

Articles

An Introduction to Mass Spectrometry Applications in Biological Research*

Received for publication, November 4, 2003

Erin J. Finehout and Kelvin H. Lee‡

From the School of Chemical and Biomolecular Engineering, Cornell University, Ithaca, New York 14853

Mass spectrometry involves the measurement of the mass-to-charge ratio of ions. It has become an essential analytical tool in biological research and can be used to characterize a wide variety of biomolecules such as sugars, proteins, and oligonucleotides. In this review, a brief history of mass spectrometry is discussed, and the basic principles of the technology are introduced. A summary of some current applications is provided, as are examples of recently published research. The current methods used to identify, quantify, and characterize proteins and peptides are then reviewed. The range of applications of mass spectrometry is considerable and only promises to grow as the technology continues to improve.

Keywords: Mass Spectrometry, protein characterization, proteomics.

The definition of a mass spectrometer may seem simple: it is an instrument that can ionize a sample and measure the mass-to-charge ratio of the resulting ions. However, the versatility of this function has allowed it to become a vital tool in a wide range of fields, including biological research. This versatility arises from the fact that mass spectrometers can give qualitative and quantitative information on the elemental, isotopic, and molecular composition of organic and inorganic samples [1]. Furthermore, samples can be analyzed from the gas, liquid, or solid state, and the masses that can be studied range from single atoms (several Da) to proteins (over 300,000 Da) [2].

The application of mass spectrometry (MS)¹ to biology began in the 1940s, when heavy stable isotopes were used as tracers to study processes such as CO₂ production in animals [3]. Since that time, advances in technology have increased the range of sample types that can be ionized and the range of masses that can be measured, which has diversified the applications of MS. The biological applications of MS currently encompass such diverse areas as screening newborns for metabolic disorders [4], comparing protein expression levels between cells grown in different media [5], determining the bioavailability of minerals in food [6], and studying how pharmaceutical drugs are

metabolized *in vivo* [7]. As an example of the growing importance of MS, a recent search of PubMed (an archive of citations from life science journals [8]) for the phrase “mass spectrometry” resulted in over 65,900 total hits, with over 6,500 of these articles published in the year 2002 alone. The current importance of MS to biological research is highlighted by the 2002 Nobel Prize in Chemistry, which was awarded to John Fenn and Koichi Tanaka “for their development of soft desorption ionization methods for mass spectrometric analysis of biological macromolecules” [9].

The goal of this review is to familiarize the reader with MS and some of its current applications to biological research. The total number of applications is too large to cover in a single review, so we provide a general overview with an emphasis on protein characterization, one of the newest and most widespread applications of biological MS.

A BRIEF HISTORY OF THE BIOLOGICAL APPLICATIONS OF MS

J. J. Thompson constructed the first mass spectrometer in 1912. The early mass spectrometers were primarily used by physicists to study the atomic weight of the elements and the natural relative abundance of elemental isotopes [1]. Although these early instruments could not be used to study biomolecules, it was not long before they were used to study heavy isotopes as tracers in biological systems [3]. Improvements in the range of masses that could be analyzed and the types of sample that could be vaporized made possible the study of organic compounds in 1956. In 1959, MS was first used to sequence peptides and oligonucleotides, and in 1962 it was used to study the structure of nucleotides [10].

Ionization of larger molecules such as proteins was not possible until 1981, when the fast atom bombardment ionization method was introduced [11]. The ability to ionize

* E. J. F. is supported by a National Science Foundation Graduate Fellowship. We gratefully acknowledge support for this work by National Science Foundation Grant BES 0120315, U.S. Department of Agriculture (USDA) Grant SCA 58-1907-1-146, and the USDA Agricultural Research Service.

‡ To whom correspondence should be addressed. E-mail: kh19@cornell.edu.

¹ The abbreviations used are: MS, mass spectrometry; CID, collision-induced dissociation; ESI, electrospray ionization; ICAT, isotope-coded affinity tag; IRMS, isotope ratio mass spectrometry; LC, liquid chromatography; MALDI, matrix-assisted laser desorption/ionization; *m/z*, mass-to-charge; TOF, time-of-flight mass analyzer.

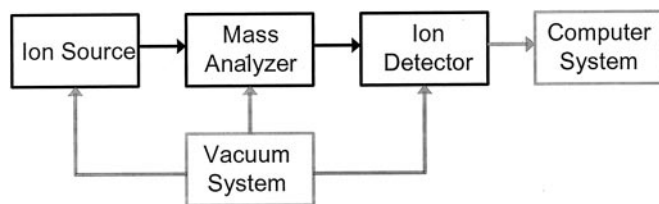


FIG. 1. Schematic of the basic components of a mass spectrometer.

larger molecules was further improved with the advent of electrospray ionization (ESI) by Fenn and coworkers in 1988 [12]. The electrospray ion source was easily connected to on-line liquid chromatography (LC), which made possible the analysis of complex mixtures. Time-of-flight (TOF) mass analysis had been developed and commercialized by 1956, but relatively poor mass resolution was a problem until improvements were made in the early 1970s [13]. An ionization technique for biological molecules that could be used with TOF analysis was introduced in 1991. This ion source, matrix-assisted laser desorption/ionization (MALDI), was the result of work in Germany by Hillenkamp, Karas, and coworkers [14] and in Japan by Tanaka and coworkers [15]. Like ESI, the MALDI ion source was capable of ionizing and vaporizing large molecules such as proteins.

Protein structure was first studied by MS in 1990 [16], and the peptide mass fingerprinting technique used to identify proteins via MS was developed in 1993. Shortly thereafter, the ability to identify proteins was further enhanced by the development of software programs to search protein mass spectra data against online databases of amino acid sequences [17]. Isotope-coded affinity tags (ICAT), developed in 1999, made possible the quantification of individual proteins in a complex mixture. This technology allowed the simultaneous comparison of the expression levels of all proteins in cells grown in different media [5].

INSTRUMENTATION

All mass spectrometers contain at least three major components: an ion source, a mass analyzer, and an ion collection/detection system. The instrument must also be connected to a computer system to process and record the data and a vacuum pump to control the pressure within the mass spectrometer (see Fig. 1). One property shared by almost all internal components is that they are maintained at pressures far below atmospheric (10^{-6} to 10^{-8} torr) [18]. Low pressure is necessary to limit the number of ion collisions, which would alter the path of the ions and possibly produce unwanted reaction products or loss of charge [19]. The specific device used in each of the components depends on the type of information desired and the properties of the sample. In this review, we consider in detail only a few of the devices commonly used in protein MS. (For a comprehensive review of the types of instrumentation used, the reader is directed to Refs. 1 and 19.)

Ion Sources—The purpose of the ion source is to ionize, and in some cases vaporize, the sample. Some of the most common source types used in biological research today include electron impact ionization [1], chemical ionization

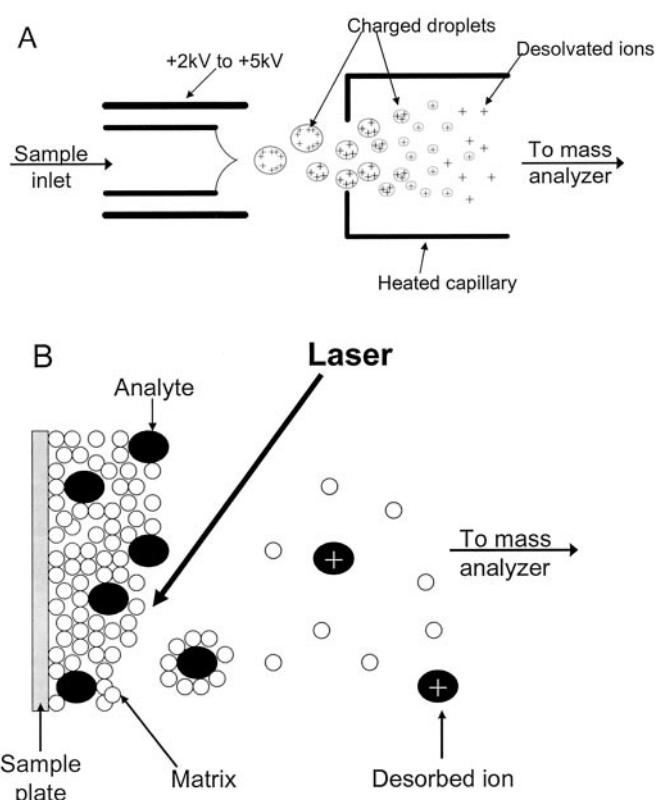


FIG. 2. Common ion sources for protein identification. A, The electrospray ion source. The liquid sample exits a capillary on which a voltage is applied. This process ionizes the sample and causes the exiting liquid to form a spray of small droplets. B, The MALDI source. A laser is pulsed at a mixture of sample and matrix molecules that have been co-crystallized.

[20], thermal ionization [1], ESI [12], and MALDI [14, 15]. The choice of ion source depends on the sample properties and the degree of ionization and fragmentation desired. Some experiments require that the sample be ionized intact, and others require that some molecular fragmentation occur.

In an ESI source, the sample is ionized when the inlet stream (liquid phase) is emitted from a capillary that has a voltage applied to it. This process forms a spray of highly charged droplets, which are then desolvated as they pass through several stages of decreasing pressure. In some cases, the droplets also pass through a heated capillary to improve desolvation (Fig. 2A). One reason behind the widespread use of ESI is that it is compatible with on-line separations such as LC. Also, ESI ionization often results in multiply charged ions. Because mass spectrometers measure the mass-to-charge ratio (m/z), having multiple charges increases the range of masses that can be analyzed [21].

MALDI is the most recently developed of the aforementioned ion sources. In MALDI, the sample is co-crystallized within an organic matrix such as sinapinic acid or α -cyano-4-hydroxycinnamic acid [21]. Co-crystallization is achieved by mixing a solution of the sample analyte with a solution of the matrix. The mixture is then applied to a metal target plate and allowed to dry. The resulting crystals are irradiated with laser pulses at a wavelength at which the matrix has high spectral absorption. This proc-

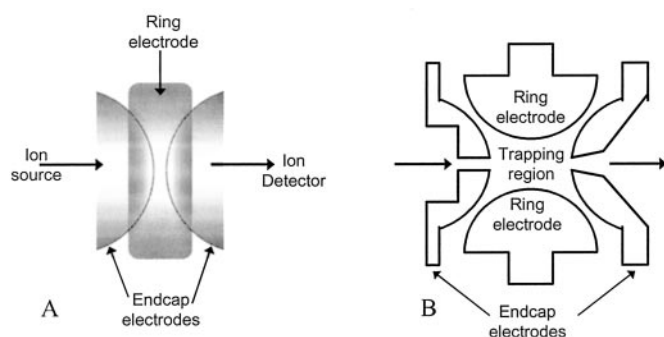


FIG. 3. **Ion trap mass analyzer.** *A*, schematic of the external view showing the ring electrode and endcap electrodes. *B*, cross-section of an ion trap showing the trapping region.

ess desorbs the mixture and photoexcites the matrix. The excited matrix then ionizes the analyte via proton transfer. The result of this process is analyte ions in the gas phase, the majority of which are singly charged [19] (Fig. 2B).

Mass Analyzers—After the sample has been ionized, the beam of ions is focused and directed into a mass analyzer, which separates the ions based on their m/z . The most common mass analyzers are the magnetic sector, the quadrupole, the ion trap, the TOF, and the Fourier transform-ion cyclotron resonance analyzer [19]. Here, the TOF and the ion trap will be discussed in more detail.

The performance of a mass analyzer is described by three characteristics: the upper mass limit, the transmission, and the resolution. The upper mass limit is the largest m/z that can be measured. The transmission is the fraction of ions produced at the source that reach the detector. The resolution is a measure of the ability to differentiate two m/z that are close together. If δm is the smallest difference between two mass peaks (m and $m + \delta m$), the resolution is $\delta m/m$ [19].

The m/z measurements for TOF analyzers are based on the equation:

$$KE = 0.5 mv^2$$

where KE is the kinetic energy of the ion, m is the mass of the ion, and v is the velocity of the ion. The ions are initially accelerated in an electric field with potential V , which results in a final kinetic energy of zV , where z is the charge on the ion. The ions then enter a flight tube (which is a field-free region) of length L , at the end of which there is an ion detector. Because the flight tube is field-free, all the KE of the ions results from their initial acceleration. Therefore, from the previous equation:

$$zV = 0.5 mv^2 \text{ and } m/z = (2Vt^2)/L^2$$

where t is the time it takes the ion to reach the detector. Because V and L are known, the m/z of an ion can be determined from the amount of time that passes between when the laser pulses (TOF analyzers are often paired with MALDI sources) and when the ion reaches the detector. TOFs have a high resolution and an extremely large mass range compared with other mass analyzers [22].

An ion trap is composed of a ring electrode and two end-cap electrodes (Fig. 3). Initially, the field within the trap is such that all ions that enter the analyzer begin a

stable oscillation within the analyzer and are “trapped.” The rf frequencies of the fields are then ramped up to eject ions of increasing m/z . The ions travel to the detector and, based on the frequency being used at the time of detection, the m/z of the ion can be calculated. An ion trap has high sensitivity and transmission but also a limited mass range and lower resolution compared with the TOF analyzer [22].

Ion Detection/Collection Systems—The original mass spectrometers used photolithic plates as detectors. However, the design of Nier in 1947 initiated the use of electronic detectors (hence the term mass “spectrometry” instead of “spectroscopy”) [3]. For isotope ratio measurements and some inorganic MS, Faraday cup collectors are used as the detector. Otherwise, the most commonly used detectors are the electron multiplier and the microchannel plate [21]. In both of these detectors, ions strike a metal plate, leading to a cascade of electron emissions that result in a measurable current [1].

Tandem MS—Tandem MS (also known as MS/MS or MS^2) is a method that isolates ions with a selected m/z (precursor ions), fragments them, and then measures the m/z of the resulting product ions. Tandem MS is widely used in both organic MS and biomolecular MS to elucidate structural details, or, in the case of peptides, an amino acid sequence. Tandem MS can be performed using an ion trap device or by using two mass analyzers separated by a collision cell. In an ion trap, all ions are ejected from the “trap” except those with the m/z of interest. The analyzer is then pressurized with a nonreactive gas such as helium, and ions are fragmented via collision-induced dissociation (CID). The fragments are then sequentially ejected and their masses determined [21]. Using an ion trap, multiple stages of fragmentation (MS^n) are possible. The analyzer can isolate an ion of interest, perform CID fragmentation, isolate a fragment, induce further fragmentation via CID, isolate a secondary fragment, etc. These subsequent fragmentation steps can reveal additional information about the structure of the analyte.

A recent addition to tandem MS instrumentation is the MALDI TOF/TOF [23]. In this instrument, two TOF analyzers are separated by a pressurized collision cell, which fragments the precursor ion and then reaccelerates the product ions before they enter the second TOF analyzer. This instrument benefits from both the high mass accuracy associated with TOF mass analyzers and the ability to do CID.

APPLICATIONS OF MS IN BIOLOGICAL RESEARCH

The range of MS applications to biology is extensive. Here we divide them into three basic categories: isotope-ratio MS, small organic molecule MS, and macromolecular MS. For each of these groups we discuss common applications and present summaries of recently published papers.

Isotope-Ratio MS—Isotope-ratio MS (IRMS) is a technique that measures the relative stable isotopic abundance of elements. The elemental isotope ratio can be analyzed for a complex system (bulk IRMS), for a specific compound within a mixture, or for a specific position within a compound. As an example, the $^{13}C/^{12}C$ ratio can be

measured for a tissue sample, for a specific fatty acid within the tissue, or for a specific carbon position within the fatty acid. IRMS can reveal information regarding the origin or state of a complex system [24]. For instance, the ratio of $^{13}\text{C}/^{12}\text{C}$ in apple juice can indicate if the juice has been adulterated with corn syrup [25]. The ratios of $^{13}\text{C}/^{12}\text{C}$, $^{15}\text{N}/^{14}\text{N}$, and $^{87}\text{Sr}/^{86}\text{Sr}$ can pinpoint the geographic origin of ivory, which can help conservationists prevent future poaching [26].

In biomedical research, IRMS is frequently used with projects involving stable isotopic tracers. For these experiments, compounds containing elements with abnormal stable isotope ratios are introduced into a system of interest (e.g. by feeding or by injection). The change in the isotopic ratio of the element in the system is then recorded with IRMS. Stable isotope tracers can be used to study the absorption, retention, and utilization of a nutrient *in vivo* [27]. Stable isotope tracers can also be used to study energy expenditure [3] or as a clinical tool for disease diagnosis [28]. IRMS uses magnetic sectors as the mass analyzers and Faraday cups as the ion detectors. The ion source is usually an electron impact or thermal ionization source [27].

An article by Davidsson and coworkers [29] describes an experiment to test the effectiveness of two different iron absorption enhancers (Na_2EDTA and ascorbic acid) to increase the iron absorbed by children in Peru's school meal program. On the first day, children were fed a breakfast "shake" containing one of the enhancers and $^{57}\text{FeSO}_4$. The next day, the same children were fed a breakfast "shake" containing either the original enhancer at a different concentration or a second enhancer and $^{58}\text{FeSO}_4$. Blood samples were taken before the first meal was given and 2 weeks after the second meal was consumed. The ratio of $^{58}\text{Fe}/^{57}\text{Fe}$ in the blood was measured to determine which meal allowed more iron absorption. It was found that ascorbic acid at a concentration of 70 mg/meal increased the iron absorbed to 115% of the daily requirement. This double stable isotope technique is useful because two variables can be tested in the same patient at the same time.

Small Organic Molecules—The mass spectrometric analysis of organic compounds gives information on the molecular mass, chemical formula, chemical structure, or quantity of the analyte. Based on the measured m/z and their peak intensities, the formula and chemical structure can be determined manually [30] and/or by comparison with a reference database of spectra. As an example, Lavermicocca *et al.* [31] used MS to identify novel antifungal compounds produced by *Lactobacillus plantarum* strain 21B (a bacterium used in making sourdough bread). Components of a culture filtrate were separated and tested for antifungal activity. Those that showed the highest activity were then characterized using gas chromatography-MS. The spectra were compared with those in a MS spectra library, and the compounds present were identified.

MS is also a valuable tool in the high-throughput analysis of compounds created in combinatorial libraries, such as those created to discover new drugs. If the mass spectrometer is connected to an LC system with appropriate

columns, both structural information and information regarding the binding affinity of a molecule can be obtained [7]. MS can also be used to study drug metabolism by identifying the primary metabolites from *in vitro* studies and by measuring them *in vivo* to determine pharmacokinetic parameters [7]. An example of this type of experiment is a study done by Oo *et al.* [32]. In this study, researchers administered oseltamivir (a treatment for influenza) to healthy patients between 1 and 12 years of age. The concentrations of oseltamivir and one of its metabolites were then monitored over time in the urine and plasma using MS. The pharmacokinetic data obtained allowed appropriate dosing schedules for children to be developed.

For instrumentation, small organic molecule MS often uses an electron impact or chemical ionization ion source. A wide variety of mass analyzers are used, and tandem MS is often employed to increase the structural information obtained.

Macromolecules—MS analysis of macromolecules includes the study of proteins, peptides, oligonucleotides, oligosaccharides, and lipids. We defer a discussion of applications to the study of proteins and peptides to "Protein Characterization." The research on the analysis of oligonucleotides has focused on studying modified oligonucleotides that may not be compatible with the current enzymatic techniques used to sequence DNA [33]. McLafferty and coworkers were able to sequence oligonucleotides up to 100 nucleotides in length using an ESI source coupled to a Fourier transform-ion cyclotron resonance analyzer [34]. MS can also be used to analyze DNA modifications such as methylation [35].

The characterization of oligosaccharides is more difficult than that of proteins and oligonucleotides because of the isomeric nature of the subunit and its ability to form branched structures. However, the structures of a wide variety of oligosaccharides have been determined by using tandem MS [19].

Many classes of molecules within the "lipid" category, including fatty acids, acylglycerols, and steroids, have been characterized by MS. The characterization of fatty acylcarnitines in blood can be used to screen newborns for hereditary fatty acid oxidation disorders. MS is used to measure the chain length and abundance of fatty acylcarnitines present in the blood. By comparing the resulting spectrum to that of a normal infant, doctors can diagnose if a fatty acid oxidation disorder is present, and in some cases they can determine the exact enzyme in which the newborn is deficient [4].

Most macromolecular research relies on the use of ESI or MALDI as an ionization source. A variety of mass analyzers, including the ion trap and TOF, are used. Tandem MS is often employed to elucidate structural details or sequence information.

PROTEIN CHARACTERIZATION

MS has become a vital tool in proteomic research. It can give information on the identity of a protein, the amount of the protein that is present, and the modifications the protein contains.

Protein Identification—The most commonly used MS

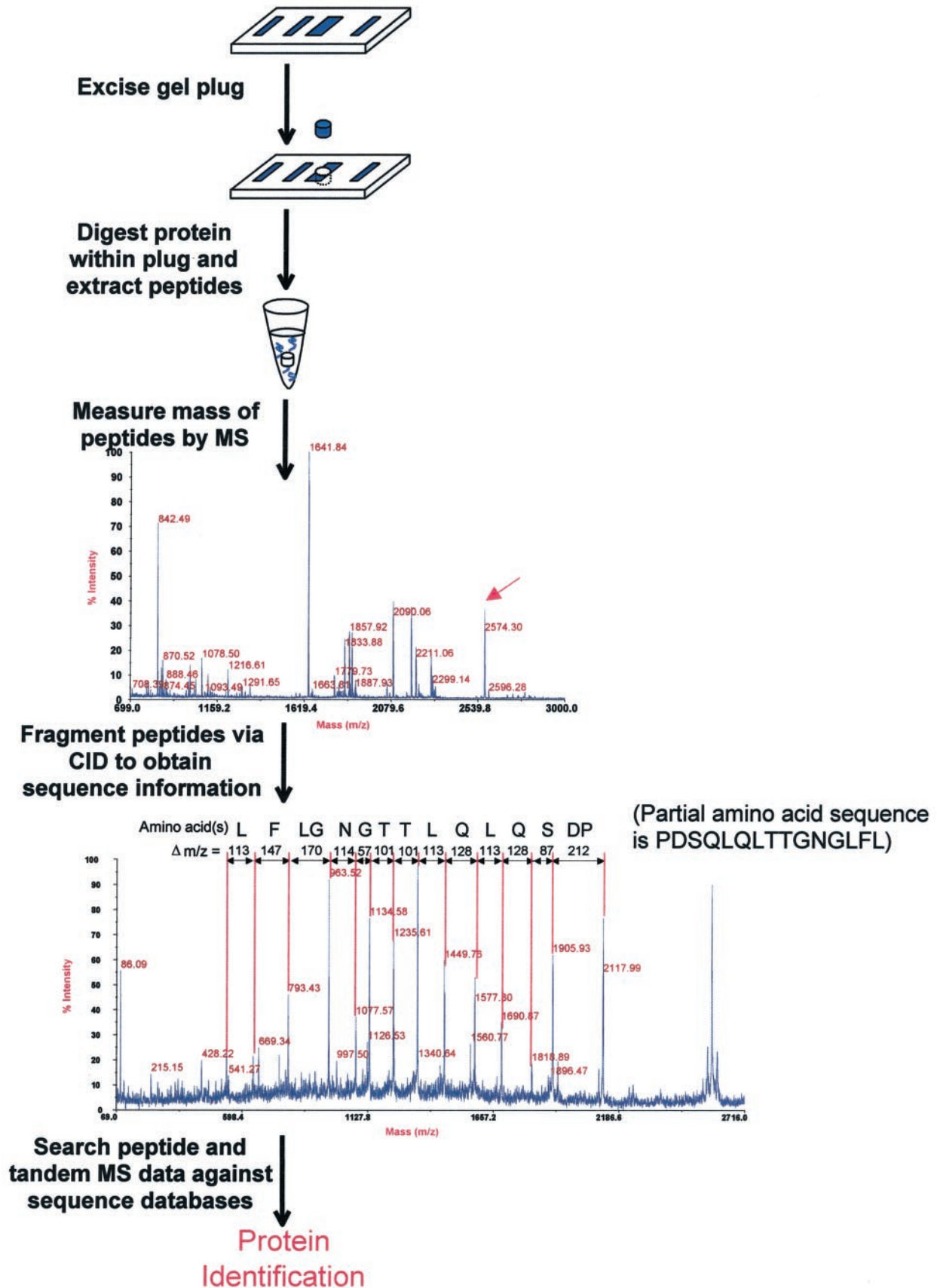


FIG. 4. **Overview of the process of protein identification.** A plug is excised from an electrophoresis gel, and the protein in the gel plug is digested and extracted. The masses of the peptides in the extract are then measured by MS to obtain the peptide mass fingerprint. Next, peptides can be selected to undergo fragmentation via tandem MS. Here, the peptide with $m/z = 2574$ (labeled with an arrow) has been selected and fragmented. Finally, both the MS and tandem MS data are searched against protein sequence databases to determine the protein that was present in the gel spot. In this example, data was collected using a MALDI-TOF/TOF (4700 Proteomics Analyzer; Applied Biosystems, Foster City, CA).

method to identify proteins is a combination of peptide mass fingerprinting and amino acid sequencing via tandem MS (see Fig. 4). Using these methods, sensitivity is routinely in the femtomole range [36]. The amino acid sequence of each protein is unique, and therefore the set of peptide masses resulting from proteolytic cleavage provides a “fingerprint” of the protein. In the peptide mass fingerprinting technique, an unknown protein is reduced and alkylated (to break any disulfide bonds) and then digested enzymatically using a sequence-specific proteolytic enzyme such as trypsin. Next, the masses of the resulting peptides are measured using MS. For the spectra in Fig. 4, a MALDI ion source was used, which results in almost all of the ions being singly charged (see “Ion Sources”).

The list of peptide masses from the spectra is entered into a search program, such as Mascot [37] or ProFound [38], along with information about the enzyme used, accuracy of the mass measurement, and possible protein modifications that may be present. The software then searches through databases of amino acid sequence information, (e.g. Swiss-Prot [39] or NCBI nr [8]). For each protein entry in the database, the software uses the amino acid sequence to predict the peptide masses that would result from digestion with the user-specified enzyme. The software then compares the list of peptide masses measured (using MS) to the list of predicted peptide masses and calculates the probability of a match. The result is a list of possible protein identifications and the probability that each identification is correct.

The probability that the match is correct can be increased by performing tandem MS on one or more of the original peptides. In tandem MS, peptides of a specific m/z , selected from the peptide mass fingerprint, are isolated from peptides of all other m/z . The isolated peptides are then fragmented by collisions with a gas such as helium that has been introduced into the mass spectrometer. The new peptide pieces formed by the CID are then measured. The weakest bond in a peptide is the one between the amino acids. Therefore, for low-energy CID, the resulting MS/MS spectrum is a series of peaks that represent peptides that differ only in the number of amino acids they contain. By measuring the mass difference between each peak, one can determine the amino acid sequence of the original peptide. High-energy CID also causes side-chain fragmentation. This information can be useful to differentiate between amino acids of the same molecular mass (e.g. leucine and isoleucine) [21].

In Fig. 4, the peptide with an m/z of 2574.30 was chosen for tandem MS analysis. The peptide was isolated and fragmented and the mass spectrum of the resulting fragments is shown. The partial amino acid sequence can be found by calculating the difference in mass between peaks. For example, the mass difference between the peaks with $m/z = 1577.80$ and $m/z = 1449.76$ is 128.04. This is within error of the mass of glutamine (Q), which is 128.06 Da. The probability of fragmentation, however, is not the same for all of the amide bonds [22]. This is evident by the wide range of peak intensities in the tandem MS data and the fact that some peaks are missing. For instance, due to the structure of proline, fragmentation at the

C-terminal end of proline is rare [22]. This is why in Fig. 4 we do not see a peak at $m/z = 2020.94$, which is equal to $m/z = 2117.99$ minus the mass of proline.

Although some work has been done to identify simultaneously all the proteins in a complex mixture, protein separation prior to enzymatic digestion is often performed. Proteins are separated using LC [40] or gel electrophoresis (either one- or two-dimensional gel electrophoresis) [36]. If separated by gel electrophoresis, the protein band or spot is excised, the proteins within are digested “in-gel,” and the resulting peptides are extracted for MS analysis [36, 41].

Protein Quantification—Proteomic research often requires knowing not only what proteins are being expressed by an organism, but also the level of protein expression. A common research technique is to compare the protein expression levels between multiple systems. In the past, quantitative analyses of protein levels using MS have been done by the metabolic labeling of the proteins in one system (e.g. cell lysates) with a heavy isotope such as ^{15}N [36]. However, this procedure is limited to cells and tissues compatible with metabolic labeling. A new method for labeling that does not have this limitation is the ICAT method.

In the ICAT approach, the ICAT reagent derivatizes the side chains of the cysteinyl residues. Proteins from one system are derivatized with a “light” form of the reagent and proteins from the other system with a “heavy” form. In the original ICAT procedure, the heavy form contained eight ^2H atoms [5]; however, in the current method, the heavy form includes nine ^{13}C atoms [42, 43]. The ICAT reagent has four parts: an affinity tag (biotin), a thiol-specific reactive group for covalent attachment to cysteine side chains, a linker, and an isotope tag (which contains ^{13}C in the heavy form) (see Fig. 5). The tagged protein mixtures are combined and enzymatically digested. The digest is fractionated using cation exchange chromatography, which also removes any neutral species from the tryptic peptides. The tagged peptides (those that contain a cysteine) in each fraction are then separated from the others on an affinity column, and the biotin portion of the tag is cleaved off. The peptides are then separated using LC and analyzed by MS [43]. The ratio of the intensities of the light- and heavy-labeled peptides gives a measurement of the ratio of protein expression in the two systems. The peptide is then analyzed with tandem MS to identify the parent protein.

Proteins with Post-translational Modifications—The function and activity of a protein is determined in part by any post-translational modifications that may be present. Over 200 distinct types of covalent modifications have been reported; the most common of these include phosphorylation, glycosylation, and ubiquitination [44]. Here the use of MS to study phosphorylation will be reviewed.

It is estimated that over 30% of proteins are phosphorylated [44]. The conventional biochemical approach to studying phosphopeptides requires the use of radioactive labels and Edman sequencing [45], but a variety of approaches have been developed to study phosphorylation using MS. If the amino acid sequence of the protein is already known, the sample is digested and the peptides

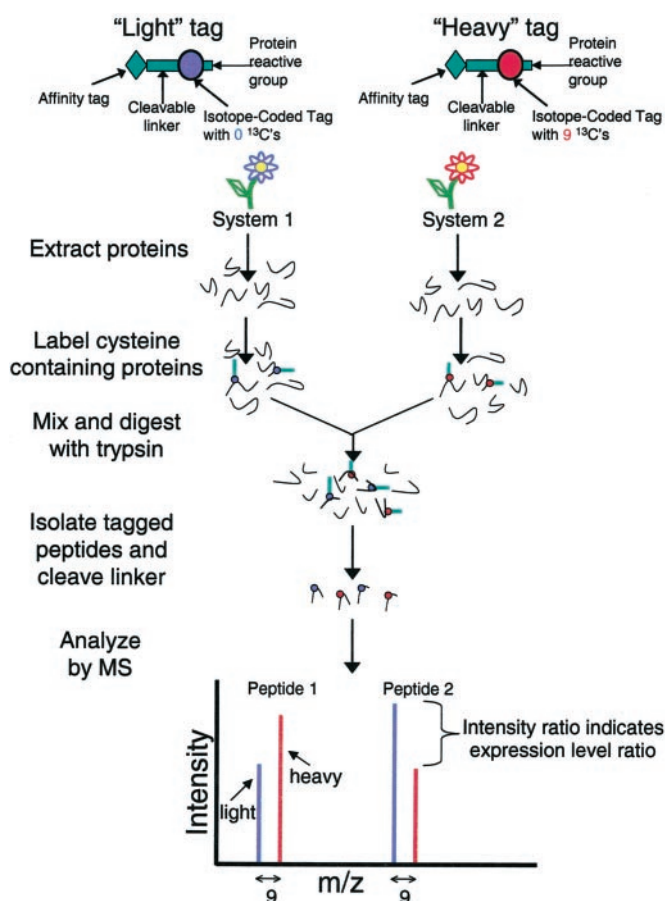


FIG. 5. **Relative protein quantification using ICAT.** The ICAT reagents contain four main components: an affinity tag, a cleavable linker, an isotope tag (which contains zero or nine ¹³C atoms for the "light" and "heavy" tags, respectively), and a group that reacts with cysteines. The proteins from one system are labeled with the "light" tag and the proteins from another system with the "heavy" tag. The proteins are mixed and digested with trypsin. The fragments containing cysteine residues are then isolated, the linker is cleaved to remove the affinity group, and the resulting peptides are analyzed with MS. The ratio of the "light" to "heavy" peaks in the spectra gives a measure of the ratio of protein expression between the systems.

analyzed with MS. The m/z of the peptides measured with MS are compared with the predicted values. An increase in a peptide mass by 80 or 160 Da indicates covalent modification by phosphate (HPO_3) (mono- and diphosphorylated species, respectively) [44]. The phosphorylated peptide can then be analyzed using MS/MS to determine the exact position of the modification. If the protein is unknown, the sample is divided after digestion, and one portion is treated with a phosphatase. This step removes any phosphate groups, and a comparison of spectra before and after phosphatase treatment will indicate which peptides are modified. The modified peptides can be analyzed with MS/MS to identify the parent protein and the site of phosphorylation [45].

The presence of phosphorylation can also be determined using different scanning modes on the mass spectrometer. In "neutral loss scanning," the mass spectrometer scans for an ion that when fragmented results in a specified neutral loss (change in mass but not in charge) [19]. When searching for phosphorylated peptides, the

tandem MS spectra are searched for a fragment that has a mass 98 Da lower than the original peptide mass (with no change in charge). This neutral loss indicates a loss of H_3PO_4 due to β -elimination. This process cannot detect phosphotyrosines, however, due to stability of the β -protons in the benzene ring [44]. In "precursor ion scanning," the mass spectrometer scans for a specific product ion after fragmentation [19]. For phosphotyrosines, the peptides are fragmented using CID and scanned for a positive fragment ion with 216 m/z (phosphotyrosine immonium ion). For other phosphorylations, the peptides are fragmented using CID and scanned for a negative ion with 79 m/z due to the release of PO_3^- [44].

MS has been used to study a multitude of post-translational modifications, often employing the same approaches discussed here for phosphorylation: treatment with an enzyme to remove the modification (with subsequent spectra comparison), neutral loss scanning, or precursor ion scanning. It should be noted that for modifications by heterogeneous molecules (such as glycans), the structure of the modification could also be analyzed with MS [46].

CONCLUSIONS

The use of MS was originally limited to the study of elemental isotopes. However, as the instrumentation, computer technology, and associated software have improved, the range of MS applications has increased. The wide variety of sample types that can now be analyzed, and the breadth of information that can be obtained, have helped MS to permeate into an extensive range of research areas. The result is that MS has become an essential analytical tool in biological research.

REFERENCES

- [1] H. H. Willard, L. L. Merritt, Jr., J. A. Dean, F. A. Settle, Jr. (1988) *Instrumental Methods of Analysis*, Wadsworth Publishing Co., Belmont, CA, pp. 465–507.
- [2] B. Guo (1999) Mass spectrometry in DNA analysis, *Anal. Chem.* **71**, 333–337R.
- [3] J. R. de Laeter (2001) *Applications of Inorganic Mass Spectrometry*, John Wiley & Sons, Inc., New York.
- [4] D. H. Chace, T. A. Kalas, E. W. Naylor (2002) The application of tandem mass spectrometry to neonatal screening for inherited disorders of intermediary metabolism, *Annu. Rev. Genomics Hum. Genet.* **3**, 17–45.
- [5] S. P. Gygi, B. Rist, T. J. Griffin, J. Eng, R. Aebersold (2002) Proteome analysis of low-abundance proteins using multidimensional chromatography and isotope-coded affinity tags, *J. Proteome Res.* **1**, 47–54.
- [6] K. Y. Patterson, C. Veillon (2001) Stable isotopes of minerals as metabolic tracers in human nutrition research, *Exp. Biol. Med.* **226**, 271–282.
- [7] N. Mano, J. Goto (2003) Biomedical and biological mass spectrometry, *Anal. Sci.* **19**, 3–14.
- [8] NCBI homepage (accessed September 21, 2003) www.ncbi.nlm.nih.gov.
- [9] Nobel Prize in Chemistry 2002 (accessed September 21, 2003) www.nobel.se/chemistry/laureates/2002/index.html.
- [10] ASMS Posters on the History of Mass Spectrometry (accessed September 18, 2003) masspec.scripps.edu/information/history/perspectives/posters.html.
- [11] M. Barber, R. S. Bordoli, R. D. Sedgwick, A. N. Tyler (1981) Fast atom bombardment of solids as an ion source in mass spectroscopy, *Nature* **293**, 270–275.
- [12] J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, C. M. Whitehouse (1989) Electrospray ionization for mass spectrometry of large biomolecules, *Science* **246**, 64–71.
- [13] M. A. Grayson (accessed September 10, 2003) *The Application of Mass Spectrometry to Biological Problems: A Historical Perspective*, wunmr.wustl.edu/~msf/ASMS2003/Mike_Grayson_abstract_2003.pdf.

- [14] M. Karas, F. Hillenkamp (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons, *Anal. Chem.* **60**, 2299–2301.
- [15] K. Tanaka, H. Waki, Y. Ido, S. Akita, Y. Yoshida, T. Yoshida (1988) Protein and polymer analysis up to m/z 100,000 by laser ionization time-of-flight mass spectrometry, *Rapid Commun. Mass Spectrom.* **2**, 151–153.
- [16] S. K. Chowdhury, V. Katta, B. T. Chait (1990) Electrospray ionization mass spectrometric peptide mapping: A rapid, sensitive technique for protein structure analysis, *Biochem. Biophys. Res. Commun.* **167**, 686–692.
- [17] D. J. C. Pappin, P. Hojrup, A. J. Bleasby (1993) Rapid identification of proteins by peptide-mass fingerprinting, *Curr. Biol.* **3**, 327–332.
- [18] R. Bakhtiar, F. L. Tse (2000) Biological mass spectrometry: A primer, *Mutagenesis* **15**, 415–430.
- [19] E. De Hoffmann, J. Charette, V. Stroobant (1996) *Mass Spectrometry, Principles and Applications*, John Wiley & Sons, France.
- [20] M. S. B. Munson, F. H. Field (1966) Chemical ionization mass spectrometry I. General introduction, *J. Am. Chem. Soc.* **88**, 2621–2630.
- [21] P. James, ed. (2001) *Proteome Research: Mass Spectrometry*, Springer-Verlag, Germany.
- [22] M. Kinter, N. E. Sherman. (2000) *Protein Sequencing and Identification Using Tandem Mass Spectrometry*, John Wiley & Sons, Inc., New York.
- [23] K. F. Medzihradsky, J. M. Campbell, M. A. Baldwin, A. M. Falick, P. Juhasz, M. L. Vestal, A. L. Burlingame (2000) The characteristics of peptide collision-induced dissociation using a high-performance MALDI-TOF/TOF tandem mass spectrometer, *Anal. Chem.* **72**, 552–558.
- [24] J. T. Brenna, T. N. Corso, H. J. Tobias, R. J. Caimi (1997) High-precision continuous-flow isotope ratio mass spectrometry, *Mass Spectrom. Rev.* **16**, 227–258.
- [25] I. T. Platzner (1997) *Modern Isotope Ratio Mass Spectrometry*, John Wiley & Sons, Inc., New York.
- [26] N. J. van der Merwe, J. A. Lee-Thorp, J. F. Thackeray, A. Hall-Martin, F. J. Kruger, H. Coetzee, R. H. V. Bell, M. Lindeque (1990) Source-area determination of elephant ivory by isotopic analysis, *Nature* **346**, 744–746.
- [27] T. Walczyk (2001) The potential of inorganic mass spectrometry in mineral and trace element nutrition research, *Fresenius J. Anal. Chem.* **370**, 444–453.
- [28] A. Canete, Y. Abunaji, G. Alvarez-Calatayud, M. DeVicente, J. A. Gonzalez-Holguera, M. Leralta, J. M. Pajares, J. P. Gisbert (2003) Breath test using a single 50-mg dose of ^{13}C -urea to detect *Helicobacter pylori* infection in children, *J. Pediatr. Gastroenterol. Nutr.* **36**, 105–111.
- [29] L. Davidsson, T. Walczyk, N. Zavaleta, R. Hurrell (2001) Improving iron absorption from a Peruvian school breakfast meal by adding ascorbic acid or Na_2EDTA , *Am. J. Clin. Nutr.* **73**, 283–287.
- [30] F. W. McLafferty, F. Turecek (1993) *Interpretation of Mass Spectra*, University Science Books, Sausalito, CA.
- [31] P. Lavermicocca, F. Valerio, A. Evidente, S. Lazzaroni, A. Corsetti, M. Gobetti (2000) Purification and characterization of novel antifungal compounds from the sourdough *Lactobacillus plantarum* strain 21B, *Appl. Environ. Microbiol.* **66**, 4084–4090.
- [32] C. Oo, G. Hill, A. Dorr, B. Liu, S. Boellner, P. Ward (2003) Pharmacokinetics of anti-influenza prodrug oseltamivir in children aged 1–5 years, *Eur. J. Clin. Pharmacol.* **9**, 9.
- [33] K. A. Sannes-Lowery, S. A. Hofstadler (2003) Sequence confirmation of modified oligonucleotides using IRMPD in the external ion reservoir of an electrospray ionization fourier transform ion cyclotron mass spectrometer, *J. Am. Soc. Mass Spectrom.* **14**, 825–833.
- [34] D. P. Little, T. W. Thannhauser, F. W. McLafferty (1995) Verification of 50- to 100-mer DNA and RNA sequences with high-resolution mass spectrometry, *Proc. Natl. Acad. Sci. U. S. A.* **92**, 2318–2322.
- [35] A. Humeny, C. Beck, C. M. Becker, A. Jeltsch (2003) Detection and analysis of enzymatic DNA methylation of oligonucleotide substrates by matrix-assisted laser desorption ionization time-of-flight mass spectrometry, *Anal. Biochem.* **313**, 160–166.
- [36] H. W. Lahm, H. Langen (2000) Mass spectrometry: A tool for the identification of proteins separated by gels, *Electrophoresis* **21**, 2105–2114.
- [37] D. N. Perkins, D. J. Pappin, D. M. Creasy, J. S. Cottrell (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data, *Electrophoresis* **20**, 3551–3567.
- [38] W. Zhang, B. T. Chait (2000) ProFound: An expert system for protein identification using mass spectrometric peptide mapping information, *Anal. Chem.* **72**, 2482–2489.
- [39] B. Boeckmann, A. Bairoch, R. Apweiler, M. C. Blatter, A. Estreicher, E. Gasteiger, M. J. Martin, K. Michoud, C. O'Donovan, I. Phan, S. Pilbout, M. Schneider (2003) The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003, *Nucleic Acids Res.* **31**, 365–370.
- [40] D. A. Wolters, M. P. Washburn, J. R. Yates, 3rd (2001) An automated multidimensional protein identification technology for shotgun proteomics, *Anal. Chem.* **73**, 5683–5690.
- [41] E. J. Finehout, K. H. Lee (2003) A comparison of automated in-gel digest methods for femtomole level samples, *Electrophoresis* **24**, 3508–3516.
- [42] Applied Biosystems (accessed September 28, 2003) www.appliedbiosystems.com.
- [43] K. C. Hansen, G. Schmitt-Ulms, R. J. Chalkley, J. Hirsch, M. A. Baldwin, A. L. Burlingame (2003) Mass spectrometric analysis of protein mixtures at low levels using cleavable ^{13}C -isotope-coded affinity tag and multidimensional chromatography, *Mol. Cell. Proteomics* **2**, 299–314.
- [44] R. E. Schweppe, C. E. Haydon, T. S. Lewis, K. A. Resing, N. G. Ahn (2003) The characterization of protein post-translational modifications by mass spectrometry, *Acc. Chem. Res.* **36**, 453–461.
- [45] X. Zhang, C. J. Herring, P. R. Romano, J. Szczepanowska, H. Brzeska, A. G. Hinnebusch, J. Qin (1998) Identification of phosphorylation sites in proteins separated by polyacrylamide gel electrophoresis, *Anal. Chem.* **70**, 2050–2059.
- [46] A. Sickmann, M. Mreyen, H. E. Meyer (2002) Identification of modified proteins by mass spectrometry, *IUBMB Life* **54**, 51–57.