Chapter 1
Introduction

Plamen Demirev and Todd R. Sandrin

Mass Spectrometry and Microbiology

Mass spectrometry (MS) is a physical method for analysis introduced more than 100 years ago. During that period, MS applications have successfully proliferated in almost all areas of science and technology—from early studies of the structure of atoms and molecules culminating with the discovery of isotopes to characterization of planetary atmospheres and surfaces and search for extraterrestrial life. MS is an indispensable tool in organic chemistry and biochemistry for structural elucidation of various classes of natural products and synthetic compounds. In the last quarter century, advances in MS methods and instrumentation have been at the forefront of efforts to map complex biological systems, including the human metabolome, proteome, and microbiome.

MS was first successfully applied to analysis of intact microorganisms more than 40 years ago (Anhalt and Fenselau 1975). These efforts have expanded and have been particularly significant after the introduction of the soft ionization MS techniques—matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI) (Fenn et al. 1989; Tanaka 2003; Karas and Hillenkamp 1988). Both techniques (recognized by the Nobel Prize in Chemistry in 2002) allow the ionization and transfer into vacuum of large, intact, nonvolatile biomolecules, such as proteins. Various types of mass analyzers—quadrupole, ion trap, time-of-flight (TOF)—have been coupled to both MALDI and ESI ion sources, allowing multiple stages (tandem) MS to be performed for structure elucidation of analytes of interest. All these instrumental developments have allowed MS to become a well-established...
method for microorganism characterization. MS has demonstrated considerable advantage as a rapid, precise, and cost-effective method for identification, compared to conventional phenotypic techniques. The method is ultimately based on detection of organism-specific “fingerprints” (or “signatures”, i.e., biomarker molecules, from either intact and/or lysed cells (Fenselau and Demirev 2001; Wilkins et al. 2005; Demirev and Fenselau 2008a, 2008b; Seng et al. 2009; Freivald and Sauer 2009; Shah and Gharbia 2010; Ho and Reddy 2010; Bizzini and Greub 2010; Sauer and Kliem 2010; Cliff et al. 2011; Welker 2011; Fenselau and Demirev 2011; Croxatto et al. 2012; Havlicek et al. 2013; Sandrin et al. 2013; DeMarco and Ford 2013; Fenselau 2013; Clark et al. 2013; Fagerquist 2013; Calderaro et al. 2014)). Different organisms exhibit different MS signatures allowing differentiation between organisms to be made. Examples of microorganism-specific biomarkers include highly expressed intact proteins, their proteolytic products, nonribosomal peptides, polar and nonpolar lipids, RNA, and DNA. Sequence/structure-specific fragments for biomarker identification are generated by tandem MS. In top-down proteomics, these biomarkers are intact proteins, while proteolytic peptides (obtained after enzymatic or chemical hydrolysis) are mapped to their precursor proteins in bottom-up/middle-down approaches. Ultimately microorganism identification relies on mapping between spectra of unknowns with signatures of known microorganisms in MS signature libraries. Such libraries are compiled either by experimentally acquiring mass spectra of reference organisms and/or by generating in silico signatures from information in genomic or proteomic databases (Pineda et al. 2000; Demirev et al. 2004).

Thousands of reports on applications of MS for microorganism characterization in research, clinical microbiology, counter-bioterrorism, food safety, environmental monitoring, and quality have been published (Havlicek et al. 2013). Regulatory bodies in Europe, the US (FDA), and elsewhere have approved MS-based assays for infectious disease diagnostics. As of mid-2015, more than 3300 commercial MALDI TOF MS systems have been deployed worldwide in hospitals and clinical laboratories. As interest has increased in this technology, the pace of discovery and development of new applications has accelerated. The technology has been shown repeatedly to be effective at rapidly discriminating, identifying, and characterizing microorganisms at the species level and above. Some of the most promising yet challenging applications of this technology require microorganism characterization at the subspecies and strain levels. Categorization of strains sharing similar traits, differentiation of closely related strains, and/or identification of a single strain by MS techniques is desired. For example, there is tremendous need in expanding this approach to rapidly identify strains of antibiotic-resistant microorganisms.

**Chapters Included in This Book**

While previous work has covered broader approaches to using MS to characterize microorganisms at the species level or above, this book focuses on strain-level and subtyping applications. Innovators, leaders, and practitioners in the field from
around the world have contributed to this comprehensive overview of current and next-generation approaches for MS-based microbial characterization at the subspecies and strain levels. Research and developments into novel MS-based assays for antibiotic resistance determination are reviewed as well.

As an introduction to the field, Basile and Mignon present in Chap. 2 a general overview of MS ionization techniques, instrumentation, and methodology currently used for the analysis of closely related bacteria. Specific properties and parameters of the types of mass analyzers used in modern MS are listed. Important factors determining the specificity in target microorganism identification and the ability to differentiate among closely related microorganisms (i.e., selectivity) are discussed in the context of strain differentiation and antibiotic resistance determination.

Sample preparation is arguably one of the most crucial steps in efforts to identify microorganisms by MS. In Chap. 3, Ho and coworkers review different sample preparation steps currently used in the context of rapid MS analysis of microorganisms. Approaches that might eliminate the need for culturing of the target organism (currently, the key rate-limiting step), while maximizing biosafety to obtain detectable signals, are emphasized. These include protocols for intact microbial cell and/or biomarker enrichment through various affinity techniques as well as cell lysis combined with biomarker solubilization. Separation techniques (e.g., liquid chromatography) may facilitate more accurate and efficient identification of strain-specific biomolecules in microbial mixtures or complex biological samples. Since MALDI-MS is the method of choice for the rapid identification of microorganisms, a discussion on the selection of MALDI matrices and matrix solvents is included as well.

In Chap. 4, Fenselau, a pioneer in the application of MS to microbiology, stresses the overriding importance of modern proteomics and bioinformatics tools in MS approaches for microorganism identification of bacteria. Utilizing genomic database information is usually faster, more efficient, and more reliable than matching to a library of experimentally collected spectra alone. Identifications can be made without controlling sample preparation or instrumental conditions, e.g., ionization. In addition, specific biomarkers can be identified for strain identification and forensic science applications. The advantages of these proteomic strategies are illustrated in the analysis of components in mixtures, genetic engineering in bacteria, and bacteria with unsequenced genomes.

Dworzanski provides in Chap. 5 an extensive overview of bottom-up shotgun proteomics for MS-based microorganism characterization. This peptide-centric technique matches product ion mass spectra of tryptic peptides against a comprehensive database of protein sequences translated from protein-encoding open reading frames found in bacterial genomes. Phylogenomic profiles of sequenced peptides are then analyzed using numerical taxonomy tools to reveal strain identities up to the subspecies level. Bottom-up proteomics also allows sequence-based subtyping of microbial strains based on identification of proteins associated with virulence, antibiotic resistance, or used in other serotyping methods.

Methods to enhance the taxonomic resolution of MALDI TOF MS to characterize bacteria to the subspecies and strain levels are reviewed in Chap. 6 by Zhang and Sandrin. They focus on several experimental factors that will improve strain-level
P. Demirev and T. R. Sandrin

characterization efforts. These factors include culture medium, sample preparation, data acquisition, and data analysis. Specific examples illustrating both successes and challenges of this approach are presented.

Sedo and Zdrahal provide in Chap. 7 specific examples of MALDI TOF MS profiling for successful differentiation between strains of the Lactobacillus acidophilus group and selected Mycobacterium spp. In these two examples, careful optimization of the culture protocols contributed to the method robustness. In addition, strains within the Acinetobacter calcoaceticus–Acinetobacter baumannii complex, Staphylococcus aureus, and Bacillus subtilis ecotypes can be successfully typed by utilizing two alternative sample preparation protocols: alternative MALDI matrix solution or microwave-assisted tryptic digestion of the intact cells.

Lasch and coworkers describe in Chap. 8 their group’s efforts to improve taxonomic resolution without compromising the simplicity and the speed of MALDI TOF MS. Such improvements may be achieved by signature database expansion with novel and diverse strains, optimization, and standardization of sample preparation and data-acquisition protocols. Further enhancement in data analysis pipelines including more advanced spectral preprocessing, feature selection, and supervised methods of multivariate classification analysis also contribute to taxonomic resolution enhancements. Strains of Staphylococcus aureus, Enterococcus faecium, and Bacillus cereus are selected to illustrate aspects of that strategy.

Efficient methods based on MALDI TOF MS to dereplicate (i.e., group together) bacterial isolates with highly similar properties have been developed and are discussed by Vandamme and coworkers in Chap. 9. The high throughput capability and low running costs for dereplication by MALDI TOF MS allow direct microorganism identification at the species level in a large number of samples and obviate the need for more labor-intensive characterization. While isolates cultured in different media under varying conditions can be identified at the species level, isolates from the same species should be carefully re-grown in standardized conditions in order to eventually select individual peaks as strain-specific markers.

In Chap. 10, McFarlane et al. utilize liquid chromatography (LC)-MS to generate intact protein expression profiles as a snapshot of expressed proteins in a wide range of bacterial samples. Subsequent top-down proteomic analysis by LC-tandem MS allows identification of expressed serovar-specific proteins, resulting from nonsynonymous single-nucleotide polymorphisms (SNPs). Closely related, unsequenced or bacterial strains with newly acquired SNPs and plasmid proteins can be successfully differentiated by this multiplexed approach.

In Chap. 11, Drissner and coworkers provide an overview of MALDI TOF and off-line LC MALDI TOF/TOF (tandem MS) methods for typing applications. They describe further a rapid procedure for tryptic peptide generation from a simple whole-cell extract. Within minutes and without the need for further sample processing they are able to differentiate each of three different Salmonella enterica subspecies based on the detection of strain-specific peptide biomarkers.

Drug-resistant strains of pathogenic organisms are some of the most persistent and difficult to eradicate clinical infections, substantially increasing patient mortality as well as healthcare costs. Novel MALDI TOF MS methods for fast and reli-
able detection of the presence of β-lactamases in drug-resistant bacterial strains are discussed by Hrabak et al. in Chap. 12. One method involves direct detection of β-lactam hydrolysis by monitoring the molecular mass of carbapenem antibiotics. Software tools for spectral interpretation to discern drug hydrolysis will allow assay automation and high throughput. Direct detection of β-lactamases (an enzyme with a molecular weight (MW) of around 29 kDa) by MALDI TOF MS (e.g., in clinical isolates of Enterobacteriaceae) provides a complementary tool for establishing drug resistance.

Functional assays that involve the combination of MS and stable-isotope labeling for establishing drug resistance are reviewed by Demirev in the final book chapter (Chap. 13). These include global or local labeling of growth media with C, N, or H isotopes in abundance ratios differing from the natural isotope abundances of these elements. Drug resistance is determined by observing characteristic mass shifts of one or more microorganism-specific biomarkers. A similar approach involves the amplification of organism-specific bacteriophages in targeted microorganisms. In this approach, the shift in biomarker masses for phages, initially proliferated in isotopically manipulated growth medium, is monitored. The advantages of these methods as well as tools for automating the data analysis are also discussed.

** Emerging MS Methods and Technologies Not Covered Here**

This book has focused on MS methods and applications that rely on generation/analysis of protein and protein-related biomarkers for subspecies typing and strain differentiation of bacteria. These applications have matured significantly as reflected in the dominant number of MALDI TOF MS instruments installed worldwide. Several MS methods not covered here but with potential to impact future clinical applications in microbiology are pointed below.

Peptide-based MS strategies for rapid virus characterization have been developed in the last 15 years. In an early proof of concept (Yao et al. 2002), the Sindbis virus AR 339 was unambiguously identified by mapping the masses of proteolytic products to a database of tryptic peptides generated in silico from a set of viruses with sequenced genomes. Animal (swine, avian) and human flu viruses have been rapidly and reliably typed by high-resolution MS mapping of peptide digests of the isolated matrix M1 protein as well as whole-virus digests (Schwahn et al. 2010; Nguyen and Downard 2013). With the development of a phylogenetics algorithm, the method has been expanded to chart the evolutionary history of the influenza virus based on spectra produced from the proteolytic digestion of hemagglutinin (a viral coat protein; Lun et al. 2013). A high degree of overlap is observed between the mass tree (i.e., generated from MS data) when compared to trees generated from the respective viral genome sequences.

A method combining nucleic acid amplification with high-resolution MS detection relies on very accurate measurement of masses of polymerase chain reaction (PCR) products to infer the base composition (Hofstadler et al. 2005; Ecker et al.
In it, “intelligent” PCR primers target broadly conserved regions between 80 and 140 base pairs that flank the variable microorganism-specific genome regions. The PCR-amplified variable regions (both forward and reverse strands) are analyzed by ESI high-resolution and high mass accuracy MS. The accurate mass information allows unambiguous base composition determination of the amplified regions. A broad set of organisms, including the major families of human and animal viruses, bacteria, and fungi, can be identified by comparison with available genome sequences in databases. The sample preparation procedure, including PCR, currently takes more than an hour. The high degree of multiplexing (more than 1500 PCR reactions per day) facilitates surveillance of a large number of clinical samples for pathogenic microorganisms as well as virulence factors and antibiotic resistance markers.

Nonprotein (including small molecule) biomarker approaches for microorganism characterization rely predominantly on the detection of lipids or lipid constituents (Heller et al. 1987, 1988; Claydon et al. 1996; Krasny et al. 2013), e.g., fatty acids (Hendricker et al. 1999; Voorhees et al. 2006), comprising up to 10% of dry cell weight. Carbohydrates (Fox et al. 2003) and heme (Demirev et al. 2002) have also been identified as biomarkers for microorganism identification by MS. Unlike proteins, correlated directly to the genome, all secondary biomarkers exhibit much higher dependence on environmental conditions, e.g., growth medium.

A laser ablation TOF mass spectrometer has been developed to identify individual airborne micrometer-sized particles, comprising a single cell or a small number of clumped cells (Tobias et al. 2005). This approach is reagent-less, and it relies on laser ablation and detection of lower mass (less than m/z 200) positive and negative ions. MS signatures for aerosolized Mycobacterium tuberculosis particles are distinct from M. smegmatis, Bacillus atrophaeus, and B. cereus particles. This technique is tested as a stand-alone airborne M. tuberculosis detector in bioaerosols from an infected patient at airborne concentrations of 1 particle/liter.

Atmospheric pressure ionization (API) techniques are among the emerging tools and approaches developed recently that allow samples, including individual colonies, to be interrogated in ambient conditions (Song et al. 2007; Meetani et al. 2007; Pierce et al. 2007; Watrous et al. 2013; Rath et al. 2013; Strittmatter et al. 2014; Hamid et al. 2014; Fang and Dorrestein 2014; Hayes and Murray 2014; Luzzatto-Knaan et al. 2015). Lipids and other secondary metabolites are the predominant biomarkers detected by desorption electrospray ionization (DESI) in MS profiling of intact untreated bacteria (Song et al. 2007; Meetani et al. 2007). Nano-DESI MS analysis of individual bacterial colonies directly from the Petri dish without any sample preparation has provided unique information on the chemical constituents of each species in vivo and in real time (Watrous et al. 2013). Strains of 28 clinically relevant bacterial species were recently analyzed by rapid evaporative ionization MS (REIMS; Strittmatter et al. 2014). In blind tests, strains cultured on different culture media have been correctly identified more than 97% of the time. Bacterial colonies, smeared onto filter paper, can be rapidly analyzed by paper spray MS without sample preparation (Hamid et al. 2014). Phospholipids—the major bio-
markers observed in both the negative and positive ion mode spectra—allow successful bacterial discrimination at the species level by this API technique.

**Perspective**

Continuing proliferation of robust MALDI TOF MS systems in clinical laboratories in hospitals is envisioned within the next 5 years. Hardware improvements—miniaturization of the TOF mass analyzer and the laser, and sample preparation modules and associated electronics—are also expected. These will be in parallel with improved instrumental parameters—mass resolving power, mass accuracy, sensitivity, as well as reduction in instrumental and analysis costs. Introduction of new types of mass analyzers (e.g., miniature ion traps) and/or ionization sources (e.g., for API) would further expand the applications of MS in clinical microbiological diagnostics and environmental monitoring. Developments that can accelerate the environmental applications of MS include smaller, commercially available, and less-expensive MS systems with efficient on-line aerosol collectors. Additional research in lab-on-a-chip (microfluidics) devices will result in novel sample preparation protocols. Further improvement of methods for analysis of microbial mixtures, specifically of closely related strains/subspecies, and compiling of “standard” instrument-independent spectral libraries would propel the entire field forward. Software improvements including novel computer bioinformatics algorithms for rapid and automated pathogen identification will be combined with further expansion of available genomic/proteomic information. MS will play an expanded role in the development of novel, rapid, reliable, and efficient methods for detection of hard-to-confirm pathogens in bodily fluids, e.g., *Borrelia*, the causative agent of Lyme disease. The transformation of MS into a viable and widespread tool for biomedical diagnostics at point-of-care has been a long-standing goal of researchers (Mann 2002). With the improvement of current and the advent of new MS methods for pathogen detection, we are coming closer to realizing that goal.

**Disclaimer**

Mention of commercial products and/or trademarks throughout this book does not imply recommendation or endorsement and is included for information purposes only. Approved regulatory and safety procedures (e.g., microorganism inactivation, work in appropriate biosafety lab, etc.) should be followed when handling pathogens.

**Acknowledgments** During the preparation of this text P.A.D. was supported in part by a Stuart Janney Sabbatical Fellowship, administered by Johns Hopkins University/Applied Physics Laboratory.
References


Introduction


P. Demirev and T. R. Sandrin


Applications of Mass Spectrometry in Microbiology
From Strain Characterization to Rapid Screening for
Antibiotic Resistance
Demirev, P.; Sandrin, T.R. (Eds.)
2016, VIII, 336 p. 68 illus., 32 illus. in color., Hardcover
ISBN: 978-3-319-26068-6