

Chapter 2

Methods and Instrumentation in Mass Spectrometry for the Differentiation of Closely Related Microorganisms

Franco Basile and Rudolph K. Mignon

Introduction

Mass spectrometer instruments can be considered as a complex chemical reaction vessel, and as such, the resulting mass spectrum (i.e., the “product” of these reactions) is directly related to all experimental parameters, including, but not limited to, sample preparation, instrument settings, and environmental conditions. Because of its highly informative data output, mass spectrometry (MS) has found many applications in the analysis and quantitation of small to large molecular weight (MW) compounds in areas of energy, environment, forensics, space exploration, and in clinical and biological laboratories, to name just a few. To this list of applications, the analysis of microorganisms has proven to be an accurate and cost-effective approach in clinical settings. Because microorganisms can be considered as a complex chemical sample, its preparation is closely related to the information being sought, and this in turn will determine the type of MS instrumentation to be used. Unfortunately, a single sample preparation protocol will not provide a compatible sample state for all types of mass spectrometers (and vice versa). This relationship between methodology and instrumentation is illustrated (albeit simplified) in Fig. 2.1, where the final sample state prior to analysis is matched with the type of sample preparation required, instrumentation(s), and required data processing.

This relationship between the final state of the sample and MS instrumentation is mainly a consequence of the type of sample inlet and ionization technique used in a particular mass spectrometer. Referring to Fig. 2.1, the analysis of intact cells by matrix-assisted laser desorption/ionization-MS (MALDI-MS), (Jaskolla and Karas 2011) one of the simplest approaches for microorganism analysis by MS, (Holland et al. 1996) requires the isolation of a pure microbial colony, which is then deposited directly onto the MALDI plate. The subsequent mass spectral profiles,

F. Basile (✉) · R. K. Mignon

Department of Chemistry, Biodection and Mass Spectrometry Laboratory, University of Wyoming, 1000 E. University Ave (Department 3838), Laramie, WY 82071, USA
e-mail: Basile@uwyo.edu

© Springer International Publishing Switzerland 2016

P. Demirev, T. R. Sandrin (eds.), *Applications of Mass Spectrometry in Microbiology*,
DOI 10.1007/978-3-319-26070-9_2

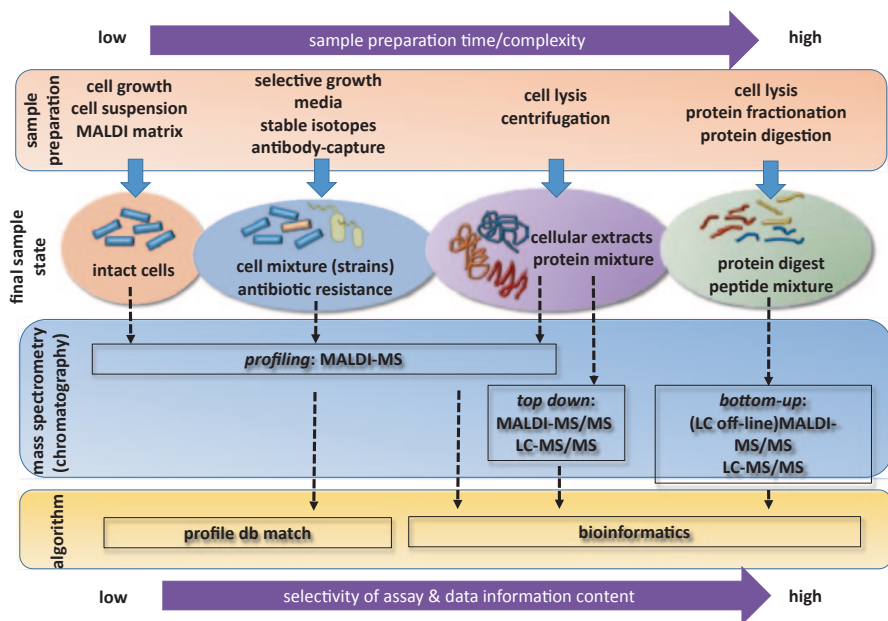


Fig. 2.1 Relationship between sample preparation time and complexity for several MS-based methods for the analysis of microorganisms

consisting mostly of ribosomal proteins, (Holland et al. 1999; Ryzhov and Fenselau 2001) are then used to classify, differentiate, and identify the microorganism. This approach requires the use of standard mass spectral databases of known microorganisms that have been acquired using the same experimental conditions. If on the other hand, one does not possess such standard mass spectral databases, a bioinformatics approach can be used. In one approach requiring a pure microbial sample, the experimentally obtained protein masses are matched to a proteome database (Demirev et al. 2001). In a second approach, a protein signal is selected for gas-phase fragmentation (tandem MS or MS/MS) and the observed ions in the tandem mass spectrum are then matched to expected fragmentation patterns of proteins contained in a proteome database (Fagerquist et al. 2010). This top-down proteomics approach can also be enhanced (i.e., more proteins detected) by the use of a liquid chromatography (LC) separation/fractionation step followed by MS/MS of the intact protein ions (McFarland et al. 2014). However, this enhancement in selectivity comes with additional sample preparation steps to extract proteins and remove other cell components incompatible with the LC step. Because of the unique ion chemistry of the protein fragmentation process and the large mass-to-charge ratios (m/z 's) of the resulting fragment ions, these top-down analyses require the use of specialized MS instrumentation that allow for the fragmentation of large protein ions and the analysis of their fragment ions with sufficient mass accuracy to provide meaningful database search results. Lastly, this bioinformatics approach can be per-

formed in a bottom-up mode where the sample preparation includes protein extraction followed by site-specific enzymatic digestion (e.g., with trypsin). The resulting complex mixture of peptides is analyzed by LC-MS/MS, and the acquired tandem mass spectrum for each peptide is matched, via a database search, to the protein originating the peptide, and if possible, its biological origin (i.e., the microorganism). The increased level of complexity for the sample preparation and/or analysis steps for both top-down and bottom-up proteomic approaches results in the highest degree of selectivity of all MS-based methods as a mixture of microorganisms can, in principle, be identified, regardless of growth conditions. Finally, the ability of MS to detect isotopologues allows the use of stable isotopes (e.g., ^{13}C , ^{15}N , or ^{18}O) to differentiate and/or quantitate biomolecules between two different cell states, as in the detection of antibiotic-resistant strains of microorganisms.

The following discussions will focus on factors affecting the ability of MS-based methods to achieve high levels of specificity and selectivity that are required in the detection of closely related bacteria and the detection of antibiotic-resistant strains, followed by a description of MS instrumentation, and examples from the current scientific literature.

Selectivity and Specificity in the Analysis of Microorganisms with MS

In the differentiation or identification of microorganisms, several factors are influential in determining the *specificity of a technique* for a target microorganism or the ability of a technique to *select among* several closely related microorganisms (i.e., selectivity). General strategies to achieve these goals include:

1. *Increasing the selectivity of the measurement to differentiate among unique features that define a certain microorganism.* This strategy may include the addition of a chromatographic step and/or increasing the mass resolution and mass accuracy of the mass spectrometer (time-of-flight (TOF), Fourier transform (FT) orbitrap or FT-ion cyclotron resonance (ICR) mass analyzers).
2. *Decreasing the overall variance of the measurement in order to detect subtle differences in traits common to all samples.* In this instance, the goal is to detect subtle differences in the pattern between two mass spectra, each obtained from different species and/or strain. Thus, the differentiation of two closely related microorganisms depends on the quantitative (relative) detection of small differences in signal strength common to both samples. Factors affecting the overall measurement variance (s^2 , where s is the standard deviation) are additive and ideally independent of each other, with the total variance of an analysis being the sum of the individual steps in the analysis

$$s_{\text{analysis}}^2 = \sum_{i=1}^n s_i^2,$$

where i is the individual step (e.g., sampling, sample preparation, measurement, data processing) in the overall analysis. In general, it has been recognized that the individual variances in the analysis follow the trend:

$$s^2(\text{sampling}) > s^2(\text{sample prep}) \gg s^2(\text{measurement})$$

Therefore, it is usually the case for most analytical protocols to focus on decreasing the variance contributions of the sampling and sample preparation steps. Manufacturers of modern chemical instrumentation, with the availability of advanced electronic components and signal processing, have considerably decreased the contribution of the measurement to the overall analysis variance. The use of automation in both sample preparation and data acquisition is key in a strategy to reduce the overall variance of the analysis. The contribution due to sampling can be reduced by increasing the number of biological samples analyzed (replicate samples).

3. *Increasing the specificity of the measurement for a target microorganism.* Factors that may increase the specificity for a target microorganism include the incorporation of a selective growth media step (antibiotic resistant), DNA amplification, antibody capture/enrichment, stable-isotope labeling, and multistage mass analyses (e.g., tandem MS or MS/MS, selective reaction monitoring or SRM, vide infra).

Approaches involving these strategies will be addressed in subsequent sections of this chapter with examples from the recent literature. However, a brief review of the MS instrumentation involved in these measurements will be presented first.

MS Instrumentation

The analysis of microorganisms with MS-based techniques involves a wide range of instrumentation, and knowledge of their capabilities and limitations is key in extracting the most information from the analysis. Two components are fundamental in defining the capabilities of any MS instrument and include the type of (i) ionization and (ii) mass analyzer used. For the techniques relevant to the characterization of biomarkers in microorganisms being discussed here, only MALDI and electrospray ionization (ESI), with the TOF, quadrupole(s), and orbitrap mass analyzers, will be described in detail. However, regardless of the type of MS instrument being used, a common operational requirement is that the final state of the sample, prior to mass analysis, be gas-phase ions of either positive or negative polarity. These gas-phase ions are then separated or sorted based on their mass-to-charge ratio or m/z , a dimensionless quantity (Price 1991; Gross 2011). (For convenience, mass, m , is expressed in terms of the unified atomic mass unit, which is defined as 1/12 the mass in kilograms of one atom of ^{12}C , u or $m_u = 1.66054 \times 10^{-27}$ kg. Thus, the quantity m is the ratio of the mass in kilograms of the ionized molecule divided by m_u or

$m = m(\text{kg})/m_u(\text{kg})$. The quantity z represents the number of elementary charges on the ion, which is also a dimensionless number) (Boyd 2008). All mass analyzers are operated under vacuum ($\sim 10^{-4}$ – 10^{-12} Torr), their magnitude depends on the mode of operation, and are required in order to avoid collisions of the analyte gas-phase ion with neutral molecules present in air (as well as avoiding arcing within components in the mass analyzer held at high voltages). This increases signal sensitivity and avoids unwanted ion–molecule reactions between the analyte ion and reactive gaseous species (e.g., oxygen).

Both MALDI and ESI are unique in their ability to form gas-phase ions from large MW molecules, biological or synthetic, without inducing fragmentations, and are thus considered to be “soft” ionization techniques (unlike “hard” ionization techniques like electron ionization (EI) which induce fragmentations *during* the ionization step) (McLafferty and Tureek 1993). ESI is considered an atmospheric pressure (AP) ionization technique since ions are generated outside the mass analyzer vacuum manifold. Although MALDI is usually conducted under vacuum in TOF-MS instruments used for bacteria identification, MALDI can also be performed under AP conditions, (Laiko et al. 2000; Madonna et al. 2003) allowing its use with instruments originally setup to use ESI, like the triple quadrupole MS.

MALDI and MALDI-MS Instrumentation

The development of MALDI by Hillenkamp and coworkers (Karas and Hillenkamp 1988) allowed for the analysis of high MW biological (e.g., proteins) and synthetic (e.g., polymers) samples without inducing fragmentation. The MALDI process relies on mixing an organic compound, termed the matrix, with the biological sample, the former in a 100:1 to 1000:1 molar excess. When the mixture is dried, the organic compound forms a heterogeneous crystalline matrix (Fig. 2.2) that surrounds and isolates individual analyte molecules in the original biological sample. Upon irradiation by a pulsed laser (UV laser in most commercial instruments), the photon energy is absorbed predominantly by the matrix compound and this electronic excitation is converted into thermal (vibrational) and translational energy, ablating (i.e., desorbing) matrix molecules as well as intact and ionized analyte molecules into the gas phase (Zenobi and Knochenmuss 1998). As such, the MALDI process is considered a pulsed ion source as it generates discrete packets of ions.

Because of the heterogeneous nature of the MALDI matrix when dry, ion yields at different locations within a MALDI matrix are not the same, leading to the description of these locations within the sample as “hot” or “cold” spots to refer to locations yielding intense or weak signals, respectively. The presence of these hot and cold signal spots within the MALDI matrix limits the usefulness of the MALDI process as a quantitative tool, imposing the need to acquire, on average, several hundred mass spectra from different locations within a sample in order to obtain a representative (average) mass spectrum.

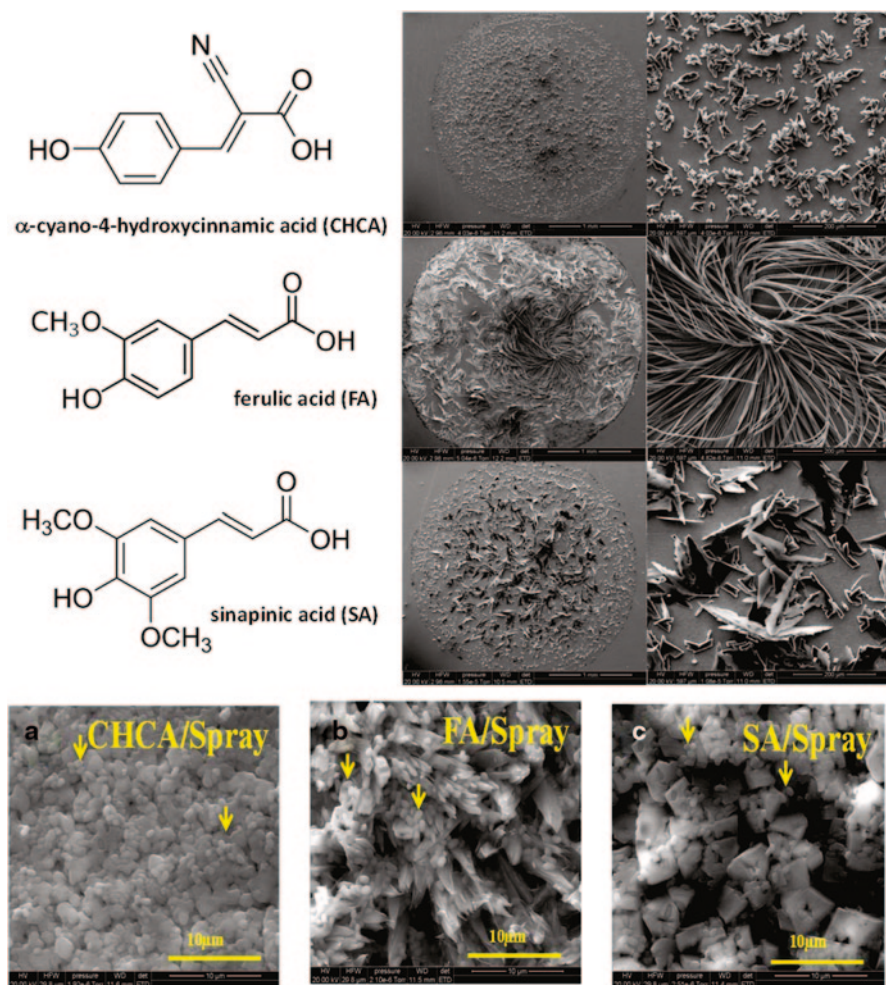


Fig. 2.2 Structures and scanning electron microscopy (SEM) photographs of different MALDI matrices deposited onto a stainless steel plate. Lower SEM photographs show *E. coli* cells co-crystallized with different MALDI matrices (matrix applied with a spray deposition technique). Arrows point to intact cells within the crystalline matrix. (Adapted from Toh-Boyo et al. 2012, copyright American Chemical Society)

The MALDI ionization process is very complex and depends heavily on the type of analyte molecule, matrix used, and laser fluence, but a recent study (Jaskolla and Karas 2011) suggests that two ionization models are mainly at play: (1) charge separation during the desorption step of *preformed ions* embedded in the crystalline matrix (a.k.a., the “Lucky Survivor” model), and (2) gas-phase protonation via ion–molecule reactions during the desorption step. In the analysis of biological molecules, the MALDI process yields primarily single-charged ions, either due to protonation or cation adduct formation (e.g., $[M+H]^+$ or $[M+Na]^+$, where M is the neutral molecule) in positive ion mode or deprotonation in negative ion mode (e.g.,

[M-H]⁺). This fact is particularly useful when analyzing a mixture of proteins as it yields a simplified mass spectrum without overlapping signals. Important to note when analyzing complex mixtures of biomolecules with MALDI is the signal suppression effect, which takes place during the ionization process. For example, in positive ion mode the signal from a highly abundant, but acidic protein may be suppressed by the presence of a low abundant, but basic protein which yields an intense signal. As a result, what is seen in the mass spectrum is neither a quantitative nor a qualitative reflection of the composition of the sample. This effect is clearly exemplified in the MALDI-MS analysis of intact bacterial cells, which mostly yields signals due to ribosomal proteins while DNA, metabolites, lipids and other high MW proteins remain undetected.

Part of the success of MALDI-MS for the analysis of microorganisms derives from the simplicity and robustness of the methodology, and in its simplest form, intact or whole cells can be deposited directly onto the MALDI plate or mixed with the matrix solution and analyzed directly. Many methods have been published describing this process, but it is believed that bacterial cells are lysed and proteins extracted into the matrix solution in the minutes before crystallization (i.e., during solvent evaporation on the plate, ~1–2 min), even though preserved cell integrity has been observed in microphotographs of the co-crystallized bacteria-matrix sample (Fig. 2.2) (Toh-Boyo et al. 2012; Madonna et al. 2000). This is backed by the fact that protocols using either solvent extraction or intact cells are both effective in producing similar protein signals, albeit with different profiles (i.e., relative peak intensities) (Basile 2011).

As mentioned earlier, it is generally agreed that the majority of the proteins observed in the analysis of bacterial cells with MALDI-MS are ribosomal proteins in the molecular mass range of 2–20 kDa (Holland et al. 1996, 1999; McFarland et al. 2014; Suarez et al. 2013). This is the case since they are abundant (almost half of the mass of growing cells), basic ($pI > 9$, easily ionized under mild acid conditions), and slightly hydrophilic in nature (easily solubilized when mixed with the matrix solution) (Ryzhov and Fenselau 2001). These facts highlight the importance of solvent composition and control of every step (i.e., exact sequence of events) (Cohen and Chait 1996) in the sample preparation protocol for MALDI-MS of bacteria, as they dictate the range of proteins detected, their observed signal strength, and overall signal pattern.

TOF Mass Analyzer

The TOF mass analyzer is suitable to measure the m/z distribution of discrete pulsed ion sources, unlike a continuous stream of ions, and for this reason it is usually coupled with MALDI, a pulsed ion source. In a TOF-MS, a discrete packet of ions with different m/z 's (generated via MALDI) are accelerated to the same kinetic energy by applying a voltage ($U \sim 10$ –25 kV; direct current, DC) to the stainless steel sample plate. These ions enter a field-free region (no voltage or magnetic fields

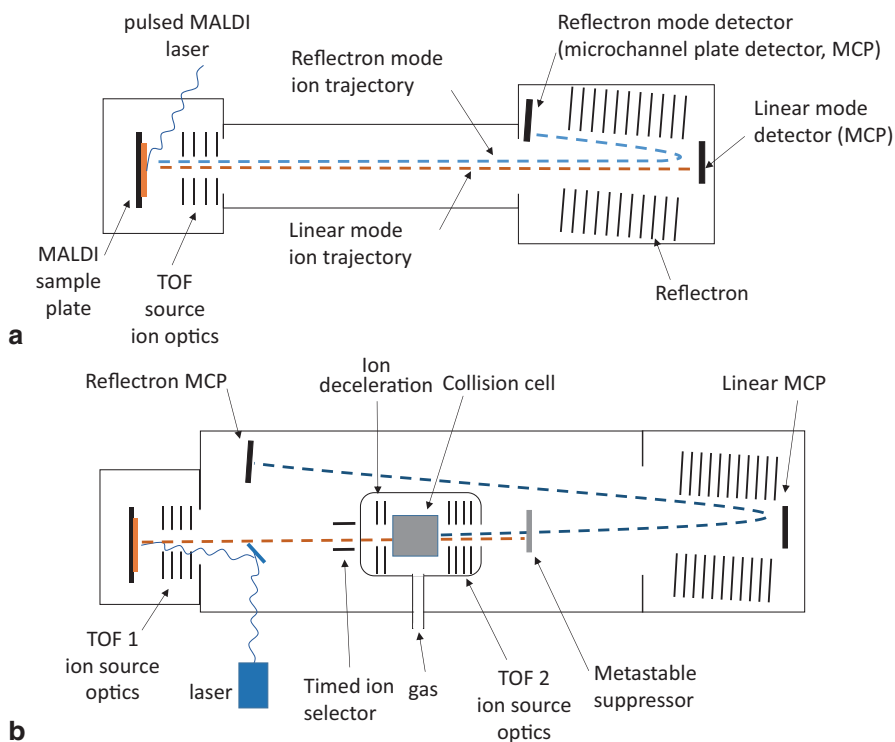


Fig. 2.3 Simplified general diagrams of **a** MALDI-TOF-MS and **b** MALDI-TOF/TOF-MS (based on the Sciex 4800/5800™ systems)

applied) where ions with small m/z 's travel faster than those with large m/z 's, and the different times to travel a predefined distance (d) forms the basis for their mass separation. The simplified relationship between TOF (t_{TOF}) and m/z is given by:

$$t_{\text{TOF}} = \frac{d}{\sqrt{2U}} \sqrt{\frac{m}{z}}.$$

In principle, the TOF-MS does not have an upper mass limit; however, in practice they are limited by the efficiency of the multichannel plate (MCP) detector (vide infra, Fig. 2.3) in converting low kinetic energy ions (i.e., large m/z 's) into a detectable electrical current, and the ability of the ionization source to produce ions of large m/z . Operationally, the relationship between m/z and t_{TOF} is established by calibration with a set of standard compounds of known m/z values for their $[\text{M} + \text{H}]^+$ ions. This calibration is dependent on matrix type and laser intensity (each affects the initial ion velocity during desorption) and the sample position within the MALDI plate (affecting the distance traveled, d , and the effective accelerating voltage, U , experienced by the desorbed ion). A simplified general diagram of a MALDI-TOF-MS instrument is illustrated in Fig. 2.3a.

Instruments based on this design (linear, but not necessarily reflectron) form part of most, if not all, of the commercially available MALDI-TOF-MS microorganism identification systems that are based on matching a mass spectrum to a mass spectral library of microorganisms (i.e., profile-based MALDI-MS). When operated in the reflectron mode to increase the mass resolution of the measurement (practical up to $\sim m/z$ 5000), current state-of-the-art TOF mass analyzers (with a properly designed MALDI ion source) specify mass accuracies in the 1 ppm or ± 0.001 (at m/z 1000). However, no peptide/protein sequence information can be derived from this mass measurement alone and an additional level of selectivity, tandem MS or MS/MS, is required to obtain this information.

Another type of MS available with the MALDI ion source and based on the TOF mass analyzer is the tandem TOF or TOF/TOF-MS. This configuration has two TOF mass analyzers configured in series, separated by a collision cell. This configuration has the capability of obtaining mass spectral protein profiles as well as sequence information of peptides (up to ~ 4000 Da), and for one manufacturer mid-sized proteins (5–15 kDa). A simplified schematic of a MALDI-TOF/TOF-MS instrument is illustrated in Fig. 2.3b (the following discussion is based on the Sciex 4800™ system (Yergey 2002). An excellent discussion of the inner workings of the Bruker MALDI-TOF/TOF-MS system can be found in Suckau et al. (2003)). Ions formed in the MALDI ion source are accelerated toward the first TOF mass analyzer (TOF1), where ions are separated according to their m/z 's. In the MS mode, ions are allowed to travel uninterrupted to either the linear or reflectron detector. In the tandem MS (MS/MS) mode, ions of a single m/z value are selected and allowed to enter the collision cell. This m/z selection is performed via a timed ion selector, with a series of voltages applied at a unique time on the path of the ion beam so as to deflect all ions except those of the desired m/z (or TOF). The selected ion is introduced into the collision cell filled with a neutral gas like argon or nitrogen, and upon collision, fragment ions are formed via collision-induced dissociation (CID). Precursor ions can also undergo fragmentation during the MALDI process, via laser-induced dissociation (LID), (Suckau et al. 2003) or after the MALDI process, via post-source decay (PSD), (Neubert et al. 2004; Fagerquist 2013) where metastable ions leave the MALDI ion source and fragment during their voyage through TOF1. In all these fragmentation events, CID, LID, or PSD, the generated fragment ions will have roughly the same velocity as the precursor ion, and thus they cannot be discriminated by their m/z 's within the TOF1 mass analyzer. The TOF/TOF instrument achieves mass separation of these fragment ions (and obtains useful sequence information) by re-acceleration of this ion packet into the second TOF mass analyzer (TOF2). This second acceleration event becomes the starting point for recording the fragment ion mass spectrum (Yergey et al. 2002). For small proteins (15 kDa), this type of instrument can be used for top-down proteomic measurements, where the fragmentation of a single protein signal from a mixture can yield sequence information about the precursor ion and has been used to discriminate proteins varying by a single amino acid in their sequence (Fagerquist et al. 2010). Finally, when coupled with offline LC and fraction collection directly onto the MALDI plate, this type of instrumentation allows for bottom-up proteomic measurements (Marcus et al. 2007; Benkali et al. 2008; Bodnar et al. 2003).

ESI and ESI-MS Instrumentation

The development of ESI as an ionization source for MS by Yamashita and Fenn (1984) allowed the formation of gas-phase ions of biomolecules in liquid samples, thus enabling the analysis of intact proteins in solutions and of samples separated by LC. A detailed discussion of ESI is beyond the scope of this chapter as many excellent reviews and books have been written on the subject (Bruins et al. 1998; Cech and Enke 2001; Cole 2008). In general, the ESI process in positive ion mode for most biomolecules (proteins and peptides) starts in an acidified solution, that is, by the formation of ions via protonation of basic groups. This is typically accomplished by the addition of a volatile organic acid like acetic acid or formic acid (1 % or 0.1 %, respectively) in a 50 % organic-aqueous solvent (methanol or acetonitrile). The solution is then driven into a metal capillary (~ 50 – $100\ \mu\text{m}$ inner diameter) connected to a power supply at 3–4 kV (DC voltage). As the liquid emerges at the open end of the capillary, the large electric field causes charge separation of the preformed ions in solution. In the case of biomolecules, an ionized peptide (positive charge) is separated from either a formate ion (HCOO^-) or the acetate ion (CH_3COO^-). The use of trifluoroacetic acid (TFA) to acidify solutions for ESI analysis is discouraged as the CF_3COO^- ion forms a strong ion pair with the positively charged biomolecule, making charge separation difficult and thus lowering the ionization efficiency of the ESI process. This accumulation of positively charged ions at the open end of the capillary causes the deformation of the liquid meniscus into what is termed a Taylor cone. Eventually, the electrostatic repulsive forces between the positive charges accumulated in the meniscus exceed the surface tension of the liquid leading to the formation of a fine jet of liquid, which breaks into fine droplets, each containing an excess of positively charged molecules. According to the ion evaporation model (IEM), (Nguyen and Fenn 2007) these droplets undergo a cascade of evaporative and Coulomb fission (charge repulsion) cycles until droplets of about 10 nm in diameter are formed. At this droplet size the effective electric field at the surface is large enough to push one or more solvated ions into the gas phase. A second ionization model, the charge residue model (CRM), describes the generation of an ion when all the solvent is evaporated from the droplet. Although there are many studies showing the prevalence of one model over the other in ESI, the consensus is that large ionized molecules (1000 u) are generated by a process closely described by the CRM. On the other hand, smaller and solvated ions can be emitted from nano-droplets by a process better described by the IEM (Wilm 2011).

In general, for positive ion mode ESI, ionization efficiency is dictated not only by the basicity of the molecule, but also by its hydrophobicity, which determines its concentration at the surface of the droplet (i.e., surface activity). As a result, not all biomolecules present in the sample are ionized with the same efficiency. That is, basic and hydrophobic molecules (with a high surface activity) tend to ionize more efficiently than basic and highly polar molecules. For example, a peptide with a high content of hydrophobic amino acids (phenylalanine, tryptophan) will experience a higher ESI ionization efficiency than a peptide of the same charge but with amino

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