After the sampling procedure has been finished, the first step of the analysis is the extraction of the mycotoxins from the biological matrix, with solvent extraction (SE) being the most frequently used. If the sample is liquid, this step is used for cleanup and/or concentration; if the sample is solid, after appropriate grounding, this step is necessary for further processing of sample extract. The solvents typically employed are mixtures of water and polar organic solvents. As the obtained extract will need to be purified or concentrated, usually with open-ended columns, the choice of solvents to extract the analyte should take this step into consideration. Pure organic solvents are preferable for cleanup with silica columns, whereas aqueous mixtures are suitable for reversed-phase or ion exchange cleanup (Patel, 2004). ELISA is an exception, as cleanup is sometimes not required (Chu, 2000).

In some cases, liquid–liquid extraction is used for sample cleanup. This technique is based on the different polarities of the analytes and the liquid matrix in an immiscible organic phase, being able to extract the compounds of interest into one solvent. These techniques are nonspecific and require large volumes of solvents. Shephard (2008) also mentioned the use of other techniques, such as the use of accelerated SE under increased pressure and supercritical fluid extraction (SFE), highlighting that the results have not justified the cost of adapting this instrumentation as an alternative to simple shaking. In the case of SFE, the polar nature of mycotoxins and their poor solubility in carbon dioxide is a major problem that requires the addition of organic solvent modifiers such as methanol or acetonitrile. Furthermore, this technique is not suitable for routine analysis because of high costs and the need for specialized equipments (Shephard, 2008).

As noted earlier, the first methods used for mycotoxin analyses frequently relied on extract cleanup on open columns packed with materials, such as silica or diatomaceous earth, which were eluted with various organic solvent mixtures. Over time, a number of alternatives that use simpler cleanup techniques have been developed, such as the use of SPE using small prepacked columns (cartridges) containing up to 500 mg of stationary phase.

Thus, the basic principle of SPE technology is a variation of chromatographic techniques, based on the differences in partitioning of analytes and interfering compounds between a mobile and a stationary phase, contained within a small disposable cartridge. The cartridges can be packed with silica gel (normal phase) or with bonded (reversed-phase, usually with octadecylsilica, C18) or ion exchange phases. The sample is loaded in one solvent, generally under reduced pressure, rinsed for removal of the maximum number of contaminants, and eluted in another solvent (Shephard, 2008). The use of these small cartridges has several advantages over the traditional liquid–liquid extraction, with the main ones being the use of considerably less solvent and the time, as they are faster in operation. In addition to cleaning sample, they can also be used to preconcentrate the sample, thus providing better detection results. The use of SPE is, thus, widespread and an integral part of many extraction and detection protocols. Several types of stationary phases should be used depending on the toxins that must be separated and quantified.

Another type of cartridges is based on the development of antibodies raised against individual mycotoxins, leading to the introduction of IACs in which a specific antibody is immobilized on a gel contained in a small column. These columns offer excellent recovery of analyte owing to the specificity gained through the use of monoclonal or polyclonal antibodies. The main disadvantage in using these materials is that high costs are involved because each column can only be used once (because of denaturation of antibodies) (Shephard, 2008). Recent developments in this area involve the combination of different antibodies into one column, which allows for the determination of more than one mycotoxin per single sample extract and cleanup. The resulting solution can be analyzed separately for each toxin (Trucksess *et al*., 2007), or a suitable gradient HPLC separation can be developed to achieve a multitoxin determination in a single chromatographic analysis.

Generic SPE, based on molecularly imprinted polymers (MIPs), represents an additional area of interest and has been investigated as a potential novel cleanup system for food analysis. This technique involves the polymerization of selected monomers in a three-dimensional network (polymer) that retains a memory of the shape and functional groups of the analyte molecule around which the polymerization occurs. Subsequent to the removal of the template, the resulting MIP is able to recognize the template (analyte) within a mixture, providing biomimetic recognition elements capable of selective binding/rebinding to the analyte with efficiencies comparable to those of antibody–antigen interactions, and effectively functioning as an artificial antibody (biomimetic receptor). However, while very promising, this method presents several problems, such as inconsistent molecular recognition characteristics, slow binding kinetics of analytes, polymer swelling in unfavorable solvents, and potential sample contamination by template bleeding (Shephard, 2008).

5.2 Separation and Quantification

There are several types of methods available for mycotoxin analysis, but the separation of the analyte from other constituents and impurities that can still be present in the extract is generally performed by chromatography.

Traditionally, the most popular method used for mycotoxin separation is TLC, which is a simple, cost-effective technique that is often used as a mycotoxin-screening assay when low detection limits are not required and which offers the advantage of screening large numbers of samples economically. There are several reviews available on the applications of TLC for mycotoxin analysis; among them is the paper published by Lin *et al*. (1998) in a special issue of the *Journal of Chromatography A*. The use of TLC analysis for mycotoxins is still popular for both quantitative and semiquantitative purposes. This is due to its high throughput of samples and low operating cost and the ease of identification of target compounds using UV–vis spectral analysis.

Several methods have been developed to obtain the best results with each separate class of mycotoxin and with time, several improvements over the classical TLC technique were introduced (Lin *et al*., 1998; Shephard, 2008), such as the use of densitometry to improve quantification, the development of high performance thin-layer chromatography (HPTLC) plates, and the use of two-dimensional analyses to improve the separation. There are several examples of published methods using these techniques for the quantification of mycotoxins, such as, for instance, a two-dimensional HPTLC method for the determination of ochratoxin A at $5\,\mathrm{ng\,g^{-1}}$ in green coffee beans (Ventura *et al*., 2005) or a reversed-phase TLC method validated for the measurement of fumonisin B1 in maize at microgram per gram levels (Shephard and Sewram, 2004).

GC has been applied to the identification and quantification of the presence of a range of mycotoxins in food samples, usually with mass spectrometry (MS and MS/MS), flame ionization detector (FID), or Fourier transform infrared spectroscopy (FTIR) detection techniques, in order to detect the volatile products (Shephard, 2008). Most mycotoxins are not volatile and, therefore, must be derivatized for GC analysis, using chemical reactions with trimethylsilyl (TMS), trifluoroacetyl (TFA), pentafluoropropionyl (PFP), or heptafluorobutyryl (HFB) reactives in order to obtain a volatile material (Shephard, 2008). As there are cheaper and faster alternatives, and considering the instability of analytes (when subject to heat), GC is not widely used, with one exception notwithstanding, namely the analysis of trichothecenes, which has been extensively studied and has generated a large number of publications (Shephard, 2008).

In the recent analysis of mycotoxins, HPLC is the most widely used technique for the identification and

quantification of the major mycotoxins in food. The polar nature of mycotoxins and their solubility in water and organic solvents such as methanol and acetonitrile implies that they can be separated using reversed-phase HPLC columns and this has resulted in a diverse array of methods (Shephard, 2008).

In essence, most of the protocols used for HPLC detection of mycotoxins are very similar. The most commonly found detection methods are UV or fluorescence detectors, which rely on the presence of a chromophore in the molecules. A number of mycotoxins, such as ochratoxin A and zearalenone, possess fluorescence bands and can be detected by fluorescence detectors.

Fluorescence detection offers a number of advantages. In measuring light emitted rather than absorbed, it can frequently achieve lower detection limits than UV detection and as analytical interferences may not absorb and fluoresce at the same wavelengths as the analyte of interest, the fluorescence chromatograms are frequently less prone to interference from coeluting compounds (Shephard, 2008).

Moreover, for those that lack a suitable chromophore, a reaction with a derivatizing agent, either pre- or postcolumn, can be used. This is the case, for instance, of the method included in the European Pharmacopoeia (2010) for determination of aflatoxins in herbal drugs, in the general Chapter 2.8.18. This method consists of three major steps: first, a single-step extraction using sonication; next, a cleanup with one IAC; and finally, separation and quantification of aflatoxin B1 by HPLC using a C18 column and a fluorimetric detection after postcolumn derivatization (with pyridinium hydrobromide perbromide (PBPB), photochemical reactor (PHRED), or electrochemically generated bromine (KOBRA)).

LC coupled with MS has been used for many years, primarily as a technique for mycotoxin confirmation. Presently, the development of bench-top instruments has opened new methodologies for routine analysis of mycotoxins, and LC–MS is the most promising technique for simultaneously screening, identifying, and measuring a large number of mycotoxins. While TLC and other HPLC techniques frequently require derivatization for sensitive detection, LC–MS provides a detection method independent of the formation of chemical derivatives or the UV absorption or fluorescence properties of the molecule.

This method offers the advantages of performing sensitive quantification of mycotoxins with a confirmatory step, allowing multiple detections from the same sample in the same run, and, also, it can be automated, which offers a major advantage over other techniques, such as TLC and ELISA (Cirlini, Dall'Asta, and Galaverna, 2012). Advances in mycotoxin detection by hyphenated chromatographic techniques/MS have been reviewed by Sforza, Dall'asta, and Marchelli (2006). The limiting factors in the use of MS as an analytical tool are the high cost of equipment, complex laboratory requirements, and limitations in the type of the solvents used in extraction and separation.

5.3   Other Methods

There are other separation techniques that can run alongside or substitute chromatographic methods. Among these, immunological assays have been used to successfully detect mycotoxins since the late 1970s. These methods are based on the interaction between an antigen (the analyte of interest) and an antibody with specificity for that analyte, that is, they are methods that rely on immunological principles (Chu, 2000). The development of antibodies for most mycotoxins, owing to their small size, must be conjugated to a carrier molecule, usually a protein (e.g., bovine serum albumin) or a polypeptide, in order to achieve immunogenicity. The conjugation process can also be responsible for decreases in assay selectivity (Shephard, 2008). ELISAs, either as plate- or membrane-based microtiters, are commercially available in kits for qualitative, semiquantitative, and quantitative analysis of the major known mycotoxins from a number of food matrices. These kits became very popular because of their relatively low cost and easy application (they generally do not require cleanup procedures and the extract containing the mycotoxin is analyzed directly), with the disadvantages of possible interferences owing to matrix effects (Goryacheva *et al.*, 2007).

Capillary electrophoresis (CE) is an instrumental technique that allows an effective separation of mycotoxins from potential interfering compounds present in the extract, based on charge- and mass-dependent migration in an electrical field, rather than chromatographic interactions between solute and stationary phases. The fast separations can be accomplished by CE in small volumes of aqueous buffer solutions, excluding the need for using large volumes of organic solvents that are

frequently needed for chromatographic separations. However, as the separation mechanism differs from chromatography, CE separations can have some disadvantages to deal with analytical impurities (in some cases, an initial cleanup step may be necessary), as well as with achieving low limits of detection owing to the limited sample amount that can be introduced into the CE capillary. The development of suitable detectors and methods, such as capillary zone electrophoresis, makes this technique comparable in sensitivity, precision, and accuracy to HPLC methods (Maragos, 2004).

A lateral-flow device (LFD), also called an immunochromatographic test, resembles ELISA principles and is a rapid immunoassay based on the interaction between specific antibodies, immobilized on a membrane strip (instead of microtiter plates), and antibody-coated dyed receptors, for example, latex or colloidal gold, that react with the analyte to form an analyte–receptor complex. Some of these LFDs combine the negative control reaction on the same strip as the sample and require only the addition of the sample solution (Pittet, 2005). The benefits of LFDs include user-friendly format, rapid response, and price. These features make strip tests ideal for applications, such as "on-site," as they can provide either an yes/no determination of the presence of the target mycotoxin or a threshold (semiquantitative) result, typically in 5–10 min, and have all reagents immobilized onto the lateral flow dipstick, without need for use of a specialized equipment. Disadvantages include a subjective interpretation and a much higher cost per test when compared with ELISA (Pittet, 2005).

The development of biosensors for the rapid, reliable, and low cost determination of mycotoxins in foodstuffs is also an area of research that has received considerable attention in recent years. A biosensor is an analytical device that incorporates a specific biological element, for example, an antibody, which creates a recognition event and a physical element that transduces the recognition event into an acoustic, electrical, or optical signal. Immunochemical biosensors that use surface plasmon resonance, quartz crystal microbalance, and screen-printed carbon electrodes have been described for the detection of the major groups of mycotoxins (Pittet, 2005). Each of these biosensors has the potential to offer a high level of convenience in terms of time of analysis but may require coupling with a cleanup technique in order to achieve adequate sensitivity.

6 CONCLUSIONS AND FUTURE PERSPECTIVES

To ensure that food and storage products have the lowest possible mycotoxin concentration, prevention of exposure, good cultural practices during production (crop rotation, tillage, choice of planting date, and management of irrigation and fertilization), harvesting, storage, transportation, marketing, processing, and regulation are required (Munkvold, 2003). As it is practically impossible to prevent all types of fungal contamination, a wide range of preservatives are used to extend the shelf life of food and storage products by inhibiting microbial growth and mycotoxin production. However, the increasingly negative consumer perception of synthetic food additives has spurred an interest in finding natural antifungal alternatives to these traditional preservatives (Zink, 1997).

The use of biopreservation techniques such as plant metabolites can extend the storage life and enhance the safety of food and crop products. The antifungal compounds present in plants are related to their defense mechanisms against infectious and pathogenic agents and may, therefore, be used to inhibit the growth of foodborne pathogens. The application of these compounds for controlling the growth of foodborne pathogens, including mycotoxicogenic fungi, requires the evaluation of the range of activity against the microorganisms, as well as the effect on products' organoleptic properties (Tiwari *et al.*, 2009).

Most studies related to the antifungal activity of secondary plant metabolites have been conducted *in vitro* using microbiological media (Tables 2 and 3), and their efficacy when applied to complex food systems is, therefore, poorly understood. In order to optimize the application of natural antifungals for food and storage product preservation, several key areas require further investigation, including targeting the microorganism of concern, using combined products to provide synergy of activity, identifying the activity/structure relationship, and understanding the processing and storage conditions of the products (Nychas and Skandamis, 2003). It seems that the challenge for practical application of these compounds is to develop optimized low dose combinations that will allow for maintenance of product safety and shelf life while minimizing undesirable flavor and sensory changes. Furthermore, studies aiming to understand the mechanism of action and the

toxicological safety of some compounds are missing, thus hampering its potential utilization for industrial purposes.

Bioactive packaging technologies have been tested in several products and represent an interesting system for the application of plant metabolites for food preservation. The active compounds may be applied by dispersion into the package, coating of the surface of the packaging material, or encapsulation (Kuorwel *et al*., 2011). The latter has been applied in order to increase the effectiveness of plant metabolites and decrease their sensory impact on foods, representing interesting starting points for industries that can try out new natural and safe materials or systems of packaging capable of prolonging the shelf life of foods without lessening their quality and hygiene. However, some secondary plant metabolites can cause unwanted reactions and toxicological evaluations should, therefore, be considered.

The filamentous fungi and their mycotoxins have existed throughout recorded history. The determination of mycotoxins in human food, therefore, will remain a necessity for the future. There are a variety of risk management options that help ensure a safe food supply, ranging from prevention of mold growth to the establishment and enforcement of regulatory limits. However, mycotoxin analysis in food commodities continues to represent a challenge to analytical chemists. The introduction of demands for mycotoxin regulation all over the world, coupled with lower regulatory limits, is increasing over the years. This requires the development of analytical methods that are more sensitive and reliable and with affordable costs. Especially in light of current European legislation set for methods of analysis and appropriate sampling (European Commission, 1998), the major question is the technical feasibility of mycotoxin monitoring; however, the possibility of putting in place prevention strategies for minimizing the risk of exposure to these agents should be maximally encouraged.

7 RELATED ARTICLES

Extraction Methodologies: General Introduction; Supercritical Fluid Extraction; Solid-Phase Microextraction (SPME) and Its Application to Natural Products; Thin-layer Chromatography, with Chemical and Biological Detection Methods; HPLC and Ultra HPLC: Basic Concepts; Innovative Strategies in the Search for Bioactive Plant Constituents; High Throughput Screening of Vegetal Natural Substances; Quality Assessment of Herbal Drugs and Medicinal Plant Products; LC and LC-MS: Techniques and Applications.

REFERENCES

Abouzied, M. M., Horvath, A. D., Podlesny, P. M., *et al*. (2002) *Food Addit. Contam.*, **19**, 755–764.

Abyaneh, M. R. (ed.) 2013. *Aflatoxins – Recent Advances and Future Prospects*. InTech: Rijeka, Croatia.

Ahmad, A., Khan, A., Akhtar, F., *et al*. (2011) *Eur. J. Clin. Microbiol. Infect. Dis.*, **30**, 41–50.

Aldred, D., Cairns-Fuller, V. and Magan, N. (2008) *J. Stored Prod. Res.*, **44**, 341–346.

ASAE. 2006. *Riscos Químicos – Situação Actual*. http://www.asae.pt/aaaDefault.aspx?f=1&back=1&codigono=596065596 5616567AAAAAAAA (accessed 05 March 2013).

Azaizeh, H. A., Pettit, R. E., Sarr, B. A., *et al*. (1990) *Mycopathologia*, **110**, 125–132.

Aziz, N. H., Farag, S. E., Mousa, L. A., *et al*. (1998) *Microbios*, **93**, 43–54.

Bakkali, F., Averbeck, S., Averbeck, D., *et al*. (2008) *Food Chem. Toxicol.*, **46**, 446–475.

Bankole, S. A. (1997) *Appl. Microbiol.*, **24**, 190–192.

Baser, K. H., and Demirci, F. (2007) Chemistry of essential oils, in Flavours and Fragrances – Chemistry*,* Bioprocessing and Sustainability, ed. R. G., Berger, Springer-Verlag, Berlin, pp. 64–65.

Ben Arfa, A., Combes, S., Preziosi-Belloy, L., *et al*. (2006) *Lett. Appl. Microbiol.*, **43**, 149–154.

Bennett, J. W., and Klich, M. (2003) *Clin Microbiol Rev.*, **16**, 497–516.

Bennis, S., Chami, F., Chami, N., *et al*. (2004) *Lett. Appl. Microbiol.*, **38**, 454–458.

Berthiller, F., Crews, C., Dall'Asta, C., *et al*. (2013) *Mol. Nutr. Food Res.*, **57**, 165–186.

Bhat, R., Rai, R. V., and Karim, A. A. (2010) *Comp. Rev. Food Sci. Food Saf.*, **9**, 57–81.

Bilgrami, K. S., Sinha, K. K., and Sinha, A. K. (1992) *Indian J. Med. Res.*, **96**, 171–175.

Bluma, R., Amaiden, M. R., and Etcheverry, M. (2008) *Int. J. Food Microbiol.*, **122**, 114–125.

Bouwmeester, H. J. (2006) *Nat. Biotechnol.*, **24**, 1359–1361.

Cairns, V, Magan, N. (2003). Impact of essential oils on growth and ochratoxin A production by *Penicillium verrucosum* and *Aspergillus ochraceus* on a wheat based substrate, in *Advances in Stored Product Protection*. Credland, P, Armitage, D. M., Bell, C. H., Cogan, P. M. (Eds.). CABI International: Wallingford, UK; 479–485.

Cairns-Fuller, V. (2004) *Dynamics and control of ochratoxigenic strains of* Penicillium verrucosum *and* Aspergillus ochraceus *in the stored grain ecosystem*, Cranfield University, Silsoe, UK PhD Thesis.

Camele, I., De Feo, V., Altieri, L., *et al.* (2010) *J. Med. Food*, **13**, 1515–1523.

Carson, C. F., and Riley, T. V. (1995) *J. Appl. Bacteriol.*, **78**, 264–269.

Carson, C. F., Mee, B. J., and Riley, T. V. (2002) *Antimicrob. Agents Chemother.*, **46**, 1914–1920.

CAST (2003) Mycotoxins*: Risks in Plant, Animal and Human Systems*, Council for Agricultural Science and Technology: Ames, IA, USA, (Report No. 139).

Chami, N., Bennis, S., Chami, F., *et al.* (2005) *Oral Microbiol. Immun.*, **20**, 106–111.

Chebli, B., Hmamouchi, M., Achouri, M., *et al.* (2004) *J. Essent. Oil Res.*, **16**, 507–511.

Chelkowiski, J. (1989). Formation of mycotoxins produced by Fusarium in heads of wheat, triticale and rye, in Fusarium mycotoxins*, taxonomy and pathogenicity*. Chelkowski, J. (ed.) Elsevier: Amsterdam, 63–84.

Chu, F. S. (1991) Current immunochemical methods for mycotoxin analysis, in Immunoassays for Trace Chemical Analysis*: Monitoring Toxic Chemicals in Humans, Food and the Environment*. Vanderlaan, M., Stanker, L. H., Watkins, B. E., Roberts, D. W. (eds.) American Chemical Society: Washington, DC, 140–157.

Chu, F. S. (2000) Mycotoxin analysis: immunological techniques, in *Foodborne Disease Handbook*, vol. **3**: Plant Toxicants. Hui, Y. H., Smith, R. A. Spoere, D. G. (eds.) Marcel Dekker: New York, 683–713.

Chulze, S. N. (2010) *Food Add. Contam.*, **27**, 651–657.

Cirlini, M., Dall'Asta, C., and Galaverna, G. (2012) *J. Chromatogr. A*, **1255**, 145–152.

Commission SFSTP, Guédon, D., Brum, M., *et al.* (2007) *STP Pharma Pratiques*, **17**, 209–225.

Conte, A., Speranza, B. S., Sinigaglia, M., *et al.* (2007) *J. Food Protect.*, **70**, 1896–1900.

Creppy, E. E. (1999) *Toxin Rev.*, **18**, 277–293.

Davidson, P. M., Naidu, A. S.. (2000). Phytophenols, in *Natural Food Antimicrobial Systems*. Naidu, A. S. (ed.) CRC Press: Boca Raton, FL; 265–293.

Dönmez-Altuntaş, H., Hamurcu, Z., Imamoglu, N., *et al.* (2003) *Nahrung*, **47**, 33–35.

Dowd, P. F. (2003) *J. Toxicol. Toxin Rev.*, **22**, 327–350.

Dudareva, N., Negre, F., Nagegowda, D. A., *et al.* (2006) *Crit. Rev. Plant Sci.*, **25**, 417–440.

EDQM (2007) *European Pharmacopoeia*, 6 edn, Council of Europe (CoE) — European Directorate for the Quality of Medicines (EDQM), Strasbourg, France.

EFSA (2011) *EFSA J.*, **9**, 2481–2668.

EFSA. (2012). EFSA Topic: Mycotoxins http://www.efsa.europa.eu/en/topics/topic/mycotoxins.htm (accessed 15 March 2013).

El-Azab, S., Abdelhamid, A. M., Hend, S., *et al.* (2010) *Toxicol. Environ. Chem.*, **92**, 383–389.

Elmasry, G., Kamruzzaman, M., Sun, D., *et al.* (2012) *Crit. Rev. Food Sci. Nutr.*, **52**, 999–1023.

Espinel-Ingroff, A., Canton, E., and Peman, J. (2009) *Curr. Fungal Infect. Rep.*, **3**, 133–141.

European Commission (1998) *Off. J.*, **L 201**17 July 1998, 0093–0101.

European Commission (2006) *Off. J.*, **L 364**20 November 2006, 0005–0024.

European Council (1985) *Official Journal*, **L 372**31 December 1985, 0050–0052.

European Council (1989) *Official Journal*, **L 186**30 June 1989, 0023–0026.

European Council (1993) *Official Journal*, **L 290** 24 November 1993, 0014–0017.

European, Pharmacopoeia (2010) Determination of aflatoxin B1 in herbal drugs, general chapter 2.8.18. European Pharmacopoeia 7. Strasbourg, France: Council of Europe; (vol. **1**) 2010.

Evans, W. C. (2009) *Trease and Evans Pharmacognosy*, 16 edn, London, Saunders Elsevier.

Fanelli, C., Taddei, F., Trionfetti-Nisini, P., *et al.* (2003) *Aspects Appl. Biol.*, **68**, 63–71.

FAO (Food and Agricultural Organization). (2004). Food and Agricultural Organization of the United Nations food and nutrition (paper 81). Worldwide regulations for mycotoxins in food and feed in 2003.

FDA (2002) *Investigative Operations Manual*, Food and Drug Administration, Washington, USA.

Figueiredo, A. C., Barroso, J., Pedro, L., *et al.* (2008) *Curr. Pharm. Design*, **14**, 3120–3140.

Fincham, J. E., Marasas, W. F., Taljaard, J. J., *et al.* (1992) *Atherosclerosis*, **94**, 13–25.

Gelderblom, W. C., Kriek, N. P., Marasas, W. F., *et al.* (1991) *Carcinogenesis*, **12**, 1247–1251.

Gelderblom, W. C. A., Smuts, C. M., Abel, S., *et al.* (1996) *Food Chem. Toxicol.*, **34**, 361–369.

Gilbert, J., and Anklam, E. (2002) *Trends Anal. Chem.*, **21**, 468–486.

Gnonlonfin, G. J. B., Hell, K., Adjovi, Y., *et al.* (2013) *Crit. Rev. Food Sci.*, **53**, 349–365.

Gonçalez, E., Felicio, J. D., and Pinto, M. M. (2001) *Braz. J. Med. Biol. Res.*, **34**, 1453–1456.

Goryacheva, I. Y., De Saeger, S., Eremin, S. A., *et al.* (2007) *Food Addit. Contam.*, **24**, 1169–1183.

Groopman, J. D., and Kensler, T. W. (2005) *Toxicol. Appl. Pharmacol.*, **206**, 131–137.

Hadacek, F., and Greger, H. (2000) *Phytochem. Anal.*, **11**, 137–147.

Harrison, L. R., Colvin, B. M., Greene, J. T., *et al.* (1990) *J. Vet. Diagn. Investig.*, **2**, 217–221.

Hitokoto, H., Morozumi, S., Wauke, T., *et al.* (1980) *Appl. Environ. Microbiol.*, **39**, 818–822.

Holmes, R. A., Boston, R. S., and Payne, G. A. (2008) *Appl. Microbiol. Biotechnol.*, **78**, 559–572.

Hope, R., Jestoi, M., and Magan, N. (2003) Multitarget environmental approach for control of growth and toxin production *by Fusarium culmorum* using essential oils and antioxidants, in *Advances in Stored Product Protection*, eds. P., Credland, D. M., Armitage, C. H., Bell, P. M., Cogan and E., Highley, CABI Publishing, Wallingford, United Kingdom, pp. 486–492.

Hope, R., Cairns-Fuller, V., Aldred, D., *et al.* (2005) *BCPC Crop Sci. Tech.*, **5B**, 429–436.

Hyldgaard, M., Mygind, T., and Meyer, R. L. (2012) *Front Microbiol.*, **3**, 12.

Inouye, S., Tsuruoka, T., Uchida, K., *et al.* (2001) *Microbiol. Immunol.*, **45**, 201–208.

Jadhav, S. J., Sharma, R. P., and Salunkhe, D. K. (1981) *Crit. Rev. Toxicol.*, **9**, 21–104.

Jarvis, B. B. (2002) *Adv. Expt. Med. Biol.*, **504**, 43–52.

Jayashree, T., and Subramanyam, C. (1999) *Lett. Appl. Microbiol.*, **28**, 179–183.

Juneja, V. K., Dwivedi, H. P., and Yan, X. (2012) *Annu. Rev. Food Sci. Technol.*, **3**, 381–403.

Karapinar, M. (1990) *Int. J. Food Microbiol.*, **10**, 193–199.

Kellerman, T. S., Marasas, W. F., Thiel, P. G., *et al.* (1990) *Onderstepoort J. Vet. Res.*, **57**, 269–275.

Kim, M.-K., Choi, G.-J., and Lee, H.-S. (2003) *J. Agric. Food Chem.*, **51**, 1578–1581.

Kong, Q. L., Song, Y. H., Zhang, L., *et al.* (2004) *Acta Agricul. Shanghai*, **20**, 68–72.

Köppen, R., Koch, M., Siegel, D., *et al.* (2010) *Appl. Microbiol. Biotechnol.*, **86**, 1595–1612.

Krämer, R. P., Hindorf, H., Jha, H. C., *et al.* (1984) *Phytochemistry*, **23**, 2203–2205.

Kuiper-Goodman, T. (1994). Prevention of human mycotoxicoses through risk assessment and risk management. In *Mycotoxins in Grain: Compounds Other Than Aflatoxin*. Miller, J. D., Trenholm, H. L. (Ed.). Eagan Press: St. Paul MN; pp. 439–469.

Kuiper-Goodman, T. (1998). Food safety: mycotoxins and phycotoxins in perspective. In *Mycotoxins and phycotoxins – developments in chemistry, toxicology and food safety*. Miraglia, M., van Edmond, H. P., Brera, C., and Gilbert, J. (Eds.). Alaken Inc: Fort Collins, Colorado; pp. 25–48.

Kuiper-Goodman, T., and Scott, P. M. (1989) *Biomed. Environ Sci.*, **2**, 179–248.

Kumar, A., Shukla, R., Singh, P., *et al.* (2008) *Innov. Food Sci. Emerg. Technol.*, **9**, 575–580.

Kuorwel, K. K., Cran, M. J., Sonneveld, K., *et al.* (2011) *J. Food Sci.*, **76**, 164–177.

Kwon-Chung, K. J., and Bennett, J. E. (1992) *Medical Mycology*, Lea & Febiger, Philadelphia, PA.

Lahlou, M. (2004) *Phytother. Res.*, **1**, 435–448.

Lattanzio, V, Lattanzio, VMT, Cardinali . (2006). Role of phenolics in the resistance mechanisms of plants against fungal pathogens and insects. Phytochemistry. In *Advances in Research*. Imperato, F. (ed.) Research Signpost: India; 23–67.

Lee, S. E., Mahoney, N. E., and Campbell, B. C. (2002) *J. Microbiol. Biotechnol.*, **12**, 679–682.

Lee, S. E., Park, B. S., Bayman, P., *et al.* (2007) *Food Addit. Contam.*, **24**, 391–397.

Lee, Y.-S., Kim, J., Shin, S.-C., *et al.* (2008) *Flavour Frag. J.*, **23**, 23–28.

Lee, Y.-S., Kim, J., Lee, S.-G., *et al.* (2009) *Pestic. Biochem. Phys.*, **93**, 138–143.

Lin, L., Zhang, J., Wang, Y., *et al.* (1998) *J. Chromatogr. A*, **815**, 3–20.

Lis-Balchin, M., Deans, S. G., and Eaglesham, E. (1998) *Flavour Frag. J.*, **13**, 98–104.

Liu, L., Xu, X., Cheng, D., *et al.* (2012) *J. Agric. Food Chem.*, **60**, 4336–4341.

López, C., Ramos, L., Bulacio, L., *et al.* (2002) *Medicina (B Aires)*, **62**, 313–316.

López-Malo, A., Alzamora, M. S., and Palou, E. (2005) *Int. J. Food Microbiol.*, **99**, 119–128.

Lopez-Reyes, J. G., Spadaro, D., Gullino, M. L., *et al.* (2010) *Flavour Fragr. J.*, **25**, 171–177.

Lucca, A. J., Mich, M., Boue, S., *et al.* (2008) *Am. J. Enol. Vitic.*, **59**, 67–72.

Mabrouk, S. S., and El-Shayeb, N. M. A. (1992) *World J. Microbiol. Biotechnol.*, **8**, 60–62.

Magan, N., Hope, R., Cairns, V., *et al.* (2003) *Eur. J. Plant Pathol.*, **109**, 723–730.

Magda, C., and Castillo, P. (2002) *EOLSS*, **7**, 1–6.

Magnussen, A., and Parsi, M. A. (2013) *World J. Gastroenterol.*, **19**, 1508–1512.

Mahmoud, A. L. (1994) *Lett. Appl. Microbiol.*, **19**, 110–113.

Mahoney, N., and Molyneux, R. J. (2004) *J. Agric. Food Chem.*, **52**, 1882–1889.

Maragos, C. M. (2004) *Food Addit. Contam.*, **21**, 803–810.

Marasas, W. F. (2001) *Environ. Health Persp.*, **109**, 239–243.

Marin, S., Magan, N., Ramos, A. J., *et al.* (2004) *J. Food Protect.*, **67**, 1792–1805.

Marquardt, R. R., and Frohlich, A. A. (1992) *J. Anim. Sci.*, **70**, 3968–3988.

Missmer, S. A., Suarez, L., Felkner, M., *et al.* (2006) *Environ. Health Persp.*, **114**, 237–241.

Mossini, S. A., and Kemmelmeier, C. (2008) *Int. J. Mol. Sci.*, **9**, 1676–1684.

Mossini, S. A., de Oliveira, K. P., and Kemmelmeier, C. (2004) *J. Basic Microbiol.*, **44**, 106–113.

Mossini, S. A., Arrotéia, C. C., and Kemmelmeier, C. (2009) *Toxins*, **1**, 3–13.

Mukherjee, P. K. (2002) *Quality Control of Herbal Drugs. An approach to evaluation of botanicals*, Business Horizons, New Delhi.

Munkvold, G. P. (2003) *Ann. Rev. Phytopathol.*, **41**, 99–116.

Nene, Y., and Thapilyal, L. (2000) *Poisoned food technique of fungicides in plant disease control*, 3 edn, New Delhi, Oxford and IBH Publishing Company.

Norton, R. A. (1999) *J. Agric. Food Chem.*, **47**, 1230–1235.

Nychas, G-JE, Skandamis, P. N.. (2003). Antimicrobials from herbs and spices. *In Natural Antimicrobials for the Minimal Processing of Foods*. Roller, S (Ed.). Woodhead Publishing Limited: Cambridge; 176–200.

Patel, P. (2004). Mycotoxin analysis: current and emerging technologies. In *Mycotoxins in food: Detection and Control*. Magan, N, Olsen, N. (Ed.). Woodhead Publishing Limited: Cambridge, England; 88–110.

Pina-Vaz, C., Rodrigues, A. G., Pinto, E., *et al.* (2004) *J. Eur. Acad. Dermatol. Venereol.*, **18**, 73–78.

Pirbalouti, A. G., Hamedi, B., Abdizadeh, R., *et al.* (2011) *J. Med. Plants Res.*, **5**, 5089–5093.

Pittet, A. (2005) *Mitt. Lebensm. Hyg.*, **96**, 424–444.

Quadri, S. H., Niranjan, M. S., Chaluvaraju, K. C., *et al.* (2013) *Intern. J. Chem. Life Sci.*, **2**, 1071–1078.

Rammanee, K., and Hongpattarakere, T. (2011) *Food Bioprocess Technol.*, **4**, 1050–1059.

Rao, A., Zhang, Y., Muend, S., *et al.* (2010) *Antimicrob. Agents Chemother.*, **54**, 5062–5069.

RASFF (Rapid Alert System for Food and Feed). (2011). The Rapid Alert System for Food and Feed 2010 Annual Report.

RASFF (Rapid Alert System for Food and Feed). (2012). The Rapid Alert System for Food and Feed 2011 Annual Report.

Reddy, L., and Bhoola, K. (2010) *Toxins*, **2**, 771–779.

Reddy, B. N., Raghavender, C. R.. (2007). Outbreaks of aflatoxicoses in India. *AJFAND* **7**: No 5 1–15.

Reddy, M. V. B., Angers, P., Gosselin, A., *et al*. (1998) *Phytochemistry*, **42**, 1515–1520.

Reddy, C. S., Reddy, K. R. N., Prameela, M., *et al*. (2007) *J. Mycol. Plant Pathol.*, **37**, 87–94.

Reddy, K. R. N., Nurdijati, S. B., and Salleh, B. (2010) *Asian J. Plant Sci.*, **9**, 126–133.

Rheeder, J. P., Marasas, W. F. O., and Vismer, H. F. (2002) *Appl. Environ. Microbiol.*, **68**, 2102–2105.

Rutqvist, L., Björklund, N. E., Hult, K., *et al*. (1978) *Appl. Environ. Microbiol.*, **36**, 920–925.

Samie, A., and Nefefe, T. (2012) *J. Med. Plants Res.*, **6**, 465–478.

Samuelsson, G. (1999) *Drugs of Natural Origin: A Textbook of Pharmacognosy*, 4 edn, Swedish Pharmaceutical Press, Stockholm.

Scott, P. M.. (1989). The natural occurrence of trichothecene. In Trichothecene Toxicosis: Pathophysiological Effects, vol. **1**. Beasley, V. R. (ed.) CRC Press: Boca Raton, FL; 2–26.

Sforza, S., Dall'asta, C., and Marchelli, R. (2006) *Mass Spectromet. Rev.*, **25**, 54–76.

Shephard, G. S. (2008) *Chem. Soc. Rev.*, **37**, 2468–2477.

Shephard, G. S., and Sewram, V. (2004) *Food Addit. Contam.*, **21**, 498–505.

Smith, J. E., Solomons, G., Lewis, C., *et al*. (1995) *Nat. Toxins*, **3**, 187–192.

Sokmen, A., Gulluce, M., Akpulat, H. A., *et al*. (2004) *Food Control*, **15**, 627–634.

Soković, M., Tzakou, O., Pitarokili, D., *et al*. (2002) *Nahrung*, **46**, 317–320.

Song, D. K., and Karr, A. L. (1993) *J. Chem. Ecol.*, **19**, 1183–1194.

Størmer, F. C.. (1992). Ochratoxin A – a mycotoxin of concern. In Handbook of Applied Mycology, vol. **5**, Mycotoxins in Ecological Systems. Bhatnagar, D., Lillehoj, E. B., Arora, D. K. (eds.) Marcel Dekker, Inc: New York; 403–432.

Stroka, J., and Anklam, E. (2002) *Trends Anal. Chem.*, **21**, 90–95.

Strosnider, H., Azziz-Baumgartner, E., Banziger, M., *et al*. (2006) *Environ Health Perspect.*, **114**, 1898–1903.

Sumalan, R.-M., Alexa, E., and Poiana, M.-A. (2013) *Chemistry Cent. J.*, **7**, 32–44.

Sweeney, M. J., and Dobson, A. D. W. (2006) *FEMS Microbiol. Lett.*, **175**, 149–163.

Tintu, I., Dileep, K. V., Augustine, A., *et al*. (2012) *Chem. Biol. Drug Des.*, **80**, 554–560.

Tiwari, B. K., Valdramidis, V. P., O'Donnell, C. P., *et al*. (2009) *J. Agric. Food Chem.*, **57**, 5987–6000.

Tripati, P., and Dubey, N. K. (2004) *Post Harvest Biol. Technol.*, **32**, 235–245.

Trucksess, M. W., Weaver, C. M., Oles, C. J., *et al*. (2007) *J. AOAC Int.*, **90**, 1042–1049.

Tzortzakis, N. G., and Economakis, C. D. (2007) *Innov. Food Sci. Emerging Technol.*, **8**, 253–258.

Ueno, Y., Iijima, K., Wang, S. D., *et al*. (1997) *Food Chem Toxicol.*, **35**, 1143–1150.

Ultee, A., Bennik, M. H. J., and Moezelaar, R. (2002) *Appl. Environ. Microbiol.*, **68**, 1561–1568.

Vale-Silva, L., Silva, M. J., Oliveira, D., *et al*. (2012) *J. Med. Mycrobiol.*, **61**, 252–260.

Van der Gaag, B., Spath, S., Dietrich, H., *et al*. (2003) *Food Control*, **14**, 251–254.

Van Egmond, H. P.. (1989). Aflatoxin M1: occurrence, toxicity, regulation. In *Mycotoxins in Dairy Products*. Van Egmond, H. P. (Ed.). Elsevier Applied Science: London; pp 11–55.

Velluti, A., Sanchis, V., Ramos, A. J., *et al*. (2003) *Int. J. Food Microbiol.*, **89**, 145–154.

Ventura, M., Anaya, I., Broto-Puig, F., *et al*. (2005) *J. Food Prot.*, **68**, 1920–1922.

Venturini, M. C., Quiroga, M. A., Risso, M. A., *et al*. (1996) *J. Comp. Pathol.*, **115**, 229–237.

Veres, K., Varga, E., Dobos, Á., *et al*. (2003) *Chromatographia*, **57**, 95–98.

Viuda-Martos, M., Ruiz-Navajas, Y., Fernández-López, J., *et al*. (2007) *J. Food Safety.*, **27**, 91–101.

Voss, K. A., Smith, G. W., and Haschek, W. M. (2007) *Anim. Feed Sci. Technol.*, **137**, 299–325.

Whitaker, T. B. (2006) *Food Addit. Contam.*, **23**, 50–61.

Wu, F., Narrod, C., Tiongco, M., *et al*. (2011) *Afla Control*, **4**, 1–20.

Zeng, W.-C., Zhu, R.-X., Jia, L.-R., *et al*. (2011) *Food Chem. Toxicol.*, **49**, 1322–1328.

Ziegler, J., and Facchini, P. J. (2008) *Annu. Rev. Plant Biol.*, **59**, 735–769.

Zink, D. L. (1997) *Emerg. Infect. Dis.*, **3**, 467–469.

Zuzarte, M., Vale-Silva, L., Gonçalves, M. J., *et al*. (2011) *Eur. J. Clin. Microbiol. Infect. Dis.*, **3**, 1359–1366.

Zuzarte, M., Gonçalves, M. J., Cavaleiro, C., *et al*. (2013) *Ind. Crop Prod.*, **44**, 97–103.

Quality Assessment of Herbal Drugs and Medicinal Plant Products

Iqbal Ahmad[1], Mohd Sajjad Ahmad Khan[2] and Swaranjit Singh Cameotra[2]

[1]*Department of Agricultural Microbiology, Aligarh Muslim University, Aligarh, India* and [2]*Institute of Microbial Technology, Chandigarh, India*

1 INTRODUCTION

Plants serve as the most valuable source for curing many diseases. The use of herbs as medicine is the oldest form of healthcare known to humanity and has been used in all cultures throughout history (Barnes, Anderson, and Philipson, 2007). There is a great demand for herbal medicines in developed and developing countries and therefore a great potential in the global market (Jones, Chin, and Kinghorn, 2006). Herbal medicines include herbs, herbal extracts and preparations, and finished products. Herbal drugs are the unprocessed parts of a plant or the whole plant. An herb can be a root, stem, leaf, bark, seed, fruit, flower, wood, rhizomes, or other plant parts, which may be entire, fragmented, or powdered. While herbal preparations are made from one or more herbs and include comminuted or powdered materials or extracts, tinctures and fatty oils of herbal materials, which may be produced by extraction, fractionation, purification, concentration, or other physical or biological processes (WHO, 2000; Shinde *et al*., 2009).

Both traditional and modern medicines rely on medicinal plants as a source of raw material. Nowadays, plant materials are used throughout the industrialized and developing world as home remedies and ingredients for the pharmaceutical industries (Bandaranayake, 2006). They are distributed worldwide, but they are most abundant in tropical countries. It is estimated that about 25% of all modern medicines are directly or indirectly derived from higher plants (Ackernecht, 1973; WHO, 2005). During recent decade, there has been increasing public interest and acceptance of natural therapies in both developing and developed countries. Owing to poverty and limited access to modern medicine, about 80% of the world's population, especially in the developing countries, use herbal medicine as their source of primary healthcare (Mukherjee, 2002; Bodeker *et al*., 2005). In these communities, traditional medical practice is often viewed as an integral part of their culture. In the West, people are attracted to herbal therapies for many reasons, the most important reason being that, like our ancestors, it is believed they will help us live healthier lives. Herbal medicines are often viewed as a balanced and moderate approach to healing.

Herbal medicines are used very commonly in various health practices or therapies of traditional medicines such as Ayurveda, Unani, Chinese, African and European medicines, Naturopathy, Osteopathy, and Homeopathy (Shinde *et al*., 2009). Traditional medicine comprises of knowledge, skills, and health practices and approaches based on theories and beliefs that are used in the maintenance of health, prevention, diagnosis, treatment, and improvement of physical and mental illnesses (WHO, 2004a,b).

There is no single herb recommended for a particular health disorder, and there is no single disease linked to a particular herb. There are different chemical compounds present in the plant that are responsible for its medicinal benefit, and these compounds are termed as *active ingredients*. Their presence depends on a number of factors including the plant species, the time and season of harvest, the type of soil, and the way the herb is prepared. In most of the cases, the active principle for pharmacological action of an herbal drug is unknown and their potency is lesser compared to synthetic drugs (Kunle *et al*., 2012). These drugs are very effectively used against the chronic diseases, and occurrence of side effects is fewer (Bandaranayake, 2006).

With the ever-increasing use of herbal medicines worldwide and the rapid expansion of the global market for these products, the safety and quality of medicinal plant materials and finished herbal medicinal products have become a major concern for health authorities, pharmaceutical industries, and the public. The production of herbal drugs involves three basic steps: (i) identification, (ii) evaluation, and (iii) standardization. Identification of herbs is based on macroscopical and microscopical features. Macroscopical feature involves odor, taste, color, size, shape, and special feature of plant and microscopically involves leaf content, trichome, stomata, and so on. Certain microscopic features and chemical test come under evaluation and standardization of herbal drugs. Evaluation of drugs means confirmation of its identity, determination of its quality and purity, and detection of adulteration (Kokate, Purohit, and Gokhale, 2006). Standardization is used to describe all measures that are taken during the manufacturing process and quality control, leading to a reproducible quality. It involves the study from germination of a plant to its clinical application. It also includes the preparation of an herbal drug to a defined content of a constituent or a group of substance with known therapeutic activity, respectively, by the addition of excipients or by mixing herbal drugs preparation (WHO, 1988, 1998). In other words, this step ensures that every packet of medicine has correct ingredient in correct amount and will induce the intended therapeutic effect.

Herbal drugs are very different from synthetic drugs, and availability of raw materials is often a problem. Standardization, stability, and quality control for herbal drugs are feasible but not easy and require various issues in consideration. This chapter has reviewed the problems associated with the preparation of herbal drugs and the guidelines recommended for the above-mentioned three steps to use herbal drugs at par with the modern synthetic drugs.

2 QUALITY CONTROL

It is important to understand concept of quality when dealing with quality control. According to ISO 9000, quality is a "degree to which a set of inherent characteristics fulfills requirements." Quality control refers to the processes involved directly or indirectly in maintaining the quality and validity of manufactured products. It is of paramount importance for safety, effectiveness, and acceptability of the product, and it is an essential operation of the pharmaceutical industry. Drugs must be marketed as safe and with the therapeutically active formulations exhibiting consistent and predictable performance.

2.1 Need of Quality Control in Regulation of Herbal Drugs

Although herbal products have become increasingly popular throughout the world, one of the impediments in its acceptance is the lack of standard quality control profile. The quality of herbal medicine, that is, the profile of the constituents in the final product has implications in efficacy and safety. However, owing to the complex nature and inherent variability of the constituents of plant-based drugs, it is difficult to establish quality control parameters, although modern analytical techniques are expected to help in circumventing this problem (Bandaranayake, 2006). In general, quality control is based on three important pharmacopeial definitions.

1. Identity – it should have one herb.
2. Purity – it should not have any contaminant other than herb.
3. Content or assay – the active constituents should be within the defined limits.

The methods of harvesting, drying, storage, transportation, and processing (e.g., mode of extraction and polarity of the extracting solvent and instability of constituents) also affect herbal quality.

The chemistry of plants involves the presence of therapeutically important constituents usually associated with many inert substances such as coloring agents, cellulose, and lignin. It is obvious that the content is the most difficult one to assess, as in most herbal drugs, the active constituents are unknown. Sometimes, markers can be used, which are, by definition, chemically defined constituents that are of interest for control purposes, independent of whether they have any therapeutic activity or not (WHO, 1992; Bandaranayake, 2006). The active principles are extracted from the plants and purified for therapeutic utility for their selective pharmacological activity. Therefore, quality control of herbal crude drugs and their constituents is of great importance in the modern system of medicine. Lack of proper standard parameters for the standardization of herbal preparation and several instances of poor quality herbs, adulterated herbs come into existence. To meet new thrust of inquisitiveness, standardization of herbals is mandatory (Raina, 2003; Kokate, Purohit and Gokhale, 2006).

3 STANDARDIZATION OF DRUG

It involves adjustment of drug preparation to a defined content of a constituent or a group of substances with known therapeutic activity by adding excipients or by mixing herbal drugs or herbal drug preparations. Standardization describes all measures taken during the manufacturing process and quality control, leading to reproducible quality of a particular product that ensures a predefined amount of quantity, quality, and therapeutic effect of ingredients in each dose (WHO, 1988; Zafar, Panwar, and Sagar Bhanu, 2005; Madhav *et al.*, 2011). It is the process of developing and agreeing on technical standards. Specific standards are worked out by experimentation and observations, which would lead to the process of prescribing a set of characteristics exhibited by the particular herbal medicine. Hence, standardization is a tool in the quality control process (Kunle *et al.*, 2012).

3.1 Standardization of Herbal Drugs

Herbal product cannot be considered scientifically valid if the drug tested has not been authenticated and characterized in order to ensure reproducibility in the manufacturing of the product. Moreover, many dangerous and lethal side effects have recently been reported, including direct toxic effects, allergic reactions, effects from contaminants, and interactions with herbal drugs (Zafar, Panwar, and Sagar Bhanu, 2005; Vaidya and Devasagayam, 2007). Growing need for standardization and quality control of herbal medicines is recognized by the World Health Organization (WHO). Standardization of botanicals offers many obstacles, and there are several challenges such as controversial identity of various plants and deliberated adulteration of plant material. Several additional problems are also inherent for standardization of herbal drugs as follows.

1. Herbal drugs are usually mixtures of many constituents.
2. The active principle(s) is (are), in most cases, unknown.
3. Selective analytical methods or reference compounds may not be available commercially.
4. Plant materials are chemically and naturally variable.
5. Chemo-varieties and chemo-cultivars exist.
6. The source and quality of the raw material are variable.

Standardization of raw materials used requires that each and every step, involved in the preparation of an herbal formulation, has to be authenticated, such as area of the collection, parts of the plant to be collected, the regional situation, phytomorphology, botanical identity, and microscopic and histological analyses.

Other parameters required for standardization of herbal drugs include quantitative microscopy, taxonomic identity, foreign matter chromatographic and spectroscopic evaluations, determination of heavy metals, pesticide residues, microbial contamination, and radioactive contamination.

Therapeutic activity of an herbal formulation depends on its phytochemical constituents (Choudhary and Sekhon, 2011). The development of authentic analytical methods, which can reliably profile the phytochemical composition, including quantitative analyses of marker/bioactive compounds and other major constituents, is a major challenge to scientists. In view of the above, standardization is an important step for the establishment of a consistent biological activity, a consistent
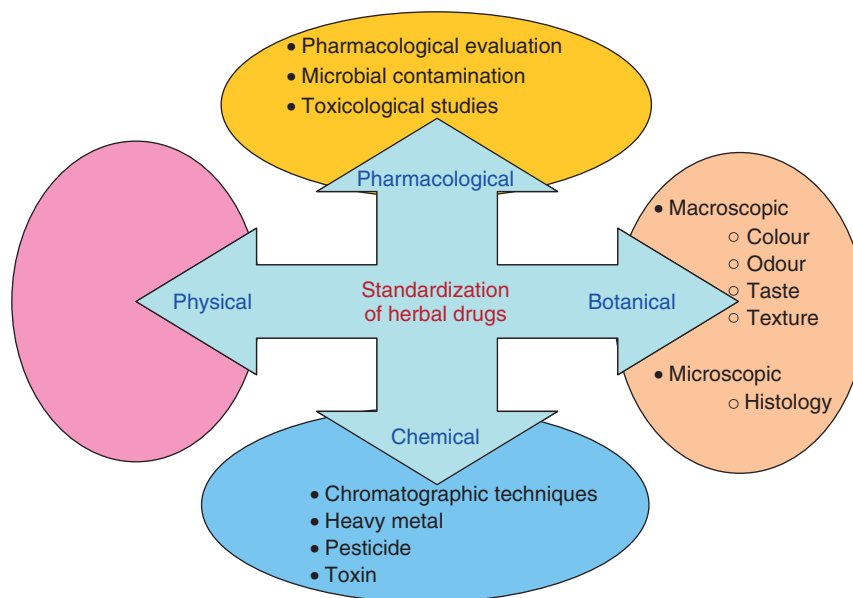
Figure 1   Steps involved in standardization of herbal drugs.

chemical profile, or simply a quality assurance program for production and manufacturing of an herbal drug (Patra *et al*., 2010). The authentication of herbal drugs and identification of adulterants from genuine medicinal herbs are essential for both pharmaceutical companies and public health, as well as to ensure reproducible quality of herbal medicine (Strauss, 2002). A schematic representation of herbal drug standardization is shown in Figure 1.

4   NEED OF QUALITY ASSURANCE IN REGULATION OF HERBAL DRUGS

Quality assurance is defined as the fulfillment of all the requirements, including legal and experience based, which are connected with all aspects of manufacturing of high quality herbal medicinal products. It starts from the beginning, the specification, the processing, and the procurement of herbal starting material, follows the procedures and quality considerations surrounding the intermediates and ends with devising and monitoring the final production steps toward the final medicinal product (WHO, 1996; Kunle *et al*., 2012).

Quality assurance is therefore defined as a network. It encompasses the control and documentation mechanisms, which insure that the multitude of regulations pertaining to and used in practice of the pharmaceutical industry is adhered to. Few of these relevant guidelines are good harvesting practice (GHP), good agricultural practice (GAP), good manufacturing practice (GMP), good laboratory practice (GLP), and good storage practice (GSP) (WHO, 1996; Bauer, 1998; Mathe and Mathe, 2008).

Quality assurance is the responsibility of not only the manufacturers but also the regulatory bodies. The assurance of the safety of an herbal drug requires monitoring of the quality of the consumer information on the herbal remedy. The quality of pharmaceuticals has been a concern of the WHO since its inception.

5   PARAMETERS FOR QUALITY CONTROL OF HERBAL DRUGS

Drug evaluation is performed to maintain the quality control of a drug. It means confirmation of its identity and determination of quality and purity of the herbal drug. Evaluation of crude drug is necessary

because of three main reasons: biochemical variations in the drug; deterioration due to treatment and storage; substitution and adulteration as a result of carelessness; and ignorance or fraud or variability caused by differences in growth, geographical location, and time of harvesting. For the quality control of a traditional medicine, the traditional methods are procured and studied, and documents and the traditional information about the identity and quality assessment are interpreted in terms of modern assessment or monograph in herbal pharmacopeia (Kokate, Purohit and Gokhale, 2006; Gupta and Sharma, 2007; Ansari, 2011).

One of the major problems faced by the herbal industry is the unavailability of rigid quality control profiles for herbal materials and their formulations. In India, the Department of Ayush, Government of India, launched a central scheme to develop standard operating procedures for the manufacturing process to develop pharmacopeial standards for ayurvedic preparations.

One of the impediments in the acceptance of the herbal products worldwide is the lack of standard quality control profiles (WHO, 1992; Shinde *et al*., 2009). Furthermore, the constituents responsible for the claimed therapeutic effects are frequently unknown or only partly explained. This is further complicated by the use of combination of herbal ingredients as being used in traditional practice. Most of the herbal formulations, especially the classical formulations of traditional medicine, are polyherbal. It is common to have as many as five different herbal ingredients in one product. Each formulation contains 10–20 or more ingredients; a few have even 50–75 ingredients. Many preparations are either liquid or semisolid (Shinde *et al*., 2009). For such formulations, it is very difficult to establish parameters for quality control. Even official standards are not available. Thus, batch-to-batch variation starts from the collection of raw material itself in the absence of any reference standard for identification. These variations multiply during storage and further processing. The unique processing methods followed for the manufacture of these drugs turn the single drugs into very complex mixture, from which separation, identification, and analysis of the components are very difficult. Hence, for herbal drugs and products, standardization should encompass the entire field of study from cultivation of medicinal plant to its clinical application (Kunle *et al*., 2012). The major parameters that are to be considered while considering quality control of herbal drugs are given in the following sections.

5.1 Identity of Plant Material

Authenticity, purity, and assay are important aspects of the standardization and quality control. As the name implies, authenticity relates to proving the material is true and corresponds to right identity. According to *WHO General Guidelines for Methodologies on Research and Evaluation of Traditional Medicines*, first step in assuring quality, safety, and efficacy of traditional medicines is correct identification. Chemical analysis is so far the best method for standardization, detection of contamination, plant identification, and authentication of medicinal plants. Molecular biology techniques can also be applied to authentication of medicinal plants as complimentary techniques (Donald and Arthur, 2002; Shinde and Dhalwal, 2007).

5.2 Botanical Parameters

5.2.1 Variations in Botanicals

Consistency in composition and biological activity are essential requirements for the safe and effective use of therapeutic agents. However, botanical preparations rarely meet this standard, as a result of problems in identifying plants, genetic variability, variable growing conditions, differences in harvesting procedures and processing of extracts, and, above all, the lack of information about active pharmacologic principles (Donald and Arthur, 2002).

The therapeutic or toxic components of plant vary depending on the part of the plant used as well as the stages of ripeness (Anna and Stephen, 1997). Products from different manufacturers vary considerably and it is not possible to control all the factors that affect the plant's chemical composition (Michael, 1999). As botanicals are prone to contamination and deterioration, there may be batch-to-batch variation in composition. Each herb contains a large number of diverse compounds and it is not possible to analyze for presence or absence for all compounds.

5.2.2 Organoleptic Evaluation or Morphological Evaluation

This type of evaluation of drug is procured by the organs of sense such as the skin, the eye, the tongue, the nose, and the ear. In addition, the morphological evaluation includes macroscopic evaluation of drugs by color, odor, taste, size, shape, texture, and so on. These are the techniques of qualitative evaluation based on the study of morphological and sensory profiles of whole drugs (WHO, 1998; Ahirwal, Ahirwal, and Ram, 2006; Surekha *et al*., 2011). For example, the fractured surfaces in cinchona, quillia, and cascara barks and quassia wood are important characteristics. Aromatic odor of umbelliferous fruits and sweet taste of licorice are the examples of this type of evaluation where odor of drug depends on the type and quality of odorous principles (volatile oils) present. The important diagnostic characters are as follows: shape of drug may be cylindrical (sarsapilla), subcylindrical (podophyllum), conical (aconite), fusiform (jalap), and so on; size represents length, breadth, thickness, diameter, and so on; and color means external color that varies from white to brownish black (Ahirwal, Ahirwal, and Ram, 2006; Kamboj, 2012).

5.2.3 Microscopic Evaluation

It involves detailed examination of the drug, and it can be used to identify the herbal drugs by their known histological characters. It is mostly used for qualitative evaluation of herbal preparation as a whole or in powder forms with the help of microscope (WHO, 1998; Kokate, Purohit and Gokhale, 2006; Ansari, 2011). The features under study include various cellular tissues, trichomes, stomata, starch granules, calcium oxalate crystals, and aleurone grains. The crude drugs can also be identified microscopically by cutting the thin transverse section (TS) and longitudinal section (LS), especially in case of wood and by staining them with proper staining reagents; for example, starch and hemicelluloses are identified by blue color with iodine solution and all lignified tissues give pink stain with phloroglucinol and hydrochloric acid. Mucilage is stained pink with ruthenium red, which can be used to distinguish cellular structure (Kamboj, 2012).

5.3 Physicochemical Parameters

5.3.1 Physical Evaluation

Physical constants are sometimes considered to evaluate certain drugs. These include moisture content, specific gravity, optical rotation, refractivity, melting point, viscosity, and solubility in different solvents. All these physical properties are useful in identification and detection of constituents present in plants. In addition, they also include foreign matter, total ash, acid-insoluble ash, water soluble ash, swelling and foaming indexes, successive extractive values, moisture content, viscosity, pH, disintegration time, friability, hardness, flow capacity, flocculation, sedimentation, and alcohol content (Bele and Khale, 2011). To determine ash content, the plant material is burnt and the residual ash is measured as total ash and acid-insoluble ash. Total ash is the measure of the total amount of material left after burning and includes ash derived from the part of the plant itself and acid-insoluble ash. The latter is the residue obtained after boiling the total ash with dilute hydrochloric acid, and burning the remaining insoluble matter. The second procedure measures the amount of silica present, especially in the form of sand and siliceous earth. Determination of moisture content helps in the estimation of volatile content present in the material (Kunle *et al*., 2012).

5.3.2 Chemical Evaluation

Most of the herbal drugs have definite chemical constituents to which their biological or pharmacological activity is attributed. Qualitative chemical tests are used to identify certain drugs or test their purity. The isolation, purification, and identification of active constituents are based on chemical methods of evaluation. Among various qualitative chemical tests such as estimation of acid value and saponification value, some are useful in evaluation of resins (acid value and sulfated ash), balsams (acid value, saponification value, and better values), volatile oils (acetyl and ester values), and gums (methoxy determination and volatile acidity). Preliminary phytochemical screening is a part of chemical evaluation. These qualitative chemical tests are useful in identification of chemical

constituents and detection of adulteration (Kamboj, 2012).

5.4 Determination of Foreign Matter

The raw material being used in the preparation of herbal drugs should be pure and free from foreign materials. These exogenous materials could be consisting of parts of medicinal plant materials or materials other than those named with the limits specified for the plant material concerned, any organism, part or product of it, or mineral admixtures not adhering to the medicinal plant materials, such as soils, stones, sand, and dust. Plant materials should be free from any excreta, molds, insects, and chemical residue (EMEA, 2002; WHO, 2003; Surekha *et al.*, 2011). Purity is closely linked with safe use of drugs and deals with factors such as ash values, contaminants (e.g., foreign matter in the form of other herbs), and heavy metals. However, owing to the application of improved analytical methods, modern purity evaluation also includes microbial contamination, aflatoxins, radioactivity, and pesticide residues. Analytical methods such as photometric analysis, thin layer chromatography (TLC), high performance liquid chromatography (HPLC), high performance thin layer chromatography (HPTLC), and gas chromatography (GC) can be employed in order to establish the constant composition of herbal preparations (Mukharji, 2001; Kokate, Purohit, and Gokhale, 2006; Ali, 2009).

5.4.1 Determination of Heavy Metals

Heavy metals can be either accidentally or intentionally present in drugs. Contamination by heavy metals such as mercury, lead, copper, cadmium, and arsenic in herbal drugs can be attributed to many causes including environmental pollution. Presence of such materials may lead to severe health hazards. A simple method for detection of heavy metals is based on color reactions with special reagents such as thioacetamide or diethylthiocarbamate.

Instrumentation analysis is required when metals are present in trace amounts such as atomic absorption spectrophotometry (AAS), inductively coupled plasma (ICP), and neutron activation analysis (NAA) (Surekha *et al.*, 2011; Kamboj, 2012).

5.4.2 Determination of Pesticide

Even though there are no serious reports of toxicity due to the presence of pesticides and fumigants, it is important that herbs and herbal products are free of these chemicals or at least are controlled for the absence of unsafe levels. Pesticide residues produce toxic effects such as irritation of the eye, lacrimation, salivation, blurring of vision, sweating, breathlessness, cardiac arrhythmias, respiratory paralysis, hypotension, and other side effects (EMEA, 1998; WHO, 1998; Surekha *et al.*, 2011). They are any specified substance in food, agricultural commodities, or animal feed resulting from the use of a pesticide. The term includes any derivatives of a pesticide, such as conversion products, metabolites, reaction products, and impurities, considered to be of toxicological significance (EMEA, 1998; WHO, 1990, 1998).

Herbal drugs are liable to contain pesticide residues, which accumulate from the agricultural practices such as spraying, treatment of soils during cultivation, and administering of fumigants during storage. The WHO and the Food and Agricultural Organization (FAO) set limits of pesticides, which are usually present in the herbs. Mainly, pesticides such as dichlorodiphenyltrichloroethane (DDT), benzene hexachloride (BHC), toxaphene, and aldrin cause serious side effects in human beings if the crude drugs are mixed with these agents (WHO, 1990; Shrikumer *et al.*, 2006; WHO, 2007). However, it may be desirable to test herbal drugs for broad groups in general, rather than for individual pesticides. Many pesticides contain chlorine in the molecule, which, for example, can be measured by analysis of total organic chlorine. In an analogous way, insecticides containing phosphate can be detected by measuring total organic phosphorus. Samples of herbal material are extracted by a standard procedure, impurities are removed by partition and/or adsorption, and individual pesticides are measured by GC, MS, or GC–MS. Some simple procedures have been published by the WHO and the *European Pharmacopoeia* has laid down general limits for pesticide residues in medicine (EMEA, 1998, WHO, 1998, 2004a, 2007; Ali, 2009).

5.4.3 Determination of Microbial Contaminants

Herbs and herbal materials may be associated with a variety of microbial contaminants, such as bacteria, fungi, and viruses often originating in soil or derived from manure. While a large range of bacteria and fungi form the naturally occurring microflora of medicinal plants, aerobic spore-forming bacteria frequently predominate. Proliferation of microorganisms may result from failure to control the moisture levels of herbal medicines during transportation and storage, as well as from failure to control the temperatures of liquid forms and finished herbal products (Shrikumer *et al*., 2006).

This microbiological background depends on several environmental factors and exerts an important impact on the overall quality of herbal products and preparations (Bele and Khale, 2011; Kunle *et al*., 2012). The presence of *Escherichia coli*, *Salmonella* spp., and molds may indicate poor quality of production and harvesting practices. Microbial contamination may also occur through handling by personnel who are infected with pathogenic bacteria during harvest/collection, post-harvest processing, and the manufacturing process. This should be controlled by implementing best practice guidelines such as good agricultural and collection practices (GACP) and GMP (WHO, 2002, 2004a,b, 2006).

Laboratory procedures investigating microbial contaminations are laid down in the well-known pharmacopeias, as well as in the WHO guidelines. Limit values can also be found in the sources mentioned. In general, a complete procedure consists of determining the total aerobic microbial count, the total fungal count, and the total Enterobacteriaceae count, together with tests for the presence of *E. coli*, *Staphylococcus aureus*, *Pseudomonas aeruginosa*, *Shigella* spp, and *Salmonella* spp. The *European Pharmacopoeia* also specifies that *E. coli* and *Salmonella* spp. should be absent from herbal preparations. The allowed contamination level may also depend on the method of processing of the drug. For example, higher contamination levels are permitted if the final herbal preparation involves boiling with water (Bandaranayake, 2006; Kunle *et al*., 2012).

5.4.4 Determination of Toxins

In addition to microbial load in the herbal preparation, the presence of toxins is of serious concern. Endotoxins are found mainly in the outer membranes of certain Gram-negative bacteria and are released only when the cells are disrupted or destroyed. They are complex lipopolysaccharide molecules that elicit an antigenic response, cause altered resistance to bacterial infections, and have other serious effects. Thus, tests for their presence on herbal medicines should be performed in dosage forms for parenteral use, in compliance with the requirements of national, regional, or international pharmacopeias.

The presence of mycotoxins in plant material can pose both acute and chronic risks to health. Mycotoxins are usually secondary metabolic products that are nonvolatile, have a relatively low molecular weight, and may be secreted onto or into the medicinal plant material. The mycotoxins produced by species of fungi including *Aspergillus*, *Fusarium*, and *Penicillium* are the most commonly reported. They comprise four main groups, namely, aflatoxins, ochratoxins, fumonisins, and tricothecenes, all of which have toxic effects. Aflatoxins have been extensively studied and are classified as Group 1 human carcinogens by the International Agency for Research on Cancer (IARC, 1993). Aflatoxins in herbal drugs can be dangerous to health even if they are absorbed in minute amounts (FAO, 2003: Bandaranayake, 2006). Aflatoxin-producing fungi sometimes build up during storage. Procedures for the determination of aflatoxin contamination in herbal drugs are published by the WHO. After a thorough clean-up procedure, TLC is used for confirmation. The other microbial toxins such as bacterial endotoxins and mycotoxins also play a detrimental role in the final preparation of herbal drugs. Withering of plant materials in the storage stage leads to increased enzymic activity, transforming some of the constituents to other metabolites not initially found in the herb. These newly formed compounds along with the molds may have adverse effects (WHO, 2007; Surekha *et al*., 2011).

5.4.5 Toxicological Parameters

Toxicity investigation will also be required because the analysis alone is unlikely to reveal the contributions to toxicity itself. In assessing toxicity of an

herbal medicine, the dose chosen is very important (TDR, 2005; Balammal, Sekar Babu, and Reddy, 2012). Toxicity assessment involves one or more of the following techniques – *in vivo* techniques, *in vitro* techniques, cell line techniques, microarray, and other modern standardization techniques to adequately model toxicity. Evaluation of the toxic effects of plant constituents of herbal formulation requires detailed phytochemical and pharmacological studies. It is, however, safe to assume that, based on human experiences in various cultures, the use of toxic plant ingredients has already been largely eliminated, and recent reports of toxicity could largely be due to misidentification and overdosing of certain constituents (ICDRA, 1991). Substitution and misidentification of herbal substances, documented or regulatory approaches, development of monitoring and surveillance systems, assessment of toxicity, and risk assessment approaches are employed to assess toxicity in herbal drugs.

5.5 Efficacy of Herbal Medicine

The efficacy of medicine is the measure of its ability to improve health and well-being. The use of herbal drugs is often justified by their long history of usage since ancient times. However, age-old wisdom does not necessarily guarantee that the product in question is efficacious with reasonable specificity. The term efficacious has a relative meaning as it may be interpreted differently by the practitioners of traditional medicine (of which herbal medicine is a type) and the proponents of the so-called modern medicine (conventional medicine). Traditional medicine usually takes a holistic approach where the physical, spiritual (which includes mental), and most often social well-being of an individual are treated (Payyappallimana, 2009). Thus, the medicinal value of an herbal product may be intimately related to its nutritional and psychological aspects. Although this philosophy is not consciously followed by many traditional practitioners nowadays, it remains a natural corollary of their system that the efficacy of their intervention cannot simply be judged by physiological and biochemical indicators. Modern medicines, on the other hand, are relatively more focused on particular diseases based on specific etiopathological entities and usually emphasize only physical well-being. In recent times,

the trend is to develop more of an integrated approach of treatment involving physical, mental, and social well-being (Barnes, 2003).

5.5.1 Assessment of Efficacy

Herbal medicines are inherently different from conventional pharmacological treatments; however, presently, there is no way to assess their efficacy other than by currently used conventional clinical trial methodologies, in which the efficacy is conventionally assessed by clinical, laboratory, or diagnostic outcomes. Clinical outcomes include the parameters such as improved morbidity, reduced pain or discomfort, improved appetite and weight gain, reduction of blood pressure, reduction of tumor size or extent, and improved quality of life (Rasheed, Reddy, and Roja, 2012). Laboratory or other diagnostic outcomes include parameters such as reduction of blood glucose, improvement of hemoglobin status, reduction of opacity as measured by radiological or imaging techniques, and improvement in electrocardiogram (ECG) findings (Choudhary and Sekhon, 2011; Rasheed, Reddy, and Roja, 2012).

5.6 Pharmacological Parameters

For herbal products, analysis of the active pharmaceutical ingredients may be best approached by analysis of one or more hypothesized active ingredients, analysis of a chemical constituent that constitutes a sizable percentage of total ingredients, and chemical fingerprinting of ingredients. Some drugs have specific biological and pharmacological activities that are utilized for their evaluation. Actually, this activity is due to specific type of constituents present in the plant extract. For evaluation, the experiments were carried out on both intact and isolated organs of living animals. With the help of bioassays (testing the drugs on living animals), strength of the drug in its preparation can also be evaluated (Kokate, Purohit and Gokhale, 2006; Ansari, 2011). Some important biological evaluations are pharmacognostic, antimicrobial, antifertility, hypoglycemic, and neuropharmacological problems.

6 ANALYTICAL METHODS FOR EVALUATION OF HERBAL DRUG USING MODERN TECHNIQUES

In general, the methods for quality control of herbal medicines involve analytical inspection using chromatographic and spectroscopic techniques. They include TLC, HPLC, GC, ultraviolet (UV), infrared (IR), Fourier transform infrared (FT-IR), atomic absorption spectroscopy (AAS), fluorimetry, near infrared (NIR), and spectrophotometer. HPLC fingerprinting includes recording of the chromatograms, retention time of individual peaks, and the absorption spectra (recorded with a photodiode array detector) with different mobile phases. Similarly, gas–liquid chromatography (GLC) is used for generating the fingerprint profiles of volatile oils and fixed oils of herbal drugs. Furthermore, the recent approaches of applying hyphenated chromatography and spectrometry such as high performance liquid chromatography–diode array detection (HPLC–DAD), gas chromatography–mass spectrometry (GC–MS), capillary electrophoresis-diode array detection (CEDAD), high performance liquid chromatography–thin layer chromatography (HPTLC), high performance liquid chromatography–mass spectroscopy (HPLC–MS), and high performance liquid chromatography–nuclear magnetic resonance spectroscopy (HPLC–NMR) could provide the additional spectral information, which will be very helpful for the qualitative analysis and even for the on-line structural elucidation (Ong, 2002; Liang, Xie, and Chan, 2004; Nikam *et al.*, 2012).

On the other hand, the methods of extraction and sample preparation are also of great importance in preparing good fingerprints of herbal medicines. A single herbal medicine may contain many natural constituents, and a combination of several herbs might give rise to interactions with hundreds of natural constituents during the preparation of extracts. Therefore, the fingerprints produced by the chromatographic instruments, which may present a relatively good integral representation of various chemical components of herbal medicines, are of high importance (Bilia *et al.*, 2002; Choi *et al.*, 2002; Kamboj, 2012).

6.1 Chromatographic Fingerprinting and Marker Compound Analysis

A chromatographic fingerprint of an herbal medicine is a chromatographic pattern of the extract of some common chemical components of pharmacologically active and/or chemical characteristics. It is suggested that with the help of chromatographic fingerprints obtained, the authentication and identification of herbal medicines can be accurately conducted even if the amount of the chemically characteristic constituents are not similar in different samples of the same preparation. A good quality control requires consideration of multiple constituents in the herbal drugs, and not individually considers only one or two marker components for evaluating the quality of the products (Zhang, 2004). However, in any of the herbal product and its extract, there are hundreds of unknown components and many of them are in low amount. Moreover, there usually exists variability within the same herbal materials. Hence, it is very important to obtain reliable chromatographic fingerprints that represent pharmacologically active and chemically characteristic components of the herbal medicine (Bele and Khale, 2012; Kunle *et al.*, 2011).

6.1.1 Thin Layer Chromatography

Thin layer chromatography is simply known as *TLC*. It is one of the most popular and simple chromatographic techniques used for the separation of compounds. In the phytochemical evaluation of herbal drugs, TLC is being employed extensively because (i) it enables rapid analysis of herbal extracts with minimum sample clean-up requirement, (ii) provides qualitative and semi quantitative information of the resolved compounds, and (iii) enables the quantification of chemical constituents.

In TLC fingerprinting, the data that can be recorded using a HPTLC scanner includes the chromatogram, retardation factor (Rf) values, the color of the separated bands, their absorption spectra (λ max), and shoulder inflections of all the resolved bands. All of these, together with the profiles on derivatization with different reagents, represent the TLC fingerprint profile of the sample. The information so generated has a potential application in the identification of an authentic drug, in excluding the adulterants and in maintaining the quality and consistency of the drug (Nikam *et al.*, 2012).

6.1.2   High Performance Thin Layer Chromatography

HPTLC technique is widely employed in pharmaceutical industry in process development, identification, and detection of adulterants in herbal products, helps in identification of pesticide content and mycotoxins, and in quality control of herbs and health foods (Soni and Naved, 2010). It has been well reported that several samples can be run simultaneously using a smaller quantity of mobile phase than in HPLC (Jianga *et al.*, 2010). It has also been reported that mobile phases of pH 8 and above can be used for HPTLC. Another advantage of HPTLC is the repeated detection (scanning) of the chromatogram with the same or different conditions. Consequently, HPTLC has been investigated for simultaneous assay of several components in a multicomponent formulation (Thoppil, Cardoza, and Amin, 2001). The simultaneous estimation of withaferin A and beta-sinosterol-D-glucoside in four Ashwagandha formulations (Shanbhag and Khandagale, 2011). *Syzygium jambolanum* DC. (Myrtaceae) was quantitatively estimated in terms of stability, repeatability, accuracy, and phytoconstituents such as glycoside, tannins, ellagic acid, and gallic acid by HPTLC (Kshirsagar *et al.*, 2008).

With this technique, authentication of various species of plant is possible, as well as the evaluation of stability and consistency of their preparations from different manufacturers. Various workers have developed HPTLC method for phytoconstituents in crude drugs or herbal formulations such as bergenin, catechin, and gallic acid in *Bergenia cilliata* Moench (Saxifragaceae) and *Bergenia lingulata* Wall. Engl. (Saxifragaceae) (Dhalwal *et al.*, 2008). HPTLC technique was reported for simultaneous estimation of gallic acid, rutin, and quercetin in *Terminalia chebula* Retz. (Combretaceae).

6.1.3   High Performance Liquid Chromatography

Preparative and analytical HPLCs are widely used in pharmaceutical industry for the isolation and purification of herbal compounds. There are basically two types of preparative HPLC: low pressure HPLC (typically under 5 bar) and high pressure HPLC (pressure >20 bar) (Chimezie *et al.*, 2008). The important parameters to be considered are resolution, sensitivity, and fast analysis time in analytical HPLC, whereas both the degree of solute purity as well as the amount of compound that can be produced per unit time, that is, throughput or recovery in preparative HPLC (Rao and Anna, 2009). This is very important in pharmaceutical industry of today because new products (natural and synthetic) have to be introduced to the market as quickly as possible. The availability of such a powerful purification technique makes it possible to spend less time on the synthesis conditions (Bhutani, 2003; Marston, 2002).

The combination of HPLC and liquid chromatography–mass spectrometry LC–MS is presently powerful technique for the quality control of Chinese medicine, that is, licorice (Manisha *et al.*, 2011). Kankasava is a fermented polyherbal formulation prepared with Kanaka and other ingredients (Vogel *et al.*, 2005). It is used in chronic bronchitis, asthmatic cough, and breathlessness. Kankasava is analyzed by reversed phase high performance liquid chromatography (RP-HPLC). It is a simple, precise, accurate RP-HPLC method developed for the quantitative estimation of atropine in Kankasava polyherbal branded formulations.

6.1.4   Liquid Chromatography–Mass Spectroscopy (LC–MS)

LC–MS is an analytical technique that combines the physical separation capabilities of liquid chromatography (LC) with the mass analysis capabilities of MS. It is a powerful technique used in many applications, which has very high sensitivity and selectivity. Generally, its application is oriented toward the specific detection and potential identification of chemicals in a complex mixture. It has become the method of choice in many stages of drug development (Shankar and Singh, 2011). Recent advances include electrospray, thermospray, and ionspray ionization techniques, which offer unique advantages of high detection sensitivity and specificity and liquid secondary ion mass spectroscopy; later, laser mass spectroscopy with 600 MHz offers accurate determination of molecular weight proteins and peptides. Isotopes pattern can be detected by this technique (Bhutani, 2003).

6.1.5 Liquid Chromatography–Nuclear Magnetic Resonance (LC–NMR)

The combination of LC and nuclear magnetic resonance (NMR) offers the potential of unparalleled chemical information from analytes separated from complex mixtures. LC–NMR improves speed and sensitivity of detection and found useful in the areas of pharmacokinetics (PK), toxicity studies, drug metabolism, and drug discovery process. The combination of chromatographic separation technique with NMR spectroscopy is one of the most powerful and time-saving method for the separation and structural elucidation of unknown compounds and mixtures, especially for the structure elucidation of light and oxygen-sensitive substances. The on-line LC–NMR technique allows the continuous registration of time changes as they appear in the chromatographic-run automated data acquisition, and processing in LC–NMR improves speed and sensitivity of detection. The recent introduction of pulsed field gradient technique in high resolution NMR and three-dimensional technique improves application in structure elucidation and molecular weight information. These new hyphenated techniques are useful in the areas of PK, toxicity studies, drug metabolism, and drug discovery process (Patil and Rajani, 2010; Nikam et al., 2012).

6.1.6 Liquid Chromatography–Infrared Spectroscopy (LC–IR)

The hyphenated technique developed from the coupling of liquid chromatography and infrared spectroscopy is known as *liquid chromatography-infrared spectroscopy* (*LC-IR*). LC-IR is an important technique as it shows absorption peaks of functional groups in mid-IR region that helps in structural identification of compounds present in a sample. The detection technique of IR is comparatively slower than other techniques such as MS or NMR. Two approaches used in these techniques are flow cell approach and solvent elimination approach.

6.1.7 Gas Chromatography–Mass Spectrometry (GC–MS)

GC equipment can be directly interfaced with rapid scan mass spectrometer of various types. GC and GC–MS are unanimously accepted methods for the analysis of volatile constituents of herbal medicines, owing to their sensitivity, stability, and high efficiency. Especially, the hyphenation with MS provides reliable information for the qualitative analysis of the complex constituents (Guo et al., 2006; Teo et al., 2008). With GC-MS, not only a chromatographic fingerprint of the essential oil of the herbal medicine can be obtained, but also the information related to its most qualitative and relative quantitative compositions is obtained. Used in the analysis of the herbal medicines, there are at least two significant advantages for GC–MS, that is, (i) with the capillary column, GC–MS has in general very good separation ability, which can produce a chemical fingerprint of high quality and (ii) with the coupled mass spectroscopy and the corresponding mass spectral database, the qualitative and relatively quantitative composition information of the herb investigated could be provided by GC–MS, which will be extremely useful for further research for elucidating the relationship between chemical constituents in herbal medicine and its pharmacology in further research. Thus, GC–MS should be the most preferable tool for the analysis of the volatile chemical compounds in herbal medicines (Gong et al., 2003; Li et al., 2003; Kamboj, 2012).

GC–MS instruments have been used for identification of hundreds of components that are present in natural and biological systems (Sharma, Gaurav, and Balkrishna, 2009; Nikam et al., 2012). An effective fast and accurate capillary GC method was employed for determining the organochlorine pesticide residues. It has been used for identification of a large number of components present in natural and biological systems (Lee and Edward, 1999; Rasheed, Reddy, and Roja, 2012).

6.1.8 Gas Chromatography-Flame Ionization Detector (GC-FID)

A number of detectors are used in gas chromatography. The most common are the flame ionization detector (FID) and the thermal conductivity detector (TCD). Coupling capillary column gas chromatographs with FT-IR spectrometer provides a potent means for separating and identifying the components of different mixtures (Sharma, Gaurav, and Balkrishna, 2009). Both are sensitive to a wide range

of components, and both work over a wide range of concentrations. While TCDs are essentially universal and can be used to detect any component other than the carrier gas (as long as their thermal conductivities are different from that of the carrier gas, at detector temperature), FIDs are sensitive primarily to hydrocarbons, and are more sensitive to them than TCD. However, an FID cannot detect water. Both detectors are also quite robust. As TCD is nondestructive, it can be operated in series before an FID (destructive), thus providing complementary detection of the same analytes (Nikam et al., 2012).

6.1.9   Supercritical Fluid Chromatography (SFC)

Supercritical fluid chromatography (SFC) is a hybrid of gas and liquid chromatographies that combines some of the best features of each. It permits the separation and determination of a group of compounds that are not conveniently handled by either GC or LC. The supercritical fluid and microbore liquid chromatography offer potential applications for the drug analysis. In this, the mobile phase is a gas (CO_2) maintained at its supercritical state, that is, above its critical temperature and pressure. The SFC mobile phase has low viscosity, approximating that of a gas, and high diffusivity between those of capillary GC and LC (Nikam et al., 2012).

SFC has been applied to a wide variety of materials including natural products, drugs, food, and pesticide. (Matthew and Henry, 2006; Nikam et al., 2012). These compounds are either nonvolatile or thermally labile so that GC procedures are inapplicable or contain no functional group that makes possible detection by the spectroscopic or electrochemical technique employed in LC (Patil and Rajani, 2010).

6.2   DNA Fingerprinting

Deoxyribonucleic acid (DNA) analysis has been proved as an important tool in herbal drug standardization. This technique is useful for the identification of phytochemically indistinguishable genuine drug from substituted or adulterated drug. It has been reported that DNA fingerprint genome remains the same irrespective of the plant part used, whereas the phytochemical content will vary with the plant part used, physiology, and environment (Shikha and Mishra, 2009). DNA is the fundamental building component of all living cells. Human characteristics, traits, and physical features are determined by the specific arrangement of DNA base-pair sequences in the cell. The distinct arrangement of adenine, guanine, thymine, and cytosine (called *DNA nucleotides*) regulates the production of specific proteins and enzymes via the *Central Dogma Theory*. Central Dogma Theory can be defined as the fundamental theory of molecular biology that genetic information flows from DNA to ribonucleic acid (RNA) to proteins. This concept of fingerprinting has been increasingly applied in recent decades to determine the ancestry of plants, animals, and other microorganisms. Genotypic characterization of plant species and strains is useful as most plants, although belonging to the same genus and species, may show considerable variation between strains. Additional motivation for using DNA fingerprinting on commercial herbal drugs is the availability of intact genomic DNA from plant samples after they are processed. Adulterants can be distinguished even in processed samples, enabling the authentication of the drug (Mihalov, Marderosian, and Pierce, 2000; Breithaupt, 2003). The other useful application of DNA fingerprinting is the availability of intact genomic DNA specificity in commercial herbal drugs, which helps in distinguishing adulterants even in processed samples (Lazarowych and Pekos, 1998; Shaw et al., 2009).

6.3   Genetic Marker

A genetic marker is a gene or DNA sequence with a known location on a chromosome and associated with a particular gene or trait. It can be described as a variation, which may arise because of mutation or alteration in the genomic loci that can be observed. It may be a short DNA sequence such as a sequence surrounding a single base-pair change (single nucleotide polymorphism SNP) or a long one such as minisatellites. Some commonly used types of genetic markers are restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP), random amplification of polymorphic DNA (RAPD), variable number tandem repeat (VNTR), micro satellite polymorphism, single nucleotide polymorphism (SNP), short tandem repeat (STR), and single feature polymorphism (SFP) (Shaw et al., 2009).

They can be further categorized as dominant or codominant. Dominant markers allow for analyzing many loci at once, that is, RAPD. A primer amplifying a dominant marker could amplify at many loci in one sample of DNA with one polymerase chain reaction (PCR) reaction. Codominant markers analyze one locus at a time. A primer amplifying a codominant marker would yield one targeted product (Shaw and But, 1995; Gomez *et al*., 2002).

6.3.1   Role of Genetic Marker in Herbal Drug Technology

6.3.1.1   Genetic Variation/Genotyping

It has been well documented that geographical conditions affect the active constituents of the medicinal plant and hence their activity profiles. Many researchers have studied geographical variation at the genetic level. Estimates of genetic diversity are also important in designing crop improvement programs for management of germplasm and evolving conservation strategies. RAPD-based molecular markers have been found to be useful in differentiating different accessions of neem collected from different geographical regions (Khanuja *et al*., 2002; Zhang *et al*., 2003). Germplasm analysis to study genetic diversity is another important area in which a lot of efforts have been put in. Fingerprinting of crops such as rice, wheat, chickpea, pigeon pea, and pearl millet is being carried out extensively (Ramakrishna *et al*., 1994; Khanuja *et al*., 2002; Sze *et al*., 2008).

6.3.1.2   Authentication and Detection of Adulteration of Medicinal Plants

DNA-based techniques have been widely used for authentication of plant species of medicinal importance. This is especially useful in case of those that are frequently substituted or adulterated with other species or varieties that are morphologically and/or phytochemically indistinguishable (Srivastava and Mishra, 2009). Dried fruit samples of goji or *Lycium barbarum* L. (Solanaceae) were differentiated from related species using RAPD markers. The RAPD technique has also been used for determining the components of a Chinese herbal prescription, yu-pingfeng san. In this study, the presence of three herbs (*Astragalus membanaceus* Fisch. Bge (Fabaceae), *Ledebouriella seseloides* Wolff (Apiaceae), and *Atractylodes macrocephala* Koidz

(Asteraceae)) in the formulation has been detected using a single RAPD primer (McCouch *et al*., 1988).

Sequence characterized amplified region (SCAR), arbitrarily primed polymerase chain reaction (AP-PCR), RAPD, and RFLP have been successfully applied for differentiation of these plants and to detect substitution by other closely related species. For example, *Panax ginseng* is often substituted by *Panax quinquefolius* (American ginseng) (Shaw and But, 1995; Lau *et al*., 2001).

6.3.1.3   Quality Control and Standardization of Medicinal Plant Materials

To ensure efficacy, selection of the correct chemo type of the plant is necessary even when there are many known chemotypes of a plant species, selection of the right chemo type to which clinical effects are attributed is difficult. DNA markers are reliable for informative polymorphism as the genetic composition is unique for each species and is not affected by age, physiological condition, and environmental factors. DNA can be extracted from fresh or dried organic tissue of the botanical material; hence, the physical form of the sample for assessment does not restrict detection. Various types of DNA-based molecular techniques are utilized to evaluate DNA polymorphism, which are hybridization-, PCR-, and sequencing-based methods (Cai *et al*., 1999; Sze *et al*., 2008).

6.4   Metabolomics Technique

DNA-based species identification alone is not sufficient for quality control of herbal medicines because plants are the products of both the genome and the environment. The metabolome is the end product of the biochemical chain DNA–RNA–protein–metabolite.   Metabolomics deals with the qualitative/quantitative identification of all the metabolites present in a living organism. The metabolites of an herbal medicine play a central role in mediating pharmacological effects. This technique has been used for identification of active phytoconstituents from herbal medicine (Shyur and Yang, 2008; He *et al*., 2011).

The use of metabolomics to screen for biomarker patterns and reveal biochemical processes during the post-genome era has increased contemporaneously with progress in global systems biology (Chan,

Pasikanti, and Nicholson, 2011). Metabolomics, which is the study of metabolite profiles in a biological system under a given set of conditions, has become an approach to understand the basic principles of relating chemical patterns in biology and systems biology (Nicholson and London, 2008). Metabolomics, with the capability of simultaneous analysis of hundreds and thousands of variables, meets the requirements for the evaluation of multicomponent herbal medicines *in vivo* and therefore bridges the gap between herbal medicine and molecular pharmacology (Wang *et al*., 2005). Utilizing a metabolomics platform to interpret the efficacies or toxicities of herbal medicines has been a key focus of recent herbal medicine researches (Wang *et al*., 2007; Zhao *et al*., 2008; Chen *et al*., 2008; Sun *et al*., 2009). However, an integrated approach involving metabolomics, pharmacology, and PK has not been explored and is greatly needed for herbal medicine researches.

Metabolomic approach was employed to identify the chemical constituents in *Sophora flavescens*, Aiton (Fabaceae), which were further analyzed for their effect on Pregane X receptor activation and Cytochrome P3A regulation (Wang *et al*., 2010). The greater potential of metabolomics has been reported in the development of active secondary metabolites from medicinal plants as novel or improved phytotherapeutic agents (Wang *et al*., 2010; He *et al*., 2011). The recent studies showed that NMR-based metabolomics approach combined with orthogonal projections to latent structure-discriminant analysis identified the purity of an herbal drug (Kang *et al*., 2008; Sun, Zhang, and Wang, 2012).

7 CONCLUSION

Herbal drug technology involves conversion of botanical materials into medicines where standardization and quality control with proper integration of modern scientific techniques and traditional knowledge are employed. The development of herbal medicines, as a rule, involves planning and obtaining intermediate preparations as required steps for processing vegetable raw material into a finished product, which will provide the desired pharmaceutical presentation. Herbal medicinal products may vary in composition and properties, and increasing reports of adverse reactions have drawn the attention

of many regulatory agencies for the standardization of herbal formulations. In this context, correct identification and quality assurance is an essential prerequisite to ensure reproducible quality of herbal drug or medicine, which contributes to its safety and efficacy. Standardization of herbal formulations is essential in order to assess the quality of drugs, based on the concentration of their active principles, physical, chemical, phytochemical standardizations, and *in vitro* and *in vivo* parameters. The quality assessment of herbal formulations is of paramount importance in order to justify their acceptability in the modern system of medicine.

8 RELATED ARTICLES

Selection, Identification, and Collection of Plants; Microscopic Analysis; Thin-layer Chromatography, with Chemical and Biological Detection Methods; HPLC and Ultra HPLC: Basic Concepts; LC and LC-MS: Techniques and Applications; HPLC Analysis of Alkaloids; Metabolomics; Mycotoxins Contamination in Food: Alternative Plant Preservatives, Legislation, and Detection Methods.

ACKNOWLEDGEMENT

REFERENCES

Ackerknecht, H. E. (1973) *Therapeutics: From the Primitives to the Twentieth Century*, Hafner Press, New York.

Ahirwal, B., Ahirwal, D., and Ram, A. (2006) *Souvenir, Recent Trends Herbal Ther.*, **11**, 25–29.

Ali, M. (2009) *Pharmacognosy and Phytochemistry* (3rd edn.). CBS Publishers & Distributors: Delhi, 181-182.

Anna, K. D., and Stephen, P. M. (1997) *Med. J. Aust.*, **166**, 538–541.

Ansari, S. H. (2011) *Essentials of Pharmacognosy*, Birla Publications Pvt Ltd, New Delhi, pp. 10–16.

Balammal, G., Sekar Babu, M., and Reddy, J. P. (2012) *Int. J. Preclin. Pharma. Res.*, **3**, 50–63.

Bandaranayake, W. M. (2006) Quality control, screening, toxicity, and regulation of herbal drugs In *Modern Phytomedicine. Turning Medicinal Plants into Drugs*, eds. Ahmad, I., Aqil, F., Owais, M., Wiley-VCH Verlag GmbH: Weinheim, Germany, 25-57.

Barnes, J. (2003) *Br. J. Clin. Pharmacol.*, **55**, 226–233.

Barnes, J., Anderson, L. A., Phillipson, J. D. (2007) *Herbal Medicine* (3rd edn.), Pharmaceutical Press: London, 1-23.

Bauer, R. (1998) *J. Drug Inform.*, **32**, 101–110.

Bele, A. A., and Khale, A. (2011) *Int. Res. J. Pharm.*, **2**, 56–60.

Bhutani, K. K. (2003) *Ind. J. Nat. Prod.*, **19**, 3–10.

Bilia, A. R., Bergonzi, M. C., Lazari, D., *et al*. (2002) *J. Agric. Food Chem.*, **50**, 5016.

Bodeker, C., Bodeker, G., Ong, C. K., *et al*. (2005) *WHO Global Atlas of Traditional, Complementary and Alternative Medicine*, World Health Organization, Geneva.

Breithaupt, H. (2003) *EMBO Rep.*, **4**, 10–12.

Busse, W. (2000) *Drug Inform. J.*, **34**, 15–23.

Cai, Z. H., Li, P., Dong, T. T., *et al*. (1999) *Planta Med.*, **65**, 360–364.

Calixto, J. B. (2000) *Braz. J. Med. Biol. Res.*, **33**, 179–189.

Chan, E. C., Pasikanti, K. K., and Nicholson, J. K. (2011) *Nat. Protoc.*, **6**, 1483–1499.

Chen, M., Ni, Y., Duan, H., *et al*. (2008) *Chem. Res. Toxicol.*, **21**, 288–294.

Chimezie, A., Ibukun, A., Teddy, E., *et al*. (2008) *Afr. J. Pharm. Pharmacol.*, **2**, 29–36.

Choi DW, Kim JH, Cho SY, Kim DH, Chang SY. 2002. Regulation and quality control of herbal drugs in Korea. *Toxicology* **181**/182: 581-586.

Choudhary, N., and Sekhon, B. S. (2011) *J. Pharmacog.*, **2**, 55–70.

Dhalwal, K., Sindhe, V. M., Biradar, Y. S., *et al*. (2008) *J. Food Comp. Anal.*, **21**, 496–500.

Donald, M. M., and Arthur, P. G. (2002) *N. Engl. J. Med.*, **347**, 2073–2076.

EMEA (1998) *Quality of Herbal Medicinal Products Guidelines*, European Agency for the Evaluation of Medicinal Products (EMEA), London.

EMEA (2002) *Points to Consider on Good Agricultural and Collection Practice for Starting Materials of Herbal Origin*, European Agency for the Evaluation of Medicinal Products (EMEA), London EMEA/HMPWP/31/99 Review.

FAO (2003) *Sampling Plans for Aflatoxin Analysis in Peanuts and Corn*, Food and Agriculture Organization of the United Nations(FAO Food and Nutrition Paper 55), Rome.

Gomez, L. R., Olsen, H. G., Lingaas, F., *et al*. (2002) *Genetics*, **162**, 1381–1388.

Gong, F., Liang, Y. Z., Xie, P. S., *et al*. (2003) *J. Chromatogr. A*, **1002**, 25–40.

Guo, F. Q., Huang, L. F., Zhou, S. Y., *et al*. (2006) *Anal. Chim. Acta*, **570**, 73–78.

Gupta, M. K. and Sharma, P. K. (2007) *Test Book of Pharmacognosy*, Ayurvedic formulations, Vol **II**, (Ist edn.). Pragati Prakashan, Meerut.

He, W., Mi, Y. L., Song, Y., *et al*. (2011) *J. Agric. Food Chem.*, **59**, 6339–6345.

IARC (1993 Working Group, International Agency for Research on Cancer (IARC)), *Some Naturally Occurring Substances*: *Food Items and Constituents*. *IARC Monographs on the Evaluation of Carcinogenic Risk to Humans*, World Health Organization, Lyon, France, vol. **56**.

ICDRA (1991) *6th International Conference on Drug Regulatory Authorities*, World Health Organization, Geneva.

Jianga, Y., David, B., Tu, P., *et al*. (2010) *Anal. Chim. Acta*, **657**, 9–18.

Jones, W. P., Chin, Y. W., and Kinghorn, A. D. (2006) *Curr. Drug Targets*, **7**, 247–264.

Kamboj, A. (2012) *Analytical Evaluation of Herbal Drugs, Drug Discovery Research in Pharmacognosy*, Vallisuta, O. (ed.), ISBN: 978-953-51-0213-7, InTech, Croatia.

Kang, J., Choi, M. Y., Kang, S., *et al*. (2008) *J. Agric. Food Chem.*, **56**, 11589–11595.

Khanuja, S. P. S., Shasany, A. K., Aruna, V., *et al*. (2002) *J. Med. Aromat. Plant Sci.*, **24**, 729–732.

Kokate, C. K., Purohit, A. P., Gokhale, S. B. (2006) *Pharmacognosy*, (35th edn). Nirali Prakashan: Pune, 98-114.

Kshirsagar, V. B., Deokate, U. A., Bharkad, V. B., *et al*. (2008) *Asian J. Res. Chem.*, **1**, 36–39.

Kunle, F. O., Egharevba, O. H., and Ahmadu, O. P. (2012) *Int. J. Biodivers. Conserv.*, **4**, 101–112.

Lau, D. T., Shaw, P. C., Wang, J., *et al*. (2001) *Planta Med.*, **67**, 456–460.

Lazarowych, N. J., and Pekos, P. (1998) *J. Drug Inform.*, **32**, 497–512.

Lee, M. S., and Edward, K. H. (1999) LC–MS Applications Development, Milestone Development Services, Pennington, New Jersey.

Li, X. N., Cui, H., Song, Y. Q., *et al*. (2003) *Phytochem. Anal.*, **14**(1), 23–33.

Liang, Y. Z., Xie, P., and Chan, K. (2004) *J. Chromatogr. B*, **812**, 53–70.

Madhav, N. V. S., Upadhya, K., and Bisht, A. (2011) *Ind. J. Physiol. Pharmacol.*, **3**, 235–238.

Manisha, K., Gharate, V. S., and Kasture, (2011) *Scholars Res. Lib.*, **3**, 28–33.

Marston, A. (2002) *Phytochemistry*, **68**, 2785–2797.

Mathe, A., and Mathe, I. (2008) *Acta Hort. (ISHS)*, **765**, 67–76.

Matthew, C., and Henry, R. (2006) *Anal. Chem.*, **78**, 3909.

McCouch, S. R., Kochert, G., Yu, Z. H., *et al*. (1998) *Theor. Appl. Genet.*, **76**, 815–829.

Michael, D. R. (1999) *West J. Med.*, **171**, 172–175.

Mihalov, J. J., Marderosian, A. D., and Pierce, J. C. (2000) *J. Agric. Food Chem.*, **48**, 3744–3752.

Mosihuzzaman, M., and Choudary, M. I. (2008) *Pure Appl. Chem.*, **80**, 2195–2230.

Mukharji, P. (2001) *Quality Control Methods of Herbal Drugs* (1st edn.). Business Horizon Pharmaceutical: Kolkata; 578-550.

Mukherjee, P. W. (2002) Quality Control of Herbal Drugs*:* An Approach to Evaluation of Botanicals, Business Horizons Publishers, New Delhi; India.

Nicholson, J. K., and London, J. C. (2008) *Nature*, **455**, 1054–1056.

Nikam, P. H., Kareparamban, K., Jadhav, A., *et al*. (2012) *J. Appl. Pharma. Sci.*, **2**, 38–44.

Ong, E. S. (2002) *J. Sep. Sci.*, **25**, 825–831.

Patil, P. S., and Rajani, S. (2010) *J. Adv. Sci. Res.*, **1**, 8–14.

Patra, K. C., Pareta, S. K., Harwansh, R. K., *et al*. (2010) *J. Pharm. Sci. Technol.*, **2**, 372–379.

Payyappallimana, U. (2009) *Role of Traditional Medicine in Primary Health Care*: *An Overview of Perspectives and Challenges*. World Health Report 2006 – http://www.who.int/whr/2006/en/ (accessed on 28 August 2009).

Raina, M. K. (2003) *Indian J. Nat. Prod.*, **19**, 11–15.

Ramakrishna, W., Lagu, M. D., Gupta, V. S., *et al*. (1994) *Appl. Genet.*, **88**, 402–406.

Rao, U. B., and Anna, N. P. (2009) *Afr. J. Pharm. Pharmacol.*, **3**, 643–650.

Rasheed, A., Reddy, S. B., and Roja, C. (2012) *Int. J. Phytother.*, **2**, 74–88.

Ratnaparkhe, M. B., Gupta, V. S., Venmurthy, M. R., *et al*. (1995) *Appl. Genet.*, **91**, 893–898.

Shanbhag, D. A., and Khandagale, N. A. (2011) *J. Chem. Pharma. Res.*, **3**, 395–401.

Shankar, K., and Singh, A. (2011) *Int. J. Res. Ayur. Pharm.*, **2**, 665–669.

Sharma, A. K., Gaurav, S. S., and Balkrishna, A. (2009) *Int. J. Green Pharm.*, **3**, 134–140.

Shaw, P. C., and But, P. P. (1995) *Planta Med.*, **61**, 466–469.

Shaw, P. C., Wong, K. L., Chan, A. W. K., *et al*. (2009) *Chin. Med.*, **4**, 21.

Shikha, S., and Mishra, N. (2009) *J. Chem. Pharm. Res.*, **1**, 1–18.

Shinde, V. M., and Dhalwal, K. (2007) *Pharmacog. Rev.*, **1**, 1–6.

Shinde, V. M., Dhalwal, K., Potdar, M., *et al*. (2009) *Int. J. Phytomed.*, **1**, 4–8.

Shrikumer, S., Maheshwari, U., Sughanti, A., *et al*. (2006) *WHO Guidelines for Herbal Drugs Standardization*, WHO, Geneva.

Shyur, L. F., Yang, N. S. (2008) Metabolomics for phytomedicine research and drug development. *Curr. Opin. Chem. Biol.*, **12**: 66-71.

Soni, K., and Naved, T. (2010) *Pharma Rev.*, 112–117.

Srivastava, S., and Mishra, N. (2009) *J. Chem. Pharma. Res.*, **1**, 1–18.

Straus, S. E. (2002) *N. Engl. J. Med.*, **347**, 2046–2056.

Sun, B., Wu, S., Li, L., *et al*. (2009) *Rapid Commun. Mass Spectrom.*, **23**, 1221–1228.

Sun, H., Zhang, A., and Wang, X. (2012) *Phytother. Res.*, **26**, 1466–1471.

Surekha, D., Rajan, V. S. T., Kumar, N. M., *et al*. (2011) *Int. J. Rev. Life Sci.*, **1**, 97–105.

Sze, S. C., Zhang, K. Y., Shaw, P. C., *et al*. (2008) *Biotechnol. Appl. Biochem.*, **49**, 149–154.

TDR (2005) *Operational Guidance*: *Information Needed to Support Clinical Trials of Herbal Products*, UNICEF/ UNDP/ WORLD Bank/WHO special program for Research and Training in Tropical Diseases.

Teo, C. C., Tan, S. N., Yong, J. W. H., *et al*. (2008) *J. Chromatogr. A*, **1182**, 34–40.

Thoppil, S. O., Cardoza, R. M., and Amin, P. D. (2001) *J. Pharm. Biomed. Anal.*, **25**, 5–20.

Vaidya, A. D. B., and Devasagayam, T. P. A. (2007) *J. Clin. Biochem.*, **41**, 1–11.

Vogel, H., Gonzalez, M., Faini, F., *et al*. (2005) *J. Ethnopharmacol.*, **97**, 97–100.

Wang, M., Lamers, R. J., Korthout, H. A., *et al*. (2005) *Phytother. Res.*, **19**, 173–182.

Wang, X., Su, M., Qiu, Y., *et al*. (2007) *J. Proteome Res.*, **6**, 3449–3455.

Wang, L., Li, F., Lu, J., *et al*. (2010) *Drug Metab. Dispos.*, **38**, 2226–2231.

WHO. (1988) The International Pharmacopeia, vol. **3**: Quality Specifications for Pharmaceutical Substances, Excipients, and Dosage forms. (3rd edn.). World Health Organization, Geneva.

WHO (1990) *Public Health Impact of Pesticides used in Agriculture*, World Health Organization, Geneva.

WHO (1992) *Quality Control Methods for Medicinal Plant Materials*, World Health Organisation, Geneva.

WHO (1996) *Quality Control Methods for Medicinal Plant Materials*, World Health Organization, Geneva.

WHO (1998) Guidelines for the Appropriate Use of Herbal MedicinesWHO Regional publications, Western pacific series No, WHO Regional Office for the Western Pacific, Manila, vol. 3.

WHO (2000) *General Guidelines for Methodologies on Research and Evaluation of Traditional Medicine*, WHO, Geneva, http://whqlibdoc.who.int/hq/2000/WHO_EDM_TRM 2000.1.pdf.

WHO (2002) Good practices for national pharmaceutical control laboratories, in WHO Expert Committee on Specifications for Pharmaceutical Preparations. Thirty-Sixth Report, World Health Organization(WHO Technical Report Series, No. 902), Geneva.

WHO (2003) *Guidelines on Good Agricultural Collection and Collection Practices (GACP)*, World Health Organization, Geneva.

WHO (2004a) *Guidelines on Safety Monitoring of Herbal Medicines in Pharmacovigilance System*, World Health Organization, Geneva.

WHO (2004b) Pharmaceutical starting materials certification scheme (SMACS): guidelines on implementation. Report of the WHO Expert Committee on Specifications for Pharmaceutical Preparations. Geneva, World Health Organization, WHO Technical Report Series, No. 917 (http://www.who.int/medicines/areas/quality_ safety / regulation_legislation/certifi cation/en/index.html).

WHO (2005) *Global Atlas of Traditional, Complementary and Alternative Medicine*, World Health Organisation, Geneva.

WHO (2007) *Guidelines for Assessing Quality of Herbal Medicines with Reference to Contaminants and Residues*, World Health Organization, Geneva.

Zafar, R., Panwar, R., and Sagar Bhanu, P. S. (2005) *Ind. Pharmac.*, **4**, 21–25.

Zhang, H. (2004) *J. Pharm. Biomed. Anal.*, **34**, 705–713.

Zhang, Y. B., Wang, J., Wang, Z. T., *et al*. (2003) *Planta Med.*, **69**, 1172–1174.

Zhao, X., Zhang, Y., Meng, X., *et al*. (2008) *J. Chromatogr. B: Analyt. Technol. Biomed. Life Sci.*, **873**, 151–158.

Part Six

Conclusion and Perspectives

47

Conclusion and Perspectives

Scientific progress is highly dependent on the continual discovery and extension of new laboratory methods and techniques and nowhere is this more evident than in the field of phytochemistry, plant biochemistry, and plant analysis. Natural products in general and plant-derived constituents in particular are gaining a tremendous interest for the development of drugs, cosmetics, agrochemicals, and nutraceuticals. Functional molecules from nature are attracting more and more researchers all over the world. This is perhaps partly due to severe health problems that occurred with the use of synthetic compounds such as thalidomide, a drug to fight nausea in pregnant women, which became an over-the-counter (OTC) medication in Germany around 1960. This resulted in more than 10 000 cases of malformation of the limbs in the newborn babies. Only about 40% of these children survived. More recently, the media reported another sanitary scandal in France related to the drug *Mediator*, which was used to treat hyperlipidemia and diabetes. This drug was considered safe and was for 33 years on the market. It is estimated that it caused around 2000 deaths! *The Lancet* published a very interesting article entitled "*Mediator scandal rocks French medical community*" (*The Lancet*, 377, 890–892, 2011). However, wrong ideas are also circulating such as *what is coming from nature is safe*. This is of course not true, as the most poisonous toxins are also coming from nature! Moreover, many of them are of plant origin. In European countries, intoxications with plants are frequent and occur 10 times more often than mushroom poisonings. Analytical methods for the safety assessment of plants, herbal drugs, and phytopharmaceuticals are thus of prime importance. They are also of great help in the search for new biologically active compounds, which is becoming a major preoccupation for numerous phytochemists. In fact, the plant kingdom still represents an enormous reservoir of new molecules to be discovered. Among the approximately 300,000 plant species existing, only a small percentage of species have been investigated phytochemically and an even smaller fraction has been submitted to biological or pharmacological screenings. The combination of analytical techniques with bioassays is necessary but requires a multidisciplinary approach. A landmark in the development of analytical methods was the coupling of chromatographic techniques with various spectroscopic techniques. This is known under the term hyphenation. HPLC–MS with new interfaces and very high sensitivities are now becoming routine methods for the quality control of herbal drugs and phytopharmaceuticals as well as for the quick identification of unknown plant constituents. Nuclear magnetic resonance spectroscopy (NMR) underwent an unimaginable development. In the years 1960–1970, only 60- and 100-MHz instruments were available, and the sample size required to obtain spectra of compounds with a molecular weight of about 400 was in the range 20–50 mg. Actually, many research laboratories are using 500-MHz instruments routinely for the structure elucidation of natural products with different probes (microprobes, cryoprobes, etc.) and the

sample size is reduced to the microgram scale! These instruments are also coupled with high performance liquid chromatography (HPLC) or ultra performance liquid chromatography (UPLC), and spectra of excellent quality can be obtained on-line from minute amounts of crude plant extracts. It is also possible to obtain two-dimensional spectra that greatly improve the structure elucidation process. Even NMR spectra can be run directly with samples of crude plant extracts. Superconducting magnets have also been elaborated and since very recently, 1000-MHz instruments are available. They will help to push the frontiers in biochemistry and structural biology and will allow the structure elucidation of very large molecules such as proteins and enzymes. Novel techniques for fractionating cellular constituents and isolating enzymes, nucleic acids, and proteins are revolutionizing our knowledge of the biochemical process of life. Biological methods are now available for the isolation of DNA from plant tissues and DNA barcoding can be used for the identification of plants. Proteomic analyses and mRNA gene expression data are useful tools to explore the physiology of cells.

The combination and simultaneous use of modern HPLC–MS instrumentation and high resolution NMR, including the direct coupling between HPLC and ultra HPLC with NMR is becoming essential for metabolomic studies and allows establishing rapidly the chemical fingerprints of various plants. HPLC–MS and NMR will greatly contribute to the discovery of new bioactive plant constituents, which could become drugs in the future. However, sophisticated chromatographic and spectroscopic methods will still depend on the sample preparation. Thus, extraction methodologies are of prime importance and new trends such as supercritical fluid extraction, microwave-assisted extraction, pressurized liquid extraction, and solid-phase microextraction are greatly improving the obtainment of samples of quality for further chromatographic and spectroscopic analyses. For the analysis of essential oils and other volatile plant constituents, gas chromatography coupled with mass spectroscopy remains a method of choice. Headspace sampling is a simple method to extract volatile constituents and it is used very often. Finally, planar chromatography techniques such as classical thin-layer chromatography (TLC) should not be neglected, as they are basic and cheap tools, requiring no sophisticated technology and instrumentation. TLC can be done in every laboratory and has still a future, especially in order to detect biologically active compounds in crude plant extracts or fractions. It is a reliable technique for the detection of radical-scavenging constituents, for example, and it can also be combined with biological detection instead of chemical detection. This is called *TLC bioautography*, which can be used as a screening method for the detection of antifungal and antimicrobial compounds in crude plant extracts. Enzymatic detection can also be performed *in situ* on the plate. This has been used in many laboratories for the search of inhibitors of the enzyme acetylcholinesterase. Inhibitors of this enzyme are of great interest in the fight against neurodegenerative diseases such as Alzheimer's disease. A judicious use of bioassays combined with chromatographic and spectroscopic techniques in screening programs will provide a valuable tool in the search of new therapeutics from the plant kingdom in the future, a reservoir that should be sensibly exploited before the large-scale destruction of many plant species goes too far. It is expected that the different techniques described in this handbook and their applications to different classes of plant constituents will help researchers to achieve this goal.

Index