CHAPTER 1

Natural Products Analysis: Instrumentation, Methods, and Applications

VLADIMÍR HAVLÍČEK and JAROSLAV SPÍŽEK
Institute of Microbiology, Academy of Sciences of the Czech Republic, Prague, Czech Republic

This book aims at highlighting the newest trends in analytical chemistry that have recently been, or soon will be, employed in the analysis of natural products and their complex mixtures. All contributing authors were motivated to stress the innovative aspects in emerging natural product chemistries and were asked to formulate their own personal visions clearly indicating which milestones can be achieved in their fields of expertise in a five-year frame. The book is structured according to analytical instrumental approaches used either routinely or experimentally for structure characterization and/or determination of both low- and high-molecular-weight natural products.

1.1 BOOK MOTIVATION

This book enumerates the most recent and cutting-edge analytical approaches including those that have not yet been commercialized into the rejuvenated natural products field. For example, less-traditional applications of synchrotron irradiation to small molecules are reported when referring to standard X-ray diffraction. Likewise, examples of the newest hyphenation techniques with impact on screening and secondary metabolism studies are described in cases in which well-known multidimensional NMR spectroscopy is discussed.

The revitalization of the natural product field is documented by an increase in the number of peer-reviewed articles illustrated by a Web of Science search (Figure 1.1).
The number of hits is seen to have increased threefold if the term “natural product activity” is evaluated. Antibacterial, antifungal, antineoplastic, anti-inflammatory, and other activities are also reported in patent literature. SciFinder returned constant data for the 2007–2013 period oscillating between 60 and 80 patent applications published annually. Diverse applications of natural products are also subjects of many review articles and book chapters. Interestingly, no monograph focused on instrumentation used for identification of natural products has been published in the past decade. This market gap was identified by Wiley senior editor Jonathan T. Rose: “In my opinion, given that plants and natural products are major sources for current and potential drugs, there is need for a book geared to researchers and professionals to facilitate natural product analysis, synthesis, and drug discovery. This type of book could explain the basics of natural products as pharmaceuticals, analytical tools and techniques, methods for isolation and elucidation, and applications for library design and in drug discovery. Such a book will find a welcome audience in organic and medicinal chemists, biochemists, analytical and medicinal chemists, microbiologists, and biomedical researchers.”

In this book the instrumentation represents the common denominator. The contributors were motivated to make a very brief introduction to physicochemical principles of their methods and give an up-to-date overview of the most important applications relevant to natural products. In a limited number of chapters the tutorial part was extended, giving the reader the opportunity to get acquainted with both the fundamentals and future trends in one place. Personal views and mutual instrumental evaluations will help the newcomers to find a suitable technique. For instance, whereas nuclear magnetic resonance spectroscopy is nonselective and less sensitive (“always tells the truth”), mass spectrometry is selectively sensitive (“tells you what you want to hear”).

**FIGURE 1.1** Report of published items accessed from the Web of Knowledge (Thomson Reuters) on December 31, 2013 illustrates the number of papers published annually in the field of “Natural Product Structure.”
1.2 THE BROAD FIELD OF NATURAL PRODUCTS

Chapters 2–4 represent medically oriented introductory chapters. Chapter 2 focuses specifically on fungi and malaria and defines the current microbiology challenges in the field of natural product discovery. These two application areas were deliberately selected because they are rather underestimated in the review literature. The importance of tackling antimicrobial resistance and the application of standardized combination therapies is stressed. Drug degradation products arising from enzyme-specific reactions, drug target reprogramming, or ejecting the drug out of the bacterial or fungal cells belong to known mechanisms by which microbes fight against antimicrobial drugs. In the field of drug resistance, cultivation of microorganisms in drug-containing stable isotope-labeled media are particularly promising. Mass spectrometry (MS) is then used for the determination of natural isotope shift reflecting the viability of the microorganism and its ability to consume and metabolize the labeled nutrients. The potential and limitations of NextGen or NextNext sequencing methods are briefly described in the perspective section in Chapter 3. The importance of peptidogenomic methods for the determination of virulence mechanisms of pathogens is accentuated by means of imaging mass spectrometry in Chapter 10.

The introductory segment of this book is terminated by Chapter 4, in which the major fractionation and isolation procedures of natural products are briefly outlined. Major attention is dedicated to the respective biological activities of natural products. The chapter is subdivided according to plant and marine origin of most important metabolites that have found significant medical applications. The authors faced a difficult task to select the clinically most important active principles, of both marine and plant origin, and align their pharmacokinetic and biological properties with medical applications. Attention was paid to organic compounds in different phases of biological trials. Most important applications of natural compounds in cardiovascular, infectious, cancer and other areas are summarized.

1.3 DISCOVERY PHASES

Recent applications of metabolomics, proteomics, mutagenomics, and genomics in exploiting bacterial natural products are summarized in Chapter 5. In mass spectrometry-based metabolomics, the problem of silent or cryptic NP biosynthesis pathways (the “silent parvome”) is discussed in the context of the quest for novel chemistries. Mass spectral alignment strategies are outlined (XCMS, MZMine, commercial products) and supported by principal component analysis program packs (SIMCA, MATLAB, etc.), the importance of which is documented (for example) on strain prioritization. Two proteomic approaches in natural product discoveries are reviewed. The first is the Kelleher group proteomic investigation of secondary metabolism (PrISM) utilizing the phosphopantetheinyl ejection assay [1]. The second proteomic technique is represented by an Orthogonal Active Site Identification System (OASIS) [2]. In the (meta)genomic part, the amplicon sequencing, shotgun
Natural products analysis libraries/metagenomics, and single-cell genomics methods are outlined and supported by success stories. Genome annotation pipelines are provided (CloVR-microbe, AntiSMASH, NCBI, SMART). The importance of concerted application of density functional theory and 2D NMR spectroscopy for absolute structure determination in natural products is stressed in the final part of the chapter. The applications of residual dipolar couplings in nuclear magnetic resonance (NMR), circular dichroism, and classical chemistry are also emphasized and create the bridge to molecular tools.

Some of them are further structured in tutorial Chapter 6 referring to the applications of electronic and vibrational spectroscopies. Advances and challenges in optical molecular spectroscopy of biomolecules and natural products are supported by chiroptic methods and placed in the context with surface-enhanced techniques and surface plasmon resonance (SPR) sensing. NMR users and fans will appreciate Chapter 7, a substantial part of which is dedicated to sample preparation and handling. Attention is also paid to LC-NMR setup, with most common instrumental variants and practical recipes (on-flow, stop-flow, and the combination of solid-phase extraction and MS). Their properties in terms of sensitivity, sample concentration, and sample nature are discussed in detail. Similarly, both supervised and unsupervised methods of statistical data evaluation are reported. Differential analysis is addressed in specialized subchapters dedicated to statistical heterospectroscopy, statistical total correlation spectroscopy, and other methods. The reader can benefit from public databases of NMR spectra and web servers dedicated to NMR metabolomics. Covariance NMR data processing of TOCSY and NOESY spectra is described. Virtual NMR chromatography (including its 3D variant) is used for distinguishing signals coming from small or large molecules. Food adulteration, plant extract analysis, and tens of other NMR applications in metabolomics are presented.

1.4 ABSOLUTE STRUCTURE

Chapter 8 describes the general technique of X-ray diffraction including the single-crystal and powder methods, and it covers advances in the instrumentation currently in use. The central argument that X-ray diffraction has a great potential and plays an increasingly important role in the structure determination of natural products is well documented and supported by the possibility to provide absolute structure determination, packing of molecules in the crystal, and structure determination in the presence of solvents in the crystal unit. Public academic software programs Sir2011, SuperFlip, CRYSTALS, and checkCIF are referred to and Cambridge Structural Database is stressed. The chapter discusses in a reader-friendly manner the common myths of X-ray diffraction (excessive time and high amount of samples needed for analyses, samples do not crystallize). It also provides a set of examples showing cases of natural product whose stereochemistry or absolute configuration originally suggested by other tools was completely revised or reassigned by X-ray. Practical comparison of what can be achieved with both laboratory or synchrotron sources and what can also be achieved with given crystal size and quality is reported.
as well. The chapter is concluded with a belief that the number of research groups producing a mere “amorphous white solid” will steadily decrease. Neutron diffraction and electron diffraction are briefly outlined. A short notice on the analysis of NRPS/PKS domains by crystallography is also presented in Chapter 12.

1.5 MASS SPECTRAL APPLICATIONS IN CONCERT

Although mass spectrometry is a mature technique celebrating its 100-year anniversary, some of its newer applications have revolutionized the emerging fields of peptidogenomics and metabolomics and also significantly contributed to revitalization of the natural product field. Chapters 9 to 15 are thus dedicated to both the instrumentation (inductively coupled plasma, imaging, ion mobility, affinity, ultrahigh resolution) and applications of mass spectrometry (ribosomal and nonribosomal natural products). In Chapter 9, solid or semisolid samples are probed by inductively coupled plasma (ICP) mass spectrometry with special attention to heteroelements—that are metals, metalloids, and nonmetals. For beginners in the field, the instrumental setup is briefly outlined with numerous applications to inorganic and organic matter analysis, including proteins separated by native polyacrylamide gel electrophoresis. Particular attention is paid to laser ablation, also when combined with 2D or 3D bioimaging approaches. The importance of elemental fractionation phenomena in quantitative determination is described and key variables defined (e.g., sample planarity, aerosol transport, vaporization, or ionization efficiency). Suppression of spectral interferences by collisional or dynamic reactive interactions in the gas phase is placed in context with the resolving power of a mass analyzer. The analytical limitations of ICP-MS are defined in a fair manner. The part on laser ablation ICP-MS comparison to other techniques of surface analysis will also be of interest to the reader. EMPA, XRD, XFA, XRA, XRF, PIXE, NAA, and some other instrumental tools are mutually compared and appropriate applications defined (including imaging). The chapter concludes with a critical personal view of quantitation in selected peer-reviewed papers reporting misleading results.

Imaging mass spectrometry is addressed in Chapter 10. In the introductory part, ionization techniques (SIMS, MALDI, and DESI) used for mass spectrometry imaging are reviewed in a tutorial manner while their practical limits and prospects for future technical development are described in the final, visionary part of the chapter. nanoDESI is defined as a central technique for bacterial imaging mass spectrometry. Ionization enrichment by derivatization and labeling strategies are outlined with a special attention to analysis of carbohydrates, oligonucleotides, and other less common molecules. Experimental considerations are defined with respect to applications in microbiology. Particular emphasis is devoted to ecology and elemental analysis with submicron spatial resolution. Biosynthesis, secretion, exchange, symbiotic interaction, or competitions are described. Many topics make this chapter interesting not only for analytical chemists and natural product fans but also for (micro)biologists and biochemists in general; for instance, the role of siderophores in iron piracy is outlined. The importance of peptidogenomic approaches is documented, for example, by
characterization of cannibalistic phenomena in bacteria and other bacterial or inter-kingdom interactions.

Chapter 11 is devoted to the specific exploration of primary and secondary metabolites by ion mobility mass spectrometry (IM-MS). In addition to a historical overview, fundamentals and instrumentation approaches introduce the reader into the field of natural product prioritization and dereplication without chromatographic separation. IM-MS offers a $10^{-4}$ lower peak capacity but is much faster (10 ms) compared to LC-FTMS separations. Different classes of biomolecules are separated in the order of increasing gas-phase packing efficiencies or densities: lipids < peptides/proteins < carbohydrates < oligonucleotides. IM structural separation can readily resolve isobaric species resulting from conformational isomers. Various arrangements for performing tandem mass spectrometry on IM instruments as well as the computational approaches for collision cross sections are provided. Contemporary efforts are underway to construct an atlas of conformation space to direct the rapid molecule identification. Future trends are targeted to peak resolution improvement and development of ion mobility imaging area.

Dereplication—that is, elimination of already known compounds from further investigation—is also depicted in Chapter 12. High-resolution tandem mass spectrometry combined with collisionally activated dissociation and/or electron-induced dissociation was used for the purpose. In its introductory part, the chapter gives basic information on the biosynthesis of nonribosomal peptides and polyketides providing the reader with the initial orientation in the enzyme systems and corresponding natural products subsequently “targeted” by other analytical tools (mainly NMR and X-ray) and strategies (PPant ejection, OASIS, PrISM) [3]. The second part of the chapter shows representative success stories, where tandem mass spectrometry was used as the main tool.

The direct and indirect affinity mass spectrometry assays for drug discovery are addressed in Chapter 10. Techniques such as fragment-based lead assembly, nanoelectrospray ionization, multitarget affinity specificity screening, detection of oligonucleotide-ligand complexes by electrospray ionization (ESI), and ESI-electron capture dissociation are included in the section of direct affinity mass spectrometry assays (MASS and DOLCE involving nucleic acid binding). Frontal affinity chromatography, affinity capillary electrophoresis, ultrafiltration, gel permeation chromatography, size exclusion chromatography, and automated ligand identification system are also reported.

1.6 COMPLEX STRUCTURES AND COMPLEX MIXTURES

The last part of the book contains three application chapters that show the complexity of natural product structures and indicate that state-of-the-art equipment is a prerequisite, but not a guarantee, for successful structure elucidation or even characterization. Chapter 13 is dedicated to ribosomally synthesized peptide toxins. Because these peptide natural products have been overviewed elsewhere [4], G. Norris and M. Patchett focused on characterizing glycosylated ribosomal products and began with
Richard Phillips Feynman’s principle of Science: “The first principle is that you must not fool yourself, and you are the easiest person to fool” (Caltech’s commencement address, 1974). This principle was then applied to the exciting story of glycocin reflecting its long-lasting development and final contributions of 2D NMR, circular dichroism spectroscopy, Edman sequencing, and Fourier transform ion cyclotron resonance mass spectrometry. The chapter also refers to useful specific or broader antibacterial peptide and prokaryotic glycoprotein repositories, including manually curated ones. Non-ribosomally synthesized and glycosylated antimicrobial peptides, as well as potential biological functions and roles of glycosylation in the discussed organisms, are also reviewed. The authors conclude with applications of venom glycopeptides, bacteriocins, and glycocins in biotechnology and the food industry.

Chapter 14 is dedicated to the description of the organic chemical diversity within complex biological and geochemical systems studied by ultrahigh-resolution mass spectrometry. Basic principles of ion cyclotron resonance are reported and various metabolome databases useful for searches based on elemental composition are described. The chapter also reports on basic statistics and mathematical tools for data visualization. The utility of Van Krevelen diagrams or Kendrick mass defect plots for selective displaying of compounds of interest, including endogenous small molecules or drugs and their metabolites, is demonstrated by the success story of Chlamydia-infected human cells.

The whole book is concluded with an application chapter (Chapter 15), in which the analytical armory is represented by electron and atomic force field microscopies. Basic principles and experimental setups in both techniques are briefly discussed in the introductory part. Amyloid fibrils represent the central subject of structural studies supported by recent literature including patents. Although these protein aggregates have been associated with more than 30 serious illnesses including Alzheimer’s, Parkinson’s, Huntington’s, prion diseases, or atherosclerosis, the chapter also highlights the important nontoxic biological functions of amyloids. Potential and proven properties of most important polypeptides (e.g., chaplins, rod-lins, or hydrophobins), are reported in the context of recent structural studies shedding light on processes like biofilm formation, microbial adhesion, initiation of aerial growth, and so on. Important applications include coating of catheters, improving biocompatibility of implants, detergent-resistant glass coatings, stainless steel lubrication to reduce friction, drug delivery systems, and many others. The fibrils may also be important in infection: They interfere with blood clotting and activate the immune system. Conversely, fibril disruption and detachment by D-amino acids can define new emerging applications in this fascinating field.

REFERENCES

