Current Methods for Mass Spectrometric Analysis of Intact Proteins and Peptides

Mass spectrometry has had an interesting history as an analytical chemistry instrument. While MS is popularly associated with the chemical analysis of molecules, it was developed in the late 1800s by the physicist J. J. Thomson as a technique to measure the mass-to-charge ratio of the electron [1]. For the next 50 years, MS remained almost solely in the domain of physics as a technique to study atomic structure [2]. In the 1940s, however, MS was increasingly used in industry to quantitatively measure chemicals in mixtures. The next big step in MS was the recognition that specific fragmentation patterns of small organic molecules could be used to identify unknown molecules [3]. Proteins and other large macromolecules, however, could still not be analyzed using MS owing to the difficulty in getting these molecules into the gas phase. In 1988, however, electrospray and matrix-assisted laser desorption ionization were developed, enabling large biomolecules, especially proteins, to be routinely analyzed using MS [4, 5]. These revolutionary developments have led to an exponential use of MS as a biological analytical tool. Forty years ago, MS instrumentation was sparingly used; today, many large institutions have put major investments into facilities to meet the demand for protein analysis using MS.

The most fundamental use that makes MS so invaluable in protein research is its ability to accurately identify proteins in complex mixtures. While Western blotting used to be considered the gold standard of protein identification, MS slowly overtook this technique to become the premier analytical method for this purpose. Mass spectrometers are now capable of identifying thousands of proteins within extremely complex proteome samples in a matter of hours. This capability has greatly accelerated the discovery of proteins involved in diseases (i.e., biomarkers) and protein complexes that functionally interact to carry out cellular processes. A natural evolution of global identification was the development of techniques to quantitate proteins enabling MS to reliably measure disease-related biomarkers or compare protein levels in different samples. Over the past 20 years, techniques have been developed to meet both needs. In their article, Jayathirtha et al. describe the steps involved in the identification of proteins using MS, from sample preparation to database searching and interpretation [6]. Building upon this article, Rotella et al. [7] describe the development of various techniques that utilize MS to measure the relative or absolute abundance of proteins in simple and complex mixtures. These techniques have not only enabled the development of clinical MS methods through the absolute measurement of specific proteins in biological samples but also the quantitative comparison of thousands of proteins between samples analyzed across different labs. Both articles provide a broad review of the available technologies enabling the reader to determine the specific method that may be best suited for their own research interests.

As the analytical capabilities of mass spectrometers increased, so did the amount of data they produced. This increase has spurred the co-development of software programs aimed at storing, analyzing, organizing, and data mining complex datasets. In an article, Rusconi presents a review of open file formats produced in MS data acquisition and how the Free and Open Source Software (FOSS) movement has provided opportunities for scientists to learn skills to develop software for solving MS data analysis problems [8]. Part of this development includes the detection of post-translational modifications (PTMs), which increase the complexity of the human proteome well beyond the approximately 20,000 protein-encoding genes found within the human genome. With over 200 different types, the discovery-driven ability of MS has exponentially increased the rate at which PTMs have been discovered. In the review by Yu et al., MS methods used to identify phosphorylation sites (arguably the most important PTM) within single proteins as well as across thousands of proteins in a complex sample are described [9]. In addition, this article illustrates an example showing how MS data can be used to characterize signaling events as a result of cells being treated via a specific perturbation.

While the identification and quantitation of proteins extracted from cells remain the dominant application of MS, exciting techniques that enable proteins to be characterized while still within the cell have been developed. In their review, Vu et al. [10] describe how mass spectrometry imaging (MSI) is used to directly analyze neuropeptides in thin tissue sections. They then proceed to describe recent developments that enable neuropeptides to be characterized in clinical settings.

Owing to the current pandemic, this year has been like none that we have seen before. I want to express my deepest gratitude to each author who provided an article for this special edition of Current Protein and Peptide Science. With all of the adjustments that each of you had to make to your current schedules, you still took the time to produce an edition that will educate the reader about the current and potential capabilities of MS for protein analysis.

REFERENCES


**Timothy Veenstra**  
(Guest Editor)  
School of Pharmacy  
Cedarville University  
Cedarville, OH 45314  
USA  
Tel: 937-766-7484  
tveenstra@cedarville.edu