Annu. Rev. Biophys. Biomol. Struct. 2004. 33:297–316
doi: 10.1146/annurev.biophys.33.111502.082538


# MASS SPECTRAL ANALYSIS IN PROTEOMICS

John R. Yates, III
*Department of Cell Biology, The Scripps Research Institute, La Jolla, California 92037; email: jyates@scripps.edu*

**Key Words** MALDI, ESI, tandem mass spectrometry, multidimensional separations

■ **Abstract** Mass spectrometry provides key tools for the analysis of proteins. New types of mass spectrometers that provide enhanced capability to discover protein identities and perform improved proteomic experiments are discussed. Handling the complex mixtures of peptides and proteins generated from protein complexes and whole cells requires multidimensional separations; several forms of separation are discussed. Applications of mass spectrometry–based approaches for contemporary proteomic analyses are described.



CONTENTS

INTRODUCTION . . . . . 297
MASS SPECTROMETRY . . . . . 298
Ionization Techniques . . . . . 298
New Types of Mass Spectrometers . . . . . 301
SEPARATIONS OF COMPLEX PEPTIDE AND PROTEIN MIXTURES . . . . . 307
Chromatographic Separations . . . . . 307
Capillary Electrophoresis Separations . . . . . 308
SHOTGUN PROTEOMICS OF PROTEINS . . . . . 309
TOP-DOWN ANALYSIS OF PROTEINS . . . . . 310
PROTEOMIC INFORMATICS . . . . . 311
CONCLUSION . . . . . 311

## INTRODUCTION

The year 2002 celebrated the development of two key mass spectrometry ionization techniques with a Nobel Prize in Chemistry. This award emphasizes the critical role ionization of molecules plays in the measurement of ions using mass spectrometry; ionization techniques are truly the "horses that pull the cart" in this field. In a great convergence, these new ionization techniques emerged as plans to sequence the human genome and the genomes of other organisms came into the mainstream. Although the enthusiasm for whole genome sequencing overshadowed the role of

protein biochemistry in deciphering the information derived from these projects, these ionization techniques enabled new, more sensitive, and higher-throughput approaches for the analysis of proteins and triggered the development of new and improved types of mass spectrometers. Together, these devices are making significant inroads into the translation of the data derived from genome sequences into biological discoveries.

Genome sequences are analyzed to identify genes that are translated into proteins, and this set of information has provided a sequence infrastructure to analyze mass spectrometry data of peptides and proteins. In this way, proteins can be identified, predicted protein sequences verified, modifications determined, and alternative splicing discovered. The dynamic nature of protein expression, the splicing together of different exons to create protein variations, and the use of regulatory modifications present significant challenges to translate the information derived from genomics into biological understanding. Explosive developments of technology and methodology in the field of mass spectrometry and proteomics are providing improved or new strategies to speed biological discovery. This review covers strategic developments in mass spectrometry and applications to proteomics.



## MASS SPECTROMETRY

Mass spectrometers consist of three basic components: an ion source, a mass analyzer, and an ion detector. Molecules need to be converted into gas-phase ions, and thus a method is required to transfer molecules from solution or solid phase into the gas phase as ions. Ions are then separated in a mass analyzer on the basis of mass-to-charge ratio. Ion motion in the mass analyzer can be manipulated by electric or magnetic fields to direct ions to a detector in an m/z-dependent manner. Ions strike a detector generating a record of the number of events or the electrical current created.

### Ionization Techniques

Ionization techniques are critical to convert molecules into ions that can then be manipulated within electric or magnetic fields. The challenge with biological molecules such as peptides and proteins is to convert polar, zwitterionic molecules into gas-phase ions without degradation or fragmentation.

ELECTROSPRAY IONIZATION Electrospray ionization (ESI) creates ions by spraying an electrically generated fine mist of ions into the inlet of a mass spectrometer at atmospheric pressure (17). By creating a potential difference between the capillary through which the liquid flows and the inlet to the mass spectrometer, small droplets of liquid are formed. These are transferred into either a heating device to cause evaporation of solvent or a device to induce fragmentation of the droplets

into increasingly smaller sizes. Once the droplets have reached the Rayleigh limit (the point in which charge repulsion exceeds the surface tension of the droplet), ions are desorbed from the droplet to create bare ions, which are then transmitted into the ion optics of the mass spectrometer. With proper tuning of the electrospray elements (e.g., heat, gas pressure), this desorption proceeds rapidly. ESI can readily convert solution-phase molecules into gas-phase ions; however, the ions produced are multiply charged. This situation complicates calculation of molecular weight because there is not a one-to-one relationship between m/z value and molecular weight ($z \neq 1$). Multiple charging does, however, lower the m/z value observed on the mass spectrometer, allowing the use of less sophisticated mass analyzers.

Steady advances in the application of ESI to the analysis of peptides and proteins have been made. The most notable improvements have come from reduction in the flow rate of the liquid used to create the electrospray and thus more efficiently create ions. Smith et al. (65) showed that when small-diameter capillaries are used for capillary electrophoresis (CE), the flow rate is significantly reduced and sensitivity is improved. Emmett & Caprioli (14) demonstrated the same principle with capillaries packed with reversed-phase (RP) packing material for small-scale liquid chromatography (LC). A key element to achieving electrospray stability at low flow rates is the presence of a small-diameter orifice at the end of the capillary. An added advantage to the use of separation techniques with low flow rate ESI is an increase in the concentration of the analyte as it elutes off the column. Separated molecules elute in a smaller volume of liquid at low flow rates in short periods, which increases the amount of material introduced into the mass spectrometer. Wilm & Mann (76) showed the use of low flow rates for infusion-based electrospray to produce long ionization times for small volumes of sample. Over time the utility of this method decreased as limitations of signal to noise of mixture infusion became apparent, prompting the development of methods to identify peptide signals among a high background of noise (32).

MATRIX-ASSISTED LASER DESORPTION IONIZATION To use laser energy to convert molecules into gas-phase ions has been a long-standing goal. Laser light has a specific energy level, but attempts to use this energy to effect thermal desorption of molecules into gas-phase ions were ineffective. Tanaka (70) finally demonstrated that ions could be desorbed from a glycerol-based liquid laced with a finely powdered metal. Rather than tuning the laser energy to absorption wavelengths of the analyte, a "matrix" that absorbs the energy of the laser is used to assist and promote thermal desorption. The glycerol/metal powder method quickly gave way to an approach that is based on cocrystallization of the analyte with an organic matrix, matrix-assisted laser desorption ionization (MALDI), as a more practical and reliable method to produce ions (27). A pulsed laser is used to deposit energy into the matrix, causing rapid thermal heating of the molecules to desorb molecules and ions into the gas phase. Because a pulsed laser is used, this ionization technique produces ions in packets rather than a continuous beam and thus requires a mass spectrometer that can either measure a complete mass spectrum without

scanning a mass range or trap all the ions for a subsequent mass analysis. This type of ionization has seen an interesting number of uses over the past few years.

The speed of analysis depends on how fast the laser can be pulsed. Over the past few years the typical laser used for MALDI applications has been a 20-Hz nitrogen laser (314 nm). Recently, 200-Hz lasers that increase the rate of analysis by a factor of 10 have appeared. An increase in speed improves the speed of protein identification experiments and creates new analytical applications. Walker et al. (73) first demonstrated the deposition of separations onto a sample plate used for MALDI. Preisler et al. (56) and Rejtar et al. (55) integrated CE separations with MALDI by depositing the effluent of the CE separation in a continuous stream on the sample plate for high-speed and parallel MALDI experiments. Griffin et al. (22) deposited effluent from LC separations in discrete spots on the MALDI plate that was then combined with a tandem mass spectrometer for peptide analysis. Ericson et al. (16) performed multiplexed chromatographic separations in combination with sample deposition on a MALDI sample plate. An advantage to performing separations off-line is the ability to multiplex separations and store them for future analyses. Furthermore, because the pulsing of the laser is controlled by the operator, after each pulse the data can be evaluated and decisions can be made as to how to proceed. Paralleling of off-line separations is effective when the speed of mass spectrometry analysis for each separation is dramatically faster than the separation, as a mass spectrometer is a serial analysis device.

The ability to analyze mixtures of peptides and polypeptides has presented new analytical opportunities for MALDI mass spectrometry. Sample plate surfaces have been modified with separation chemistries to allow molecules to be enriched prior to application of matrix or to immobilize affinity reagents such as antibodies to enrich for a specific molecule. In addition, MALDI can be used for molecular imaging of thin slices of tissues and cells (7, 8, 20, 38, 60, 68, 71, 78). Tissue imaging is performed by collecting thin sections from tissues tumors and then spreading them onto a MALDI sample stage. The surface of the tissue is then coated with an organic matrix. The laser is then scanned incrementally across the tissue's surface to ionize peptides and polypeptides. Ion images can be reconstructed to identify sections of the tissue where specific types of peptides or polypeptides localize. Ultimately, this method may have uses in the diagnosis of diseases (by identifying molecular patterns associated with specific types of diseases). Several analytical challenges are associated with this method. Ionization and mass analysis are typically biased toward lower-molecular-weight components in the tissue because ion suppression results from the molecular complexity of the tissue sample. Slow scanning speeds are necessary to increase ion image density across the tissue, but resolution is still insufficient to observe subcellular structures (e.g., image organelles). Ion suppression resulting from sample complexity can be relieved by separating protein mixtures prior to ionization.

Petricoin et al. (53) used separation surfaces directly on a sample plate to perform chromatographic separations with serum samples. Material retained on the surface is then analyzed by surface-enhanced desorption ionization (SELDI), a

variation on the MALDI technique, and the patterns of ions observed are analyzed using statistical techniques to identify ions diagnostic of disease. This strategy has been used to profile patient serum to identify patterns of ions that may be indicative of disease state (53, 54). Yanagisawa and colleagues (53, 79) demonstrated a similar analysis using the tissue imaging method to profile tissues of nonsmall cell lung tumors. A limitation to both approaches arises from a bias against high-molecular-weight proteins that are not observed in the analyses. At least with tissue imaging the scanning speed of the laser is slow relative to the surface area needed to be covered. Both methods, however, show promise for the discovery of biomarkers.

## New Types of Mass Spectrometers

Ionization techniques deliver ions to the mass analyzer. Both MALDI and ESI are soft ionization techniques in that ions are created with low internal energy and thus undergo little fragmentation. Mass-to-charge ratios can be readily and accurately measured for the intact ions, but this information does not provide data on the covalent structure of the ion. For peptides and proteins in particular, data related to the amino acid sequence of the molecule are desired. To generate this information, new configurations of mass spectrometers have been developed to isolate ions, fragment them, and then measure the mass-to-charge ratio of the fragments. These devices are collectively called tandem mass spectrometers. The performance of tandem mass spectrometers is differentiated in terms of mass accuracy, mass resolution, robustness, and ease of operation (Table 1). The description below focuses on reasonably new instruments or instruments whose performance and use have recently become important in the field of proteomics.

LINEAR TWO-DIMENSIONAL ION TRAP Quadrupole ion traps have been an important mass spectrometer for proteomic experiments (31). These devices are based on the three-dimensional quadrupole, which traps ions in the center of the device and then scans the ions from the trap to a detector (Figure 1*A*). Several issues limit the performance of three-dimensional ion traps (31). First, there is a limit to the number of ions that can be trapped in the device. Second, when ions are scanned from the trap, half exit in the direction of the detector and the other half exit in

**TABLE 1** Comparison of performance characteristics for new types of tandem mass spectrometers

| Instrument | Resolution | Mass accuracy | MS scan rate | MS/MS scan rate |
|---|---|---|---|---|
| Linear Ion Traps | 2–15,000 | 100–300 ppm | Moderate to fast | Moderate to fast |
| Q-QIT | 2000 | 300–500 ppm | Moderate | Moderate |
| TOF-TOF | 20–25,000 | 10–20 ppm | Fast | Slow |
| FTMS | 50–100,000 | 1–2 ppm | Moderate | Moderate |

ppm = parts per million

the opposite direction. Third, there is a limitation in mass accuracy and resolution, although a narrow mass range scan can be employed to obtain high-resolution data with improved mass accuracy. To circumvent some of these limitations, new mass spectrometers based on a two-dimensional quadrupole ion trap mass spectrometer have been developed (6, 59).

Two-dimensional quadrupole ion traps or linear ion traps can hold almost 10 times more ions than three-dimensional traps (59). This increased volume

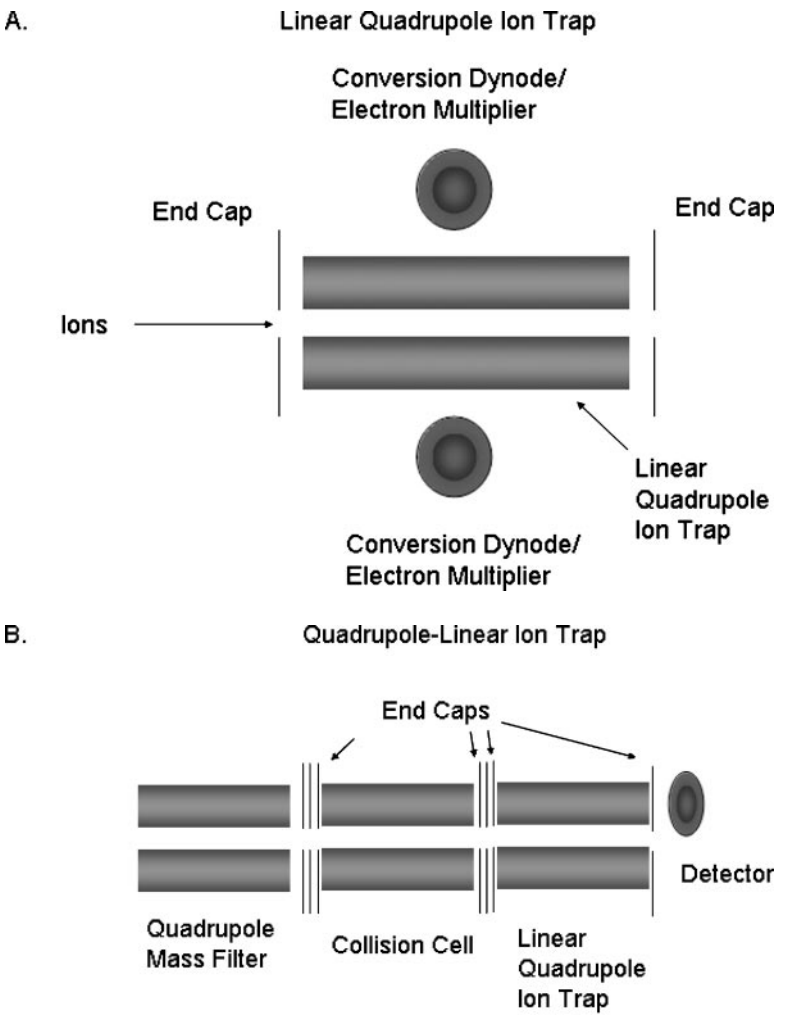


**Figure 1** (*A*) Schematic of a linear ion trap mass spectrometer. (*B*) Schematic of a quadrupole mass filter linear ion trap mass spectrometer. (*C*) Schematic of a TOF/TOF mass spectrometer. (*D*) Schematic of a linear ion trap–Fourier transform ion cyclotron resonance mass spectrometer.

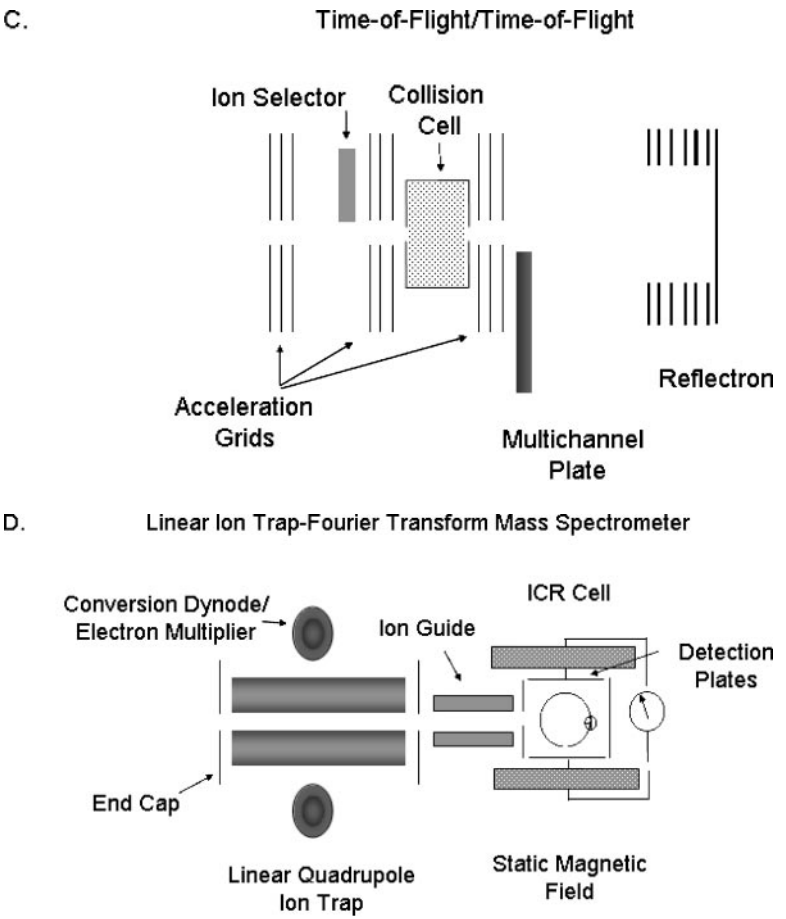


**Figure 1** (*Continued*)

significantly improves ion statistics during mass analysis. Ions are injected into the linear trap through an end cap and then ejected from it (59). Doing so allows the use of two detectors, since ions exit equally through the sides of the trap. Collecting ions with two detectors doubles the ion current collected during a scan of the m/z range. A second feature of the linear ion trap is the ability to scan at much faster speeds (15,000 AMU/s versus 5500 AMU/s), which increases the number of scans that can be acquired over the course of an LC analysis. Linear ion traps have limits to the mass resolution or accuracy that can be obtained. At normal scan speeds unit resolution (∼2000) is obtained, but slowing the scan speed can yield much higher resolutions (15,000 resolution over a 10 AMU window). As scan speeds are decreased, the mass range has to be reduced to minimize space charging—a phenomenon resulting when ions of like charge are forced closely together,

resulting in a perturbation of ideal ion motion in the electric fields. Tandem mass spectrometry (MS/MS) experiments are performed in this device by separating m/z measurement in time rather than space from ion isolation and collision-induced dissociation. MS/MS experiments benefit from better ion statistics of the linear ion trap as well as increased scan speed. Consequently, more data can be acquired at better quality over a three-dimensional ion trap, but the resolution and mass accuracy measurements of the ion trap are still limited.



QUADRUPOLE-QUADRUPOLE ION TRAP A novel use of the linear ion trap is to combine it with two quadrupoles to create a configuration similar to a triple quadrupole (37). When quadrupoles are combined with an ion trap, ions can be isolated and fragmented outside the ion trap and then accumulated in the trap for analysis of the fragment ions (Figure 1*B*). Additionally, ions can be simply passed through the mass filters and accumulated in the linear ion trap for analysis. A further benefit to this type of configuration is the ability to use alternative scan modes for the analysis of ions (37). For example, precursor ion scans can be used to identify peptide ions that produce a specific fragment ion m/z value. These types of scans have been used to identify phosphopeptides and glycopeptides (4, 5). Although these scans are useful to identify phosphopeptides, they require measurement of negative ions to identify phosphate-containing ions and then a second experiment is required to perform MS/MS on the positive ions to determine the sequence of the peptide and the site of modification. Phosphopeptides create strong ion signals as negative ions, but they don't always create strong signals as positive ions. The strong anionic nature of the phosphate decreases ionization efficiency of phosphopeptides such that the signal observed for phosphopeptides is much lower than that observed for nonphosphorylated peptides. These types of experiments are typically performed using triple quadrupole mass spectrometers, so why use a linear ion trap as the third mass analyzer? Ion traps typically show at least a 10-fold-greater improvement in sensitivity of analysis than beam instruments by trapping the ions and measuring each trapped m/z value in turn. The radio frequency–only quadrupole, which doubles as a collision cell, can be used to collect ions prior to injection into the linear ion trap used as a mass analyzer (26). Thus, the expectation is that a large increase in sensitivity should be observed. An added advantage to the linear ion trap is that product ion spectra can be acquired as a function of the internal energy of the fragment ions by monitoring the stability of the ions as a function of time (25). The result is a simplification of the tandem mass spectra for peptides.

TIME-OF-FLIGHT/TIME-OF-FLIGHT MASS SPECTROMETRY A type of mass analyzer that has been effective in proteomics is the time-of-flight (TOF) mass spectrometer. A mass spectrum is measured by determining the flight time of ions down a field-free flight tube. The time-of-flight of ions is related to their m/z values and thus a mass spectrum can be acquired. The major limitation to a TOF mass analyzer is the inability to perform true MS/MS. Post-source decay and prompt fragmentation of ions have been observed, but when used with the standard TOF instrument,

the ability to select ions for the analysis is severely limited and fragmentation of peptides using post source decay (PSD) or prompt fragmentation is not general (3, 66). Medzihradszky et al. (45) developed a tandem TOF mass spectrometer to use the high-speed capabilities of the mass analyzer to create a high-throughput tandem mass spectrometer (Figure 1*C*). The TOF/TOF instrument was equipped with a MALDI source and a high-speed laser to create ions (2). Consequently, m/z measurement is fast. A TOF mass analyzer is used in the ion selection process, and the selected ions are then transferred into a collision cell. Analysis of product ions occurs in a second TOF mass spectrometer. When performing MS/MS, however, the TOF/TOF does not show a dramatic increase in speed relative to other types of mass spectrometers (e.g., ion trap mass spectrometer) unless there is a large quantity of sample to generate a large ion signal. When sample quantities are high, ions statistics for MALDI are better and thus the instrument does not need to accumulate signal for long periods to improve signal to noise. The product ion data for peptides, however, can be used for de novo sequencing despite the increased complexity of fragmentation observed under high-energy collision conditions (82). The use of a MALDI source to create ions for the TOF/TOF can allow separation of the peptide fractionation process from the mass analysis, creating two separate workflows. In theory this could allow several separation stations to create samples for one mass spectrometer, but this only improves efficiency if the mass spectrometry analysis of each fractionated sample is faster than the separation.

FOURIER TRANSFORM ION CYCLOTRON MASS SPECTROMETRY Two important parameters in the measurement of ions by mass spectrometry are mass resolution ($M/\Delta M$) and mass measurement accuracy. Mass resolution is useful to determine ion charge states and isotopic distributions and to resolve isobaric ion (ions of the same nominal mass) species. Mass resolution is also a critical factor in mass measurement accuracy, as the narrower the ion envelope the more accurately the peak centroid can be determined. A type of mass spectrometer with the greatest capability for mass resolution and mass measurement accuracy is the Fourier transform ion cyclotron resonance mass spectrometer (FTMS) (43). These devices have been around for decades but have been experiencing a renaissance in the past few years. FTMS operates by trapping ions in a cell within a static magnetic field (Figure 1*D*). Under the influence of the field, ions assume a cyclotronic motion in the cell about the z direction of the magnetic field. m/z can be determined by measuring the frequency of motion of the ions. A breakthrough in ion cyclotron resonance (ICR) technology occurred when Comisarow and colleagues (12, 43) developed Fourier methodology to determine simultaneously the frequencies of all ions. An electronic pulse is used to perturb the motion of ions in the cell into coherence and then an electrical signal is measured as the ion motion decays to incoherence. Coherent motion of the ions produces an electrical signal measured between detection plates, whereas incoherence produces none. The time domain signal is converted to a frequency domain by Fourier transform, and m/z values are

derived from frequency of motion. Frequency can be measured with more accuracy and precision than any other physical property.

Initial advances in FTMS occurred when the ion source was removed from the ICR cell, allowing the use of ion sources such as fast atom bombardment (29). Because the ICR cell requires a high vacuum for the long flight paths of the ions, separation of the ion source and ICR cell allowed the use of ionization techniques employing higher pressures or frequent breaking of vacuum to introduce samples. More sophisticated ion guides and external ion traps were added to instruments to better control the movement and injection of ions into the ICR cell. Initial work with FTMS attempted to perform the MS/MS experiment within the ICR cell, a process involving ion isolation, fragmentation, and fragment ion measurement. Many technical and physical limitations prevented this strategy from achieving routine practicality, although nice examples have emerged using electron capture dissociation (ECD) to fragment intact proteins in the FTMS (28, 83). New efforts have decoupled the MS/MS experiment from FTMS ion measurement to allow the use of more efficient methods to isolate and fragment ions outside the FTMS. In one case, two quadrupole mass filters have been used to isolate and fragment ions, in which the ions are then collected in the final stage and injected into the ICR cell. A second strategy is to use a linear ion trap outside the FTMS to collect ions and then isolate and fragment the ions for injection into the ICR cell (75). Inside the linear ion trap, ions can be counted and the proper number of ions can be injected into the ICR cell. Overfilling the ICR cell results in space charging, in which ions are repelled from one another, thereby distorting their natural motion within the magnetic field, and underfilling the ICR cell results in poor signal to noise. These changes may result in an instrument that could create a high impact in proteomics.

The high mass accuracy and resolution of the FTMS has been used in several ways for peptide and protein analysis. By using shotgun proteomic strategies, the high mass accuracy of peptide ion measurements has been used as a method to identify the sequence of peptides through database searches. With high accuracy measurement, the molecular weights of peptide ions can be unique in a database and thus be an identifying feature for a peptide sequence (13). In general, large peptides that have been specifically cleaved are likely to be unique in a database. This strategy has been used to profile the proteins expressed in *Deinococcus radiodurans* strains under a variety of growth conditions (40, 64). A concern with this approach is that there is no convenient method to determine a peptide's cleavage specificity from the m/z value of the ion, and as such the cleavage specificity is simply assumed to match that of the enzyme used.

The analysis of intact proteins generally requires high-resolution mass measurements to resolve highly charged ions and their isotopes. Although accurate measurement of protein molecular weights can provide valuable information about a protein's sequence, such a measurement is frequently insufficient to identify a protein and additional information such as ion fragmentation is required (48, 50, 62). Horn et al. (28) have pioneered the use of ECD for protein dissociation,

and this method shows promise as a fragmentation method to identify proteins. Once a protein has been identified, the molecular weight measurement becomes extremely valuable, as it can help define posttranslational modifications in the protein (33). Kelleher et al. (33) used stable isotope labeling together with fragmentation to identify protein modifications. Unless fragmentation information has been obtained around sites of modification, it is difficult to identify the exact site of modification. A second challenge for the analysis of intact proteins is separation and delivery of the proteins to the mass spectrometer. Kelleher and colleagues (47) have used off-line fractionation of the proteins to reduce the complexity of the mixtures and deliver simplified mixtures of proteins to the instrument for analysis. Despite the challenges, analysis of intact proteins is important to understand the role of modified alternatively spliced forms of proteins in biological systems (35).

## SEPARATIONS OF COMPLEX PEPTIDE AND PROTEIN MIXTURES

A key element of proteomic analysis is the separation of peptides and proteins. Separation provides a method to simplify complex mixtures and also serves as a method to deliver molecules to the ionization source. This section discusses only chromatographic methods of protein and peptide separation, since gel electrophoretic methods for this purpose are well established.

Peptide separation techniques have come into new prominence as "shotgun proteomics" methods have gained in popularity. Shotgun proteomics involves digestion of a protein mixture into a collection of peptides, which is then separated on-line with a tandem mass spectrometer. Tandem mass spectra are collected for as many peptides as possible, and the results are then used to search through databases to identify the proteins in the original mixture. In this type of analysis the resolution of the separation is critical to the success of the analysis.

### Chromatographic Separations

Shen et al. (61, 63) have explored the use of long columns with small diameters to improve both separation and sensitivity of proteomic separation. Precolumns were also used to improve loading and resolution of separations. One-dimensional separations can be effective for the resolution of simple protein mixtures when MS/MS is used to analyze the mixtures. Smith et al. (64) used accurate mass measurements for the analysis of peptides; in this case the separation requirements are not as severe as those required for an MS/MS analysis, but sufficient resolution is still required to achieve a high dynamic range for the separation. When MS/MS is used for more complex mixtures to identify peptides, additional separation resolution is required. Multidimensional separation techniques have been used to resolve increasingly complex mixtures (49). Link et al. (39) demonstrated an integrated multidimensional separation technique to separate components of

protein complexes. The method relies on an ion exchange separation of peptides on a strong cation exchange (SCX) resin to fractionate on the basis of a charge in a batch elution mode, and then a reversed-phase (RP) separation of each fraction is collected directly off the SCX column. This approach has been used for mixtures as complex as digested yeast cell lysates. Off-line multidimensional separations have also been used for complex mixtures in which fractions are collected from an SCX separation and then each fraction is subjected to LC/MS/MS (24). For ease of analysis, the fractions are often delivered by an autosampler, although this method does not seem to offer any analytical advantages over the integrated multidimensional approach (51).

Affinity methods have also become important approaches to selectively fractionate peptide mixtures. Methods have been developed to isolate peptides on the basis of the presence of specific amino acids in the sequence or on the presence of specific posttranslational modifications. Metal-affinity chromatography (IMAC) used in combination with Cu(II) has been used to enrich peptides containing His residues (21, 30). When IMAC is combined with an Fe(III) ligand, phosphopeptides can be enriched (18). Glycosylated peptides have been isolated with lectin chromatography, and this method can be used to study glycosylation patterns for proteins (21). These methods provide a mechanism to enrich samples with specific structural features, but add additional steps that can limit dynamic range rather than extend it through losses of sample to nonspecific absorption to surfaces.

Protein separations have been achieved by combining chromatofocusing with RP chromatography using a $C_4$ or $C_8$ column (74). Chromatofocusing is a form of chromatographic isoelectric focusing that uses an ion exchange resin and a pH gradient to elute proteins by their isoelectric point. Chong et al. (11) used this approach as an initial fractionation method followed by RP separations of each fraction. Fractions can be collected or directly introduced into the mass spectrometer. Separating proteins without the use of chaotropes or detergents to maintain protein solubility can bias the types of proteins observed, and thus more effort is needed to discover which types of proteins are missed by this procedure.

## Capillary Electrophoresis Separations

Capillary electrophoresis (CE) is an alternative separation technique that relies on the movement of molecules in a fused silica capillary. Although the separation of peptides by CE has significant promise, it has yet to develop into a standard approach for proteomic analyses. The best conditions for the separation of peptides by CE are not the best conditions for ionization of peptides by ESI. To circumvent this problem, Preisler and colleagues (55, 56) used off-line deposition of CE effluent onto a MALDI plate to decouple the separation and ionization processes. When vacuum deposition of the effluent is used, thin lines of separated peptides can be deposited. Excellent sensitivity can be achieved when interfaced with MALDI, but sampling times can become an issue requiring the use of faster lasers.

Multidimensional separations are also possible with CE. Tong et al. (72) demonstrated a combined RP-CE separation for a ribosome protein complex. Peptides

were step eluted from the RP packing material inserted at the head of the column, followed by a CE separation performed at lower voltage to slow the speed of separation. Chen et al. (10) combined capillary isoelectric focusing (CIEF) with capillary RP chromatography to separate complex peptides mixtures. Peptides are resolved by their isoelectric points and then moved into a RP column for separation. A dialysis device in between the two columns is used to change the pH of the solution and remove ampholytes used in the CIEF step (81). Multidimensional separations using CE have the potential to add different separation modes to resolve complex mixtures but still require optimization to create robust and practical separation techniques.

CIEF to fractionate proteins directly into a mass spectrometer has potential to create a separation system similar to two-dimensional gel electrophoresis. Yang et al. (80) demonstrated the separation of *Escherichia coli* proteins using CIEF, followed by molecular weight analysis using mass spectrometry. The advantage to this approach is the ability to accurately measure the molecular weights of intact proteins, but the value of the data was limited because it is difficult to identify to the proteins without additional experimentation (as described above). As the mass spectrometry experiments for protein identification become more general, this approach will prove significant for the analysis of intact proteins.

## SHOTGUN PROTEOMICS OF PROTEINS

The methods described above have been combined in various forms to create a variety of methods for proteomic analysis. Proteomic analysis has for many years relied upon the use of gel electrophoresis for protein separation, but over the past few years chromatographic methods have played an increasing role in proteomics and this area is briefly covered. Two approaches have emerged from these developments: shotgun proteomics, also referred to as bottom-up proteomics, and top-down proteomics.

Shotgun proteomics relies on the digestion of protein mixtures followed by separation of the peptides and subsequent introduction into a tandem mass spectrometer. Tandem mass spectra are automatically collected for as many peptides as possible and then those spectra are used to search protein sequence databases. These approaches are well suited to the analysis of protein complexes, which comprise a discrete set of proteins with a functional relationship, and shotgun proteomics provides a sensitive technique to identify the components of complexes. As an example, Cheeseman et al. (9) isolated components of the yeast kinetochore to identify the subcomplexes. By using a method developed by MacCoss et al. (41), Cheeseman et al. (9) were able to identify not only the proteins in the complexes but also the phosphorylation sites. When site-directed mutagenesis was used, the Ipl1 kinase, which associates with the kinetochore, was disrupted to identify which sites in the complex were targets of this kinase. The analysis was done using an LC/LC/MS/MS method to ensure peptide sequences were identified along much of the protein sequences. By acquiring tandem mass spectra from peptides along

much of the protein's sequence, it can be determined if any of the residues are modified.

The LC/LC/MS/MS approach for shotgun proteomics has been used to identify proteins in more complex mixtures. Florens et al. (19) used this approach to identify proteins expressed at various stages of development in *Plasmodium falciparum*, the pathogen that causes malaria. Proteins from four different developmental stages of this organism were identified, including a separation of membrane from soluble proteins. Coordinate stage-specific expression was observed for a set of 138 proteins, suggesting mechanisms for regulation of protein transcription. A variation on this approach has been used with a selective reagent to isolate peptides containing specific types of amino acids (23). Protein expression changes resulting from a shift in carbon source in the growth media from glucose to galactose in *Saccharomyces cerevisiae* and those changes in proteins localized to microsomes in mammalian cells after chemical induction have been identified. Shotgun methods are useful for the identification of hard-to-analyze classes of proteins such as membrane proteins. Wu et al. (77) demonstrated the identification of membrane proteins from rat brain tissues, and Blonder et al. (2a) demonstrated a preparation procedure for membrane proteins, employing an organic solvent to solubilize the proteins. Recently, Schirmer et al. (58) identified proteins from the nuclear envelope with potential links to diseases such as Charcot-Marie dystrophies. An advantage to the shotgun method is that it is less biased toward membrane proteins, as solubilizing proteins becomes less of an issue.

A related method for the analysis of proteomes involves the use of accurate mass measurement to identify peptides. In a method developed by Conrads et al. (13), a shotgun proteomics method is initially employed using LC/MS/MS to identify peptides. An accurate mass measurement is then used to confirm the identification and to produce an accurate mass tag for the peptide. This tag combines both the accurate mass measurement and the peptide's retention time during RP chromatography (40).

## TOP-DOWN ANALYSIS OF PROTEINS

Although shotgun proteomics is an important strategy for the identification of proteins in mixtures, this method is insensitive to the isoforms of a protein. Proteins can be modified in different manners, including alternative splicing. An interesting finding of the genome project was the smaller than expected number of genes, suggesting the importance of covalent modifications and alternative splicing as methods to create functional diversity. New methods are needed for high-throughput and accurate measurement and identification of intact proteins. Top-down mass spectrometry approaches analyze intact proteins to identify both the protein and any structural modification.

A key element to top-down proteomics is the ability to fragment intact proteins in the mass spectrometer. Two strategies have shown promise for this purpose. The first strategy uses ECD to fragment proteins. Low-energy electrons are fired at ions captured in the ICR cell to cause their fragmentation (44). The second strategy

uses charge state reduction through ion-ion reactions and then collision-induced dissociation of a selected charge state to cause fragmentation of the ion (67). Fragmentation by ECD appears to be more complete, allowing the identification of protein modifications and amino acid sequence variations.

Meng et al. (46) have used the top-down approach to study small proteins from *Methanococcus jannaschii*. The protein mixtures were separated off-line and then introduced into an FTMS mass spectrometer by infusion. When a protein mixture is infused into an ESI source, ion signals persist for many minutes, allowing more time for the mass spectrometry analysis. To facilitate the analysis, Meng et al. (46) have developed software to search protein fragmentation spectra through sequence databases. Before a top-down approach will have a broad impact in proteomics, the issue of separation and delivery of proteins to the mass spectrometer in a large-scale manner needs to be improved and automation of the necessary mass spectrometry steps needs to occur.



## PROTEOMIC INFORMATICS

The computer analysis of mass spectrometry data has been a key tool for large-scale and high-throughput protein analysis and has enabled proteomic style studies. A principal method in shotgun proteomics analysis of protein mixtures is the database searching of MS/MS data of peptides (15). Several strategies have been developed to measure the uniqueness of a tandem mass spectrum to amino acid sequences in the database. All methods are based on predicting the fragmentation pattern of the peptide and then comparing the pattern to the fragments in the tandem mass spectrum. Scoring of a match has been based on cross-correlation, fragment ion frequencies, and hypergeometric probability (15, 52). Intensity models have also been included in correlation analyses to improve matches between sequence and spectra. Probability-based methods provide a statistical measure for the fit between sequence and spectrum (42, 52, 57). The statistical relevance for quantitative comparisons can be determined through empirical analysis of known data sets and establishing scoring cutoffs for a statistical confidence level (1, 36). Most of these programs are suitable for the analysis of protein modifications. Large sets of data have a further complication of filtering and assembling the data to establish the set of proteins identified. Tabb et al. (69) developed a software program designed specifically for the analysis data obtained from LC/LC/MS/MS. Proteomic style analyses can generate large amounts of data and thus have a need for a high level of automation in the data analysis.

## CONCLUSION

As the importance of proteomics has increased, developments in mass spectrometry have strived to produce new capabilities and improve performance. Developments have occurred in two areas: improved mass accuracy and resolution, and speed. As these developments are implemented into existing proteomic strategies and

new strategies are developed, capabilities for proteomics should improve. Two other elements of equal importance for proteomics are the ability to separate and deliver peptides and proteins to the mass spectrometer, and suitable software to analyze the mass spectrometry data. Separation methods based on multidimensional LC are likely to improve, providing more resolution, sensitivity, and dynamic range. Furthermore, integrated multidimensional separations for proteins will likely develop in the near future, superseding the off-line methods currently used.



## ACKNOWLEDGMENTS

This work was supported by NIH grants RR11823-08 and 5 R01 MH067880.

**The *Annual Review of Biophysics and Biomolecular Structure* is online at http://biophys.annualreviews.org**

## LITERATURE CITED

1. Anderson DC, Li W, Payan DG, Noble WS. 2003. A new algorithm for the evaluation of shotgun peptide sequencing in proteomics: support vector machine classification of peptide MS/MS spectra and SEQUEST scores. *J. Proteome Res.* 2:137–46
2. Bienvenut WV, Deon C, Pasquarello C, Campbell JM, Sanchez JC, et al. 2002. Matrix-assisted laser desorption/ionization-tandem mass spectrometry with high resolution and sensitivity for identification and characterization of proteins. *Proteomics* 2:868–76

2a. Blonder J, Goshe MB, Moore RJ, Pasa-Tolic L, Masselon CD, et al. 2002. Enrichment of integral membrane proteins for proteomic analysis using liquid chromatography-tandem mass spectrometry. *J. Proteome Res.* 1: 351–60

3. Brown RS, Lennon JJ. 1995. Sequence-specific fragmentation of matrix-assisted laser-desorbed protein/peptide ions. *Anal. Chem.* 67:3990–99
4. Carr SA, Huddleston MJ, Annan RS. 1996. Selective detection and sequencing of phosphopeptides at the femtomole level by mass spectrometry. *Anal. Biochem.* 239:180–92
5. Carr SA, Huddleston MJ, Bean MF. 1993. Selective identification and differentiation of N- and O-linked oligosaccharides in glycoproteins by liquid chromatography-mass spectrometry. *Protein Sci.* 2:183–96
6. Cha B, Blades M, Douglas DJ. 2000. An interface with a linear quadrupole ion guide for an electrospray-ion trap mass spectrometer system. *Anal. Chem.* 72: 5647–54
7. Chaurand P, Caprioli RM. 2002. Direct profiling and imaging of peptides and proteins from mammalian cells and tissue sections by mass spectrometry. *Electrophoresis* 23:3125–35
8. Chaurand P, Schwartz SA, Caprioli RM. 2002. Imaging mass spectrometry: a new tool to investigate the spatial organization of peptides and proteins in mammalian tissue sections. *Curr. Opin. Chem. Biol.* 6:676–81
9. Cheeseman IM, Anderson S, Jwa M, Green EM, Kang J, et al. 2002. Phospho-regulation of kinetochore-microtubule attachments by the Aurora kinase Ipl1p. *Cell* 111:163–72
10. Chen J, Lee CS, Shen Y, Smith RD, Baehrecke EH. 2002. Integration of

capillary isoelectric focusing with capillary reversed-phase liquid chromatography for two-dimensional proteomics separation. *Electrophoresis* 23:3143–48

11. Chong BE, Hamler RL, Lubman DM, Ethier SP, Rosenspire AJ, Miller FR. 2001. Differential screening and mass mapping of proteins from premalignant and cancer cell lines using nonporous reversed-phase HPLC coupled with mass spectrometric analysis. *Anal. Chem.* 73: 1219–27
12. Comisarow MB, Marshall AG. 1974. Fourier transform ion cyclotron resonance mass spectrometry. *Chem. Phys. Lett.* 25:282–83
13. Conrads TP, Anderson GA, Veenstra TD, Pasa-Tolic L, Smith RD. 2000. Utility of accurate mass tags for proteome-wide protein identification. *Anal. Chem.* 72:3349–57
14. Emmett MR, Caprioli RM. 1994. Micro-electrospray mass spectrometry: ultra-high-sensitivity analysis of peptides and proteins. *J. Am. Soc. Mass Spectrom.* 5:605–13
15. Eng JK, McCormack AL, Yates III JR. 1994. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* 5:976–89
16. Ericson C, Phung QT, Horn DM, Peters EC, Fitchett JR, et al. 2003. An automated noncontact deposition interface for liquid chromatography matrix-assisted laser desorption/ionization mass spectrometry. *Anal. Chem.* 75:2309–15
17. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. 1989. Electrospray ionization for mass spectrometry of large biomolecules. *Science* 246:64–71
18. Ficarro SB, McCleland ML, Stukenberg PT, Burke DJ, Ross MM, et al. 2002. Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. *Nat. Biotechnol.* 20:301–5
19. Florens L, Washburn MP, Raine JD, Anthony RM, Grainger M, et al. 2002. A proteomic view of the Plasmodium falciparum life cycle. *Nature* 419:520–26
20. Garden RW, Moroz LL, Moroz TP, Shippy SA, Sweedler JV. 1996. Excess salt removal with matrix rinsing: direct peptide profiling of neurons from marine invertebrates using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *J. Mass Spectrom.* 31:1126–30
21. Geng M, Ji J, Regnier FE. 2000. Signature-peptide approach to detecting proteins in complex mixtures. *J. Chromatogr. A* 870:295–313
22. Griffin TJ, Han DK, Gygi SP, Rist B, Lee H, et al. 2001. Toward a high-throughput approach to quantitative proteomic analysis: expression-dependent protein identification by mass spectrometry. *J. Am. Soc. Mass Spectrom.* 12:1238–46
23. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. 1999. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* 17:994–99
24. Gygi SP, Rist B, Griffin TJ, Eng J, Aebersold R. 2002. Proteome analysis of low-abundance proteins using multidimensional chromatography and isotope-coded affinity tags. *J. Proteome Res.* 1:47–54
25. Hager JW. 2003. Product ion spectral simplification using time-delayed fragment ion capture with tandem linear ion traps. *Rapid Commun. Mass Spectrom.* 17:1389–98
26. Hager JW, Yves Le Blanc JC. 2003. Product ion scanning using a Q-q-Q linear ion trap (Q TRAP) mass spectrometer. *Rapid Commun. Mass Spectrom.* 17:1056–64
27. Hillenkamp F, Karas M, Beavis RC, Chait BT. 1991. Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. *Anal. Chem.* 63:1193A–203A
28. Horn DM, Ge Y, McLafferty FW. 2000. Activated ion electron capture dissociation for mass spectral sequencing of larger (42 kDa) proteins. *Anal. Chem.* 72:4778–84

29. Hunt DF, Shabanowitz J, McIver RT Jr, Hunter RL, Syka JE. 1985. Ionization and mass analysis of nonvolatile compounds by particle bombardment tandem-quadrupole Fourier transform mass spectrometry. *Anal. Chem.* 57:765–68
30. Ji J, Chakraborty A, Geng M, Zhang X, Amini A, et al. 2000. Strategy for qualitative and quantitative analysis in proteomics based on signature peptides. *J. Chromatogr. B Biomed. Sci. Appl.* 745: 197–210
31. Jonscher KR, Yates I, John R. 1997. The quadrupole ion trap mass spectrometer—a small solution to a big challenge. *Anal. Biochem.* 244:1–15
32. Kast J, Gentzel M, Wilm M, Richardson K. 2003. Noise filtering techniques for electrospray quadrupole time of flight mass spectra. *J. Am. Soc. Mass Spectrom.* 14:766–76
33. Kelleher NL, Nicewonger RB, Begley TP, McLafferty FW. 1997. Identification of modification sites in large biomolecules by stable isotope labeling and tandem high resolution mass spectrometry. The active site nucleophile of thiaminase I. *J. Biol. Chem.* 272:32215–20
34. Deleted in proof
35. Kelleher NL, Taylor SV, Grannis D, Kinsland C, Chiu HJ, et al. 1998. Efficient sequence analysis of the six gene products (7–74 kDa) from the Escherichia coli thiamin biosynthetic operon by tandem high-resolution mass spectrometry. *Protein Sci.* 7:1796–801
36. Keller A, Nesvizhskii AI, Kolker E, Aebersold R. 2002. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* 74:5383–92
37. Le Blanc JC, Hager JW, Ilisiu AM, Hunter C, Zhong F, Chu I. 2003. Unique scanning capabilities of a new hybrid linear ion trap mass spectrometer (Q TRAP) used for high sensitivity proteomics applications. *Proteomics* 3:859–69
38. Li L, Garden RW, Sweedler JV. 2000. Single-cell MALDI: a new tool for direct peptide profiling. *Trends Biotechnol.* 18:151–60
39. Link AJ, Eng J, Schieltz DM, Carmack E, Mize GJ, et al. 1999. Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol.* 17:676–82
40. Lipton MS, Pasa-Tolic L, Anderson GA, Anderson DJ, Auberry DL, et al. 2002. Global analysis of the Deinococcus radiodurans proteome by using accurate mass tags. *Proc. Natl. Acad. Sci. USA* 99: 11049–54
41. MacCoss MJ, McDonald WH, Saraf A, Sadygov R, Clark JM, et al. 2002. Shotgun identification of protein modifications from protein complexes and lens tissue. *Proc. Natl. Acad. Sci. USA* 99:7900–6
42. MacCoss MJ, Wu CC, Yates JR 3rd. 2002. Probability-based validation of protein identifications using a modified SEQUEST algorithm. *Anal. Chem.* 74:5593–99
43. Marshall AG, Hendrickson CL, Jackson GS. 1998. Fourier transform ion cyclotron resonance mass spectrometry: a primer. *Mass Spectrom. Rev.* 17:135
44. McLafferty FW, Horn DM, Breuker K, Ge Y, Lewis MA, et al. 2001. Electron capture dissociation of gaseous multiply charged ions by Fourier-transform ion cyclotron resonance. *J. Am. Soc. Mass Spectrom.* 12: 245–49
45. Medzihradszky KF, Campbell JM, Baldwin MA, Falick AM, Juhasz P, et al. 2000. The characteristics of peptide collision-induced dissociation using a high-performance MALDI-TOF/TOF tandem mass spectrometer. *Anal. Chem.* 72:552–58
46. Meng F, Cargile BJ, Miller LM, Forbes AJ, Johnson JR, Kelleher NL. 2001. Informatics and multiplexing of intact protein identification in bacteria and the archaea. *Nat. Biotechnol.* 19:952–57
47. Meng F, Cargile BJ, Patrie SM, Johnson JR, McLaughlin SM, Kelleher NL. 2002. Processing complex mixtures of

intact proteins for direct analysis by mass spectrometry. *Anal. Chem.* 74:2923–29

48. Mortz E, O'Connor PB, Roepstorff P, Kelleher NL, Wood TD, et al. 1996. Sequence tag identification of intact proteins by matching tandem mass spectral data against sequence data bases. *Proc. Natl. Acad. Sci. USA* 93:8264–67
49. Opiteck GJ, Jorgenson JW, Anderegg RJ. 1997. Two-dimensional SEC/RPLC coupled to mass spectrometry for the analysis of peptides. *Anal. Chem.* 69:2283–91
50. Paša-Tolic L, Jensen PK, Anderson GA, Lipton MS, Peden KK, et al. 1999. High throughput proteome-wide precision measurements of protein expression using mass spectrometry. *J. Am. Chem. Soc.* 121:7949–50
51. Peng J, Elias JE, Thoreen CC, Licklider LJ, Gygi SP. 2003. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J. Proteome Res.* 2:43–50
52. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20:3551–67
53. Petricoin EF, Ardekani AM, Hitt BA, Levine PJ, Fusaro VA, et al. 2002. Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* 359:572–77
54. Petricoin EF 3rd, Ornstein DK, Paweletz CP, Ardekani A, Hackett PS, et al. 2002. Serum proteomic patterns for detection of prostate cancer. *J. Natl. Cancer Inst.* 94:1576–78
55. Preisler J, Hu P, Rejtar T, Moskovets E, Karger BL. 2002. Capillary array electrophoresis-MALDI mass spectrometry using a vacuum deposition interface. *Anal. Chem.* 74:17–25
56. Rejtar T, Hu P, Juhasz P, Campbell JM, Vestal ML, et al. 2002. Off-line coupling of high-resolution capillary electrophoresis to MALDI-TOF and TOF/TOF MS. *J. Proteome Res.* 1:171–79
57. Sadygov R, Yates JR. 2003. A hypergeometric probability model for protein identification and validation using tandem mass spectral data and protein sequence databases. *Anal. Chem.* 75:3792–98
58. Schirmer EC, Florens L, Guan T, Yates JR 3rd, Gerace L. 2003. Nuclear membrane proteins with potential disease links found by subtractive proteomics. *Science* 301:1380–82
59. Schwartz JC, Senko MW, Syka JE. 2002. A two-dimensional quadrupole ion trap mass spectrometer. *J. Am. Soc. Mass Spectrom.* 13:659–69
60. Schwartz SA, Reyzer ML, Caprioli RM. 2003. Direct tissue analysis using matrix-assisted laser desorption/ionization mass spectrometry: practical aspects of sample preparation. *J. Mass Spectrom.* 38:699–708
61. Shen Y, Smith RD. 2002. Proteomics based on high-efficiency capillary separations. *Electrophoresis* 23:3106–24
62. Shen Y, Xiang F, Veenstra TD, Fung EN, Smith RD. 1999. High-resolution capillary isoelectric focusing of complex protein mixtures from lysates of microorganisms. *Anal. Chem.* 71:5348–53
63. Shen Y, Zhao R, Berger SJ, Anderson GA, Rodriguez N, Smith RD. 2002. High-efficiency nanoscale liquid chromatography coupled on-line with mass spectrometry using nanoelectrospray ionization for proteomics. *Anal. Chem.* 74:4235–49
64. Smith RD, Anderson GA, Lipton MS, Pasa-Tolic L, Shen Y, et al. 2002. An accurate mass tag strategy for quantitative and high-throughput proteome measurements. *Proteomics* 2:513–23
65. Smith RD, Loo JA, Edmonds CG, Barinaga CJ, Udseth HR. 1990. Sensitivity considerations for large molecule detection by capillary electrophoresis-electrospray ionization mass spectrometry. *J. Chromatogr.* 516:157–65
66. Spengler B. 1997. Post-source decay

analysis in matrix-assisted laser desorption/ionization mass spectrometry of biomolecules. *J. Mass Spectrom.* 32: 1019–36

67. Stephenson JL, McLuckey SA, Reid GE, Wells JM, Bundy JL. 2002. Ion/ion chemistry as a top-down approach for protein analysis. *Curr. Opin. Biotechnol.* 13:57–64
68. Stoeckli M, Chaurand P, Hallahan DE, Caprioli RM. 2001. Imaging mass spectrometry: a new technology for the analysis of protein expression in mammalian tissues. *Nat. Med.* 7:493–96
69. Tabb DL, McDonald HW, Yates JR III. 2002. DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. *J. Proteome Res.* 1:21–26
70. Tanaka K. 2003. The origin of macromolecule ionization by laser irradiation (Nobel lecture). *Angew. Chem. Int. Ed. Engl.* 42:3860–70
71. Todd PJ, Schaaff TG, Chaurand P, Caprioli RM. 2001. Organic ion imaging of biological tissue with secondary ion mass spectrometry and matrix-assisted laser desorption/ionization. *J. Mass Spectrom.* 36:355–69
72. Tong W, Link A, Eng JK, Yates JR 3rd. 1999. Identification of proteins in complexes by solid-phase microextraction/multistep elution/capillary electrophoresis/tandem mass spectrometry. *Anal. Chem.* 71:2270–78
73. Walker KL, Chiu RW, Monnig CA, Wilkins CL. 1995. Off-line coupling of capillary electrophoresis and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Anal. Chem.* 67:4197–204
74. Wall DB, Kachman MT, Gong SS, Parus SJ, Long MW, Lubman DM. 2001. Isoelectric focusing nonporous silica reversed-phase high-performance liquid chromatography/electrospray ionization time-of-flight mass spectrometry: a three-dimensional liquid-phase protein separation method as applied to the human erythroleukemia cell-line. *Rapid Commun. Mass Spectrom.* 15:1649–61
75. Wilcox BE, Hendrickson CL, Marshall AG. 2002. Improved ion extraction from a linear octopole ion trap: SIMION analysis and experimental demonstration. *J. Am. Soc. Mass Spectrom.* 13:1304–12
76. Wilm MS, Mann M. 1994. Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last? *Int. J. Mass Spectrom. Ion Process.* 136:2–3
77. Wu CC, MacCoss MJ, Howell KE, Yates JR. 2003. A method for the comprehensive proteomic analysis of membrane proteins. *Nat. Biotechnol.* 21:532–38
78. Xu BJ, Caprioli RM, Sanders ME, Jensen RA. 2002. Direct analysis of laser capture microdissected cells by MALDI mass spectrometry. *J. Am. Soc. Mass Spectrom.* 13:1292–97
79. Yanagisawa K, Shyr Y, Xu BJ, Massion PP, Larsen PH, et al. 2003. Proteomic patterns of tumour subsets in non-small-cell lung cancer. *Lancet* 362:433–39
80. Yang L, Lee CS, Hofstadler SA, Pasa-Tolic L, Smith RD. 1998. Capillary isoelectric focusing-electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry for protein characterization. *Anal. Chem.* 70:3235–41
81. Yang L, Lee CS, Hofstadler SA, Smith RD. 1998. Characterization of microdialysis acidification for capillary isoelectric focusing-microelectrospray ionization mass spectrometry. *Anal. Chem.* 70:4945–50
82. Yergey AL, Coorssen JR, Backlund PS Jr, Blank PS, Humphrey GA, et al. 2002. De novo sequencing of peptides using MALDI/TOF-TOF. *J. Am. Soc. Mass Spectrom.* 13:784–91
83. Zubarev RA, Horn DM, Fridriksson EK, Kelleher NL, Kruger NA, et al. 2000. Electron capture dissociation for structural characterization of multiply charged protein cations. *Anal. Chem.* 72:563–73

*Annual Review of Biophysics and Biomolecular Structure*
*Volume 33, 2004*

# CONTENTS

ENZYME-MEDIATED DNA LOOPING, *Stephen E. Halford, Abigail J. Welsh, and Mark D. Szczelkun* 1

DISEASE-RELATED MISASSEMBLY OF MEMBRANE PROTEINS, *Charles R. Sanders and Jeffrey K. Myers* 25

CONFORMATIONAL SPREAD: THE PROPAGATION OF ALLOSTERIC STATES IN LARGE MULTIPROTEIN COMPLEXES, *Dennis Bray and Thomas Duke* 53

A FUNCTION-BASED FRAMEWORK FOR UNDERSTANDING BIOLOGICAL SYSTEMS, *Jeffrey D. Thomas, Taesik Lee, and Nam P. Suh* 75

STRUCTURE, MOLECULAR MECHANISMS, AND EVOLUTIONARY RELATIONSHIPS IN DNA TOPOISOMERASES, *Kevin D. Corbett and James M. Berger* 95

STRUCTURE, DYNAMICS, AND CATALYTIC FUNCTION OF DIHYDROFOLATE REDUCTASE, *Jason R. Schnell, H. Jane Dyson, and Peter E. Wright* 119

THREE-DIMENSIONAL ELECTRON MICROSCOPY AT MOLECULAR RESOLUTION, *Sriram Subramaniam and Jacqueline L.S. Milne* 141

TAKING X-RAY DIFFRACTION TO THE LIMIT: MACROMOLECULAR STRUCTURES FROM FEMTOSECOND X-RAY PULSES AND DIFFRACTION MICROSCOPY OF CELLS WITH SYNCHROTRON RADIATION, *Jianwei Miao, Henry N. Chapman, Janos Kirz, David Sayre, and Keith O. Hodgson* 157

MOLECULES OF THE BACTERIAL CYTOSKELETON, *Jan Löwe, Fusinita van den Ent, and Linda A. Amos* 177

TETHERING: FRAGMENT-BASED DRUG DISCOVERY, *Daniel A. Erlanson, James A. Wells, and Andrew C. Braisted* 199

THE USE OF IN VITRO PEPTIDE-LIBRARY SCREENS IN THE ANALYSIS OF PHOSPHOSERINE/THREONINE-BINDING DOMAIN STRUCTURE AND FUNCTION, *Michael B. Yaffe and Stephen J. Smerdon* 225

ROTATION OF $F_1$-ATPASE: HOW AN ATP-DRIVEN MOLECULAR MACHINE MAY WORK, *Kazuhiko Kinosita, Jr., Kengo Adachi, and Hiroyasu Itoh* 245

MODEL SYSTEMS, LIPID RAFTS, AND CELL MEMBRANES, *Kai Simons and Winchil L.C. Vaz* 269

MASS SPECTRAL ANALYSIS IN PROTEOMICS, *John R. Yates, III* 297

INFORMATION CONTENT AND COMPLEXITY IN THE HIGH-ORDER ORGANIZATION OF DNA, *Abraham Minsky* 317

THE ROLE OF WATER IN PROTEIN-DNA RECOGNITION, *B. Jayaram and Tarun Jain* 343

FORCE AS A USEFUL VARIABLE IN REACTIONS: UNFOLDING RNA, *Ignacio Tinoco, Jr.* 363

RESIDUAL DIPOLAR COUPLINGS IN NMR STRUCTURE ANALYSIS, *Rebecca S. Lipsitz and Nico Tjandra* 387

THE THERMODYNAMICS OF DNA STRUCTURAL MOTIFS, *John SantaLucia, Jr. and Donald Hicks* 415

SPIN DISTRIBUTION AND THE LOCATION OF PROTONS IN PARAMAGNETIC PROTEINS, *D. Goldfarb and D. Arieli* 441

INDEXES

Subject Index 469

Cumulative Index of Contributing Authors, Volumes 29–33 495

Cumulative Index of Chapter Titles, Volumes 29–33 498

ERRATA

An online log of corrections to *Annual Review of Biophysics and Biomolecular Structure* chapters may be found at http://biophys.annualreviews.org/errata.shtml