CHAPTER 16
Mass Spectrometry

INTRODUCTION

From the point of view of protein scientists, in the course of the 1990s mass spectrometry (MS) has developed from a fairly exotic technique practiced in only a handful of laboratories into a routine, essential tool for protein characterization. MS is complementary to other protein and peptide analysis methods such as protein sequencing, amino acid analysis, and high-performance liquid chromatography (HPLC) peptide mapping (Chapter 11). It is particularly useful for identifying post-translational modifications (Chapters 12 to 14), which are often quite difficult to distinguish by other methods. To some extent, MS is beginning to replace alternative methods such as Edman sequencing as continuing advances in instruments, software, and user-friendliness improve the sensitivity and utility of MS methods.

A detailed discussion of peptide and protein MS analysis, including sample ionization methods, basic principles and parameters, and data interpretation, is presented in UNIT 16.1. The unit also lists numerous literature references that should be especially useful to scientists seeking more detailed information on specific MS-related issues.

A key advance in the application of MS for routine analysis of peptides and proteins has been the development of commercial instruments that employ matrix-assisted laser desorption/ionization (MALDI) techniques to transfer peptides and proteins efficiently into the gas phase for time-of-flight based mass measurements. This technique has multiple advantages compared with older techniques, including the instruments’ relatively low cost and ease of operation, relatively good tolerance of nonvolatile buffer components, and ability to analyze heterogeneous samples as mixtures, since the ionization method predominantly produces a single charged species per sample component. A general discussion of MALDI-MS analysis of peptides is presented in UNIT 16.2, and UNIT 16.3 provides specific protocols for analysis of both peptides and proteins, with emphasis on sample preparation.

A frequent research question in biology is the correlation of observed function with the proteins responsible. MS-based methods have emerged as the most sensitive, rapid, and economical method for identifying proteins. The simplest approach is MALDI-MS fingerprint mapping following in-gel protease digestion of the protein of interest (UNIT 16.4). The observed peptide masses are used to search sequence databases over the Internet (UNIT 16.5) using one of several alternative algorithms. This method works best when the protein of interest is expected to be found in an available sequence database—i.e., when the complete genome of the species from which the protein was isolated has been sequenced. When the database contains only related, but nonidentical, proteins, or when the results of the fingerprint mapping search are ambiguous, partial sequence information is generally required to identify homologous proteins or to search expressed-sequence tag (EST) databases. Such partial sequence information can be obtained on a MALDI spectrometer either using fragmentation induced within the mass analyzer (UNIT 16.6) or via exopeptidase digestion of samples on the MS sample target (UNIT 16.7).

The other common mass analysis method for biomolecules is electrospray ionization (ESI)-MS (UNITS 16.1, 16.8, & 16.9). In recent years, the cost of ESI-based instruments has

Contributed by David W. Speicher

Copyright © 2000 by John Wiley & Sons, Inc.
decreased and they have become more sensitive and user friendly. Triple-quadrupole and ion-trap instruments with ESI interfaces are particularly well suited to tandem mass spectrometry (MS-MS) experiments, in which an initial mass measurement is followed by fragmentation of a target peptide and reanalysis of fragment mass spectra.

ESI-MS is much less tolerant of salts and buffers than MALDI-MS. Hence, interfering components can be removed by trapping the peptides or proteins on a small amount (typically <2 µl) of reversed-phase resin. After removing polar components with water, the peptides or proteins are eluted with a few microliters of a high-purity organic solvent such as acetonitrile. The desalted samples are then introduced into an ESI mass spectrometer using nanoliter flow rates from a borosilicate nanospray glass needle (UNIT 16.8). This sample preparation/analysis is especially well suited for detailed structural studies of simple samples containing a small number of protein or peptide components, e.g., analysis of posttranslationally modified peptides. Alternatively, samples can be introduced into ESI mass spectrometers by on-line liquid chromatography (LC) using a microscale capillary reversed-phase column incorporated into the electrospray needle (UNIT 16.9). The on-line reversed phase column can concentrate the sample, remove contaminating buffers and salts, and partially separate peptide/protein components prior to MS analysis. This is the method of choice for analyzing complex mixtures such as tryptic peptides from in-gel digestion of proteins excised from polyacrylamide gels. When on-line nanocapillary LC columns are used without post-column stream splitting, very high sensitivities can be obtained.

One of the most common uses of tandem MS (MS/MS) spectra is identification of proteins in polyacrylamide gels or purified macromolecular complexes using the SEQUEST software to search DNA and protein sequence databases (UNIT 16.10). However, database searches with SEQUEST or analogous MS/MS spectral matching algorithms are most reliable when the complete sequence of the genome for the species being studied is in the database that is searched. Even when a highly homologous protein from a closely related species is in the database, a definitive match may be missed because sequence coverage by the MS/MS data is rarely complete and nearly all substitutions change tryptic peptide masses. High-quality MS/MS spectra that do not match sequences in databases can be manually interpreted; that is, by de novo peptide sequencing (UNIT 16.11). De novo sequence interpretation is also used to locate specific residues that are posttranslationally or chemically modified, or to locate naturally occurring substitutions in a protein with a known sequence.

David W. Speicher