CMLS, Cell. Mol. Life Sci. 58 (2001) 868–884
1420-682X/01/070868-17 $ 1.50 + 0.20/0


CMLS Cellular and Molecular Life Sciences

# Review

# Mass spectrometry for protein and peptide characterisation

**A. P. Jonsson**

Department of Medical Biochemistry and Biophysics, Karolinska Institutet, 171 77 Stockholm (Sweden),
Fax: +46 8 33 74 62, e-mail: Andreas.Jonsson@mbb.ki.se

Received 17 April 2001; accepted 19 April 2001

**Abstract.** Mass spectrometry has become an important analytical tool in biological and biochemical research. Its speed, accuracy and sensitivity are unmatched by conventional analytical techniques. Identification of proteins and characterisation of their primary structure is a rapidly growing field in the post-genomic era, where matrix-assisted laser desorption/ionisation time-of-flight peptide mass fingerprinting combined with electrospray tandem mass spectrometry can efficiently solve many questions. Many recently determined genomic sequences have not been characterised at the protein level. Analysis of the amino acid sequence and characterisation of post-translational modifications are therefore important steps towards correlation of protein structure with function. This review concerns methods, instrumentation and applications of mass spectrometry in protein and peptide analysis.

**Key words.** Mass spectrometry; protein; peptide; matrix-assisted laser desorption/ionisation; electrospray ionisation; quadrupole; time-of-flight.

## A general overview of mass spectrometry

During the last decade, mass spectrometry has emerged as a major analytical technique in molecular-level biological research. There has been an explosion in instrumentation and applications to large biomolecules, particularly in the field of protein and peptide analysis [1].

Mass spectrometry has certainly added a new dimension to the accuracy in determining protein molecular weights. Improvements of many orders of magnitude can be made using mass spectrometry rather than techniques such as gel filtration and sedimentation analysis [2]. The introduction of sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) [3] allowed protein molecular weights to be determined with a mass accuracy of approximately 5–10%. This can be compared to the parts per million (ppm) and parts per billion (ppb) accuracies that are obtained with modern mass spectrometers [4–6].

The principle of mass spectrometry is measurement of the mass-to-charge ratio ($m/z$) of gas-phase ions. Mass spectrometry is almost always a destructive analytical method, meaning that the sample is consumed. Sir J. J. Thomson developed the technique in the early 1900s when he obtained mass spectra of small gaseous ions [7]. His work has recently been reviewed by I. W. Griffiths [8]. Thomson's work was further developed by F. W. Aston who measured the masses of more than one hundred stable isotopes using similar techniques [9–11].

Mass spectrometry is now a widespread technique with key functions in a number of fields, including biochemistry [12], biotechnology [13], pharmacology [14], microbiology [15] and the proteomics/functional genomics field [16, 17]. This development is a direct consequence of the introduction of the soft ionisation methods of electrospray ionisation [18–20] and matrix-assisted laser desorption/ionisation (MALDI) [21]. These ionisation methods have made it possible to ionise large thermally labile biomolecules and transfer them to the gas phase without dissociation. Non-covalent complexes of biomolecules can nowadays be analysed in the megadalton (MDa) range [22].

When the ionised molecules have been transferred to the gas phase, a number of mass analysers can be employed to determine their mass-to-charge ratio, such as quadrupole mass filters [23], double-focusing magnetic and electric sectors [24, 25], time-of-flight (TOF) [26, 27], quadrupole ion traps [28, 29] and Fourier-transform ion cyclotron resonance (FTICR) [30] mass analysers. The choice of mass analyser will reflect the needs of the specific application, and will probably take mass accuracy, resolution, sensitivity and cost into account. Two or more analysers can be combined to form a tandem mass spectrometer, such as a triple quadrupole [31] or a hybrid instrument (consisting of two different analysers), e.g. quadrupole-TOF (Q-TOF) [32] or magnetic sector-TOF [33]. A common feature of these tandem mass spectrometers is a collision cell located between the two analyser regions. This cell contains gas into which a precursor ion (e.g. a protonated peptide), which has been selected by the first mass analyser, collides and fragments in a process called collision-induced dissociation (CID) or collisionally activated decomposition [34]. The fragments produced by the tandem mass spectrometry (MS/MS) process are called product ions and they are mass-measured in the second analyser region. The nature of the product ions can subsequently be used to determine the primary structure of the peptide, i.e. the amino acid sequence [35]. This MS/MS method is widely used today in modern biochemistry, where unknown proteins can be identified by determination of a sequence tag from a proteolytic peptide of the protein [36, 37]. Determination of protein primary structure by mass spectrometry is now considered a routine technique [38, 39] and much interest is currently directed towards identification of post-translational modifications [40, 41].

Peptide mass fingerprinting (PMF) [42, 43] is another protein identification strategy in the post-genomic era [44]. This method is based on proteolytic digestion of the protein with a site-specific enzyme, e.g. trypsin, and simple mass measurement of the created peptides, often carried out on a MALDI TOF instrument. The peptide mass fingerprint is searched against databases in which the protein sequences have been theoretically digested with the same enzyme. This method is highly compatible with automation and is perhaps the best choice for high-throughput identification of large sample sets.

## Ionisation techniques

Mass spectrometry measures the mass-to-charge ratio of ionised molecules in the gas phase. Hence, the analytes need to be ionised and transferred to the gas phase prior to analysis. Earlier techniques such as electron impact (EI) [24] and chemical ionisation (CI) [45] were efficient in ionising small volatile thermally stable molecules. Fast-atom bombardment (FAB) was introduced in the early 1980s by Barber and co-workers [46] and was the first ionisation technique that could be used for routine analysis of biomolecules of mass up to a few thousand daltons [47]. The soft ionisation techniques, MALDI and electrospray, discussed in the Introduction and described below, are the revolutionising methods that have made mass spectrometry one of the most important tools to analyse large biomolecules.

### Electrospray ionisation

Electrospray is an atmospheric-pressure ionisation method that produces small charged droplets from a liquid medium under the influence of an electric field. The process itself originates from the beginning of the last century but it was Dole and co-workers [48] who first transferred large molecules to the gas phase in the late 1960s. In 1984, electrospray was used for the first time to create gas phase ions to be analysed by a mass spectrometer. The group of J. B. Fenn reported the use of electrospray mass spectrometry [18, 19] at approximately the same time as Alexandrov and co-workers [20] but it was still a few more years before the first electrospray mass spectra of large molecules were published, again from the group of Fenn [49, 50].

As Alexandrov et al. [20] reported 1984, electrospray is a very suitable method for combining chromatographic methods with mass spectrometry. This is because the electrospray process transfers ions from the solution phase to the gas phase at atmospheric pressure. In conventional electrospray, a flow of liquid, from a chromatographic system or a syringe pump, is passed through a thin conducting needle at high voltage (3–4 kV). The potential difference is applied between the needle and the counter electrode (the inlet of the mass spectrometer). The analytes will to some extent, depending on the pH of the solvent, exist in an ionised form in the liquid and the applied potential will create an accumulation of like charges at the tip of the needle. Positive or negative ions will migrate to the end of the capillary depending on the polarity of the applied field. Figure 1 illustrates the electrospray process in the positive-ion mode, where positively charged ions accumulate at the liquid surface (meniscus). Ions of opposite polarity (negatively charged ions) will migrate towards the positive capillary wall. The high density of positive charges at the tip leads to the formation of a Taylor cone [51] due to the repulsive coulombic forces between the positive ions. Under a high electric field, the repulsive forces become stronger than the surface tension at the tip of the cone, and a liquid mist is ejected. The mist breaks up in small highly charged droplets when moving towards the counter electrode (fig. 1). Both a potential and a pressure gradient will direct the droplets towards the inlet of the mass spectrometer and a counter-current

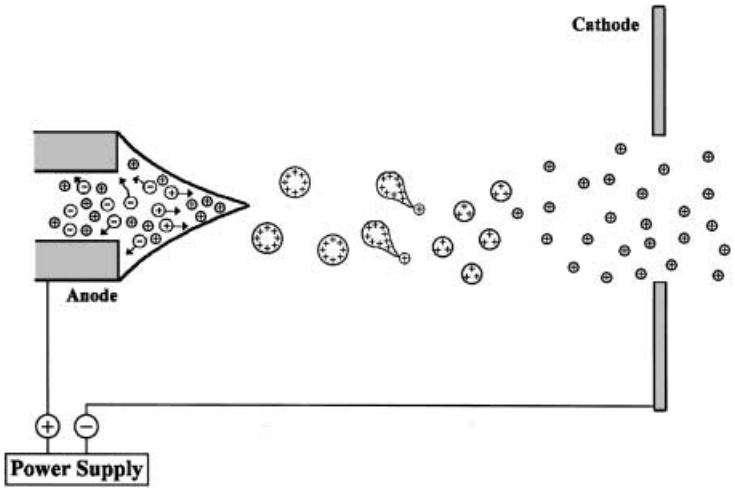


Figure 1. Schematic representation of the electrospray process. A positive potential is applied to the capillary (anode), causing positive ions in solution to drift towards the meniscus. The meniscus destabilises, leading to the formation of a Taylor cone emitting droplets with excess positive charge. Gas phase ions are formed from charged droplets in a series of solvent evaporation cycles as they are accelerated towards the entrance of the mass spectrometer (cathode). [Modified with permission from ref. 56.]

flow of gas facilitates solvent vaporisation and prevents ion cluster formation. As the solvent evaporates, the charge remains constant in the droplets and at a certain point, the surface coulombic forces exceed the surface tension forces, the Rayleigh limit [52], and the droplets break up into smaller droplets [53]. This process continues until droplets with diameters in the nanometer range are generated.

The formation of gas phase ions from the small charged droplets is not yet fully understood. Two mechanisms have been proposed. The original idea was that the solvent evaporates and droplets break up until those with only a single analyte ion are created [48]. The evaporation continues until a gas phase ion is formed. This is usually referred to as the charged-residue model. Iribarne and Thomson [54] suggested an alternative mechanism in 1976 (the ion evaporation model), in which they proposed that droplets with a radius less than 10 nm can allow field desorption, i.e. direct emission of a gaseous ion. The charge state of the ion will depend on the number of charges that are transferred from the droplet surface to the ion during desorption. The gas phase ion formation processes are still under debate, and while the ion evaporation theory might be the most accepted, a mechanism related to the charged-residue process may account for the formation of gaseous protonated macromolecules [55]. A number of useful papers and volumes have been published in which detailed descriptions of the electrospray process are discussed [56–59].

The electrospray ion source has gone through major developments since its introduction. Conventional electrospray instruments operate best at a flow rate of 3–10 µl/min, but the coupling of liquid chromatography (LC) to mass spectrometry has sometimes demanded flow rates of up to 1 ml/min. The electrospray evaporation is then facilitated by a coaxial gas flow (a nebulising gas) [60]. This type of source is generally called pneumatically assisted electrospray. On the other hand, the coupling of capillary electrophoresis (CE) to mass spectrometry gives very low flow rates and a sheath flow of organic solvent might be needed. However, this process dilutes the sample solution and sheathless interfaces have been developed [61]. Like all other analytical techniques, miniaturisation has been one of the key steps in the development of electrospray ionisation. The first report on low-flow-rate electrospray ionisation came in 1993 [62]. The following year, Emmett and Caprioli [63] followed up with a continuous infusion source operating at a flow rate of approximately 300–800 nl/min, resulting in a major sensitivity increase. This source is called the microelectrospray source and neuropeptides have been analysed down to the zeptomole ($10^{-21}$ mole) level [64]. The first electrospray source without continuous infusion was developed by Wilm and Mann [65, 66], and called nanospray. The sample is sprayed from a metal-coated capillary (needle) with an opening of 1–10 µm. This results in a flow rate of 20–50 nl/min and, hence, small sample volumes are consumed. Less than 0.5 µl can be loaded into the needle and used for more than 15 min in favourable situations. This time frame is usually enough for several MS/MS experiments. The flow rate will depend on the orifice diameter of the needle, the applied voltage (usually < 1 kV) as well as the viscosity and volatility of the solvent. The nanospray source has higher ionisation efficiency than a conventional electrospray source due to the production of smaller droplets, and it is also reportedly less sensitive to salts than conventional electrospray. Another modification that

has made electrospray sources overall more tolerant to inorganic salts is the Z-spray configuration [67]. In this set-up, the spraying device is mounted at a right angle to the inlet/cone of the mass spectrometer. The idea is that the major part of the inorganic salts will not be desorbed from the droplets and will hence travel in a straight path and not be analysed, while the organic ions will be more easily transferred to the gas phase and transported into the mass spectrometer down the pressure and potential gradients.

As mentioned above, electrospray ionisation produces multiply charged ions, and proteins with molecular weights in the 10–100 kDa mass range will in general produce an envelope of ions with *m/z* values below 2500 (fig. 2). Ion transmission is generally good in this region and mass measurement statistics are excellent due to many different charge states being observed. These features make electrospray the most suitable ionisation method for molecular-weight determination of large biomolecules. The envelope of different charge states can be converted to a true mass scale via maximum-entropy processing (fig. 2), an iterative method that calculates the mass and the abundance from the experimental *m/z* peaks [68] or via other deconvolution or transform processes.

**Matrix-assisted laser desorption/ionisation**

MALDI was introduced in the late 1980s by the group of Hillenkamp [21]. The MALDI technique is, like electrospray ionisation, referred to as 'soft' and thereby compatible with analysis of large biomolecules like proteins [69, 70].

The analytes are mixed with a saturated solution of ultraviolet-absorbing matrix. The most commonly used matrices in peptide/protein analysis are $\alpha$-cyano-4-hydroxy-cinnamic acid (for small peptides) and 3,5-dimethoxy-4-hydroxy-cinnamic acid, i.e. sinapinic acid (for larger peptides and proteins). The matrix/analyte mixture is applied to a target plate. The solvent evaporates and the matrix and the analytes co-crystallise on the target. A laser beam (commonly a nitrogen laser at 337 nm) provides light that is absorbed by the aromatic matrix molecules. Energy is subsequently transferred to the analyte that becomes desorbed into the gas phase. The ionisation mechanism is not fully understood and several suggestions are still debated [71]. Co-desorption of matrix and analyte succeeds proton transfer, which may take place in the solid phase, and also in the expanding plume of matrix and analyte ions after the laser irradiation.

The MALDI source has traditionally been coupled to TOF mass analysers because of its pulsed nature. Recent developments have made possible atmospheric pressure MALDI mass spectrometry.

MALDI produces mainly singly charged ions, and this feature means it is excellently suited for analysis of complex biological mixtures such as protein digests [43]. MALDI TOF mass spectrometry is thus the primary analytical technique in proteomics for identification of proteins separated by two-dimensional gel electrophoresis [17]. The basic amino acid residues in peptides and proteins are easily protonated and, consequently, these biomolecules are preferably analysed in the positive ion mode. The MALDI ionisation process is less sensitive to salts than electrospray ionisation. Nevertheless, salts and other impurities will cause peak broadening with the formation of adducts. This will lower the mass accuracy and limit sensitivity, and these problems highlight the importance of satisfactory sample preparation. A number of different clean-up approaches have been reported [72, 73] often including the use of microcolumns for desalting, which have recently become available commercially (ZipTip; Millipore). Several biotechnology companies have launched robotic systems that desalt the large sample sets produced, before the introduction of sample to MALDI mass spectrometry.

For a long time, MALDI TOF was considered a low-resolution mass spectrometric method. The introduction of delayed extraction [74–76] and the mass reflectron [77] has changed this view dramatically. Descriptions of these two features can be found below in the 'The time-of-flight mass analyser' section.

MALDI TOF has lately become the instrument of choice for many laboratories that are investing in mass spectrometry equipment. The major advantages are that it is an easy system to operate, it requires minimal mass spectrometric expertise to obtain data and it usually gives a quick result. The relatively low purchasing cost together with reasonable running costs are other parameters that make the system attractive from a buyers point of view. However, as with all mass spectrometric systems, the interpretation of data, whether mass spectra or a list of database search 'hits', requires a degree of expertise and provides the rate-determining step for protein analysis.

**Mass analysers**

When ions have been formed in the source, they are transported to the analyser region and separated according to their mass-to-charge ratio. A number of mass analysers are available which can be divided into two classes. The first class, electric-field mass analysers, consist of the quadrupole mass filter, the quadrupole ion trap (Paul trap) and the TOF mass analyser. The second class, magnetic-field mass analysers, comprise magnetic sectors and ion cyclotron resonance mass analysers. The different analysers vary in their mass accuracy, mass range, resolu-

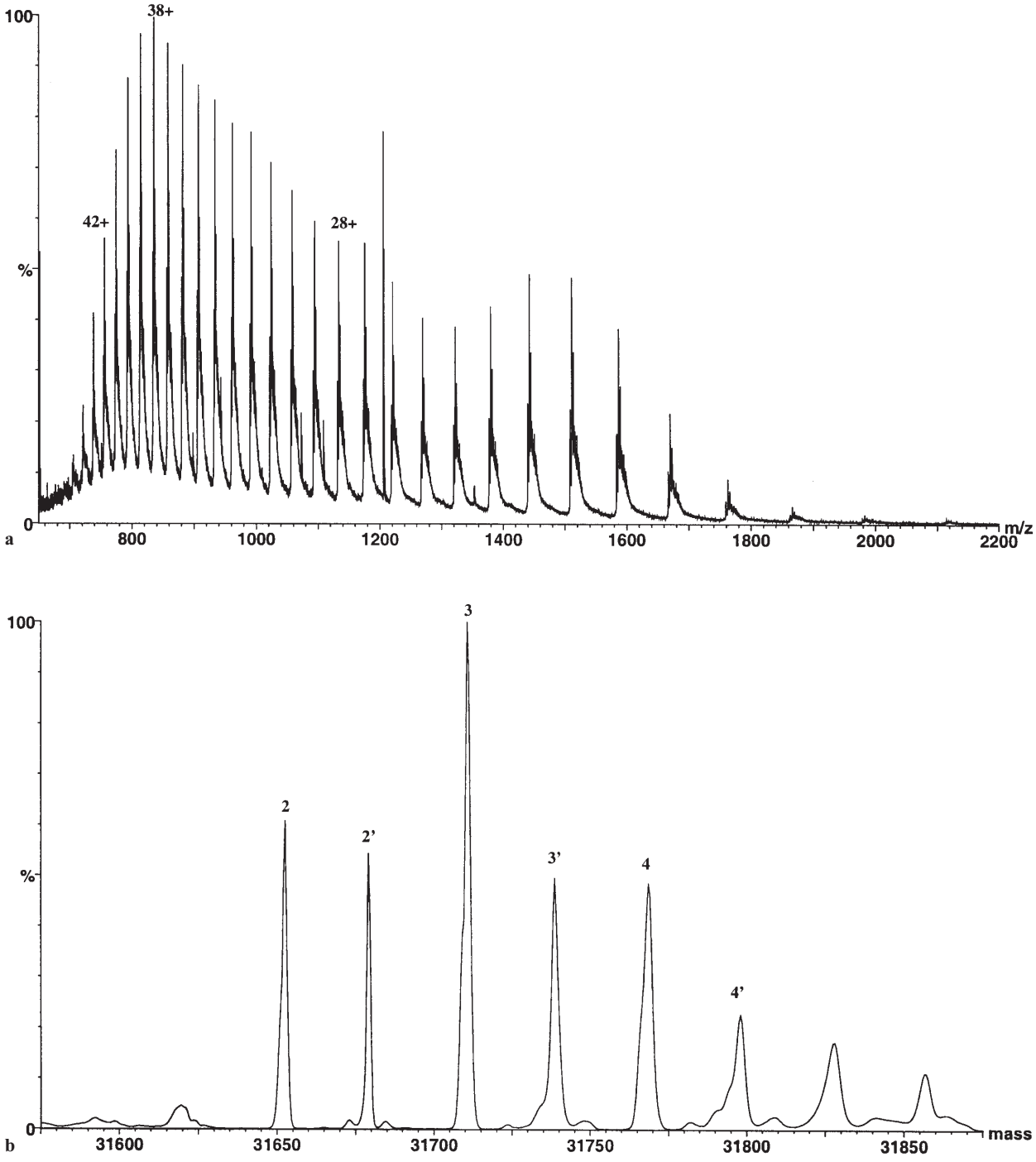


Figure 2. (*a*) Mass spectrum of the partially carboxymethylated TR-LBD. Three charge states are indicated. (*b*) Maximum-entropy-derived spectrum. The number of carboxymethylated Cys residues present in each form is indicated. An apostrophe indicates that the form is additionally formylated [4].

tion, sensitivity, speed, footprint, cost, and the choice will depend on the specific application. The quadrupole ion trap, for example, has $MS^n$ capability (the possibility to perform dissociation analysis of created product ions), and FTICR mass spectrometers have extraordinary resolution (up to $10^7$) and mass accuracy [6]. Magnetic-sector instruments are the most mature instruments on the market and have been a mainstay for many decades. Today, they are the instrument of choice for environmental-pollutant analysis. However, quadrupoles, traps and TOF instruments have largely accommodated the growth in biological applications while the market for

sector instruments remains constant. The characteristics of the quadrupole mass filter and the TOF mass analyser are described below.

### The quadrupole mass filter

The quadrupole mass filter is the most common mass analyser in use today and can be regarded as a real 'workhorse'. It was introduced in the early 1950s [23] and the technique has only seen modest developments since then. The mass filter is used extensively as both a stand-alone device and in multistage mass spectrometers like triple quadrupoles [31] and quadrupole TOF instruments [32, 78]. The quadrupole analyser is constructed of four electronically conducting cylindrical rods and is operated by the application of a combination of direct current (DC) and radio frequency (RF) voltages (fig. 3).

The mass filter establishes a two-dimensional quadrupole field between the four cylindrical electrodes with the two opposite rods connected electrically. One rod pair (+) is connected to a positive DC voltage, upon which a sinusoidal RF voltage is superimposed. The other rod pair (–) is connected to a negative DC voltage, upon which a sinusoidal RF voltage is also superimposed. Successful selection of a specific ion requires the RF and DC values to be set such that only the ion of interest has a stable trajectory through the quadrupole system. In one field direction, ions with a low *m/z* (light ions) will follow the alternating component of the field, gain energy and oscillate with increasingly large amplitudes until they hit one of the rods and are discharged. Only high-mass ions will hence be transmitted to the other end of the quadrupole. However, in the other direction of the field, ions with a high *m/z* (heavy ions) will be unstable because of the defocusing effect of the DC component. Lighter ions, on the other hand, will be stabilised by the alternating component and transferred to the other end of the quadrupole. Combining the two directions gives a mass filter that is suitable for mass analysis. Application of a suitable RF/DC ratio to the quadrupole can make it discriminate against both high- and low-mass ions to a desired degree. The mass filter can be set to transmit a single isotope or to scan over a wide *m/z* range.

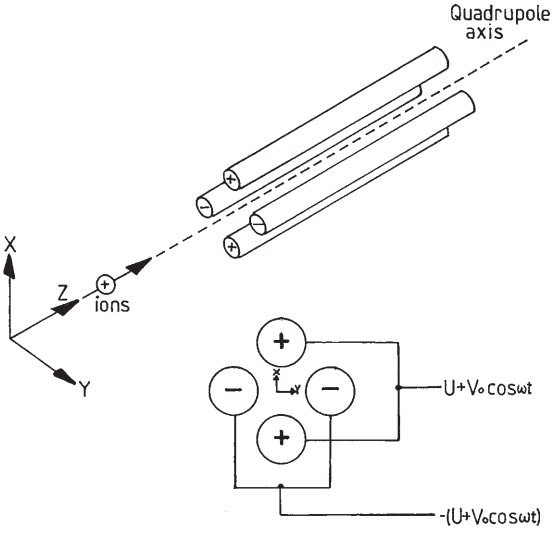


Figure 3. Schematic representation of the quadrupole mass filter and its connections.

The mass filter is a continuous analyser compared to the TOF analyser that has a pulsed nature. This feature makes the quadrupole highly compatible with continuous infusion sources, e. g. electrospray and liquid separation techniques such as high-performance liquid chromatography (HPLC) and CE. The sensitivity of the analyser for mass spectral acquisition is limited by the necessity to scan the quadrupole. The mass range of commercial instruments is now about *m/z* 4000. However, modern quadrupoles can transmit ions with *m/z* values above 10,000 [22] and by reducing the operating frequency, the mass range can be extended to *m/z* 45,000 [79]. Even though calibration is a straightforward process in mass filters (*m/z* depends linearly on RF and DC), mass accuracy has traditionally been poor. However, Green and co-workers have shown that accuracies in the range of 5 ppm can be achieved with careful operation [80]. A useful and educational review on the quadrupole mass filter was published in 1986 [81].

### The TOF mass analyser

A major development in TOF mass spectrometry came in the mid 1950s when Wiley and McLaren [26] described 'time-lag focusing' which markedly improves resolution. The principle of the TOF mass analyser is to measure the flight time of ions accelerated out of an ion source into a field-free drift tube to a detector. The flight time is related to the *m/z* values of the ions according to the following formula:

$$TOF = L\,(2U_{acc}\,e)^{-1/2}\,(m/z)^{1/2}$$

where $L$ is the drift length in the field-free region, $U_{acc}$ is the potential difference in the accelerating region, $e$ is the charge of an electron, $m$ is the mass of the ion and $z$ is its charge state. The TOF is usually measured from the time point at which the ions are accelerated out of the source to the time point when they reach the detector. The ions will separate in the TOF mass analyser according to their *m/z* ratios, light ions arriving at the detector earlier than heavy ions if they carry the same number of charges.

The ions initial spatial spread and initial velocity of the ions limit the resolving power of a TOF mass analyser. In a MALDI source, for example, the ionisation creates a burst of ions that will be at different distances from the detector (spatial spread) and have different kinetic energies. Ions with the same *m/z* value but with different distances to the detector will consequently be detected at

different time points, thus decreasing resolution. The same is true for ions with the same *m/z* value but with different initial kinetic energy. As mentioned earlier, powerful tools have been invented to compensate for these two problems.

Mamyrin and co-workers [77] introduced the mass reflectron in 1973. The mass reflectron compensates for the initial energy spread of the ions; a schematic of a TOF instrument with a reflector flight tube is shown in figure 4. In this type of instrument, the ions are not detected after a single pass through the field-free region. They are instead reflected back into the field-free drift tube by the electric field of the reflectron and detected at the same end of the tube as the initial acceleration region. Ions with a high energy will penetrate the reflectron deeper than ions with the same *m/z* value but with lower energy. Ions with the same *m/z* value can then be focused at the detector. In addition, the incorporation of a reflectron will extend the flight path with a resultant improved resolving power.

The second major instrumental development that has resulted in improved resolution in TOF is, as mentioned above, time-lag focusing. Wiley and McLaren [26] described an instrument with a two-stage accelerating region (cf. fig. 4) and, more recently, a number of related designs have been applied to MALDI TOF [74–76]. By introducing a delay between the end of the ionisation pulse and the application of the extraction pulse, the ions are sorted in space according to their original kinetic energy. An ion that starts further from the field-free region in the first electric field will then be accelerated to a slightly higher energy than an ion with the same *m/z* value but with a starting position closer to the field-free region. The second electric field accelerates the ions into the drift tube such that those with the same *m/z* value will arrive at the detector position at the same time.

The TOF analyser has a pulsed nature and has been used widely in conjunction with a pulsed ion source, e.g. MALDI [69, 70], although early TOF was interfaced to EI sources. TOF is considered to be a high-speed mass analyser. The basic cycle time for the TOF analyser is limited by the flight time of the heaviest ions. Since this is frequently in the 100-µs range, thousands of full mass spectra can be generated each second. TOF mass analysers can produce full spectra at high sample utilisation efficiencies, because all the *m/z* values in the flight tube at any time can be detected. These capabilities of fast spectral generation rate, high efficiency and a high duty cycle have made TOF a target for the return to continuous ionisation sources. Dodonov and co-workers [83] introduced the electrospray source to an orthogonal extraction TOF mass spectrometer in 1987. The continuously infused ions are pulsed at a high frequency perpendicular to their initial direction of movement and their flight times are measured. The group of Guilhaus [27, 84] has made great contributions to the area of orthogonal-acceleration TOF mass analysers. This instrumental set-up with coupling to continuous ionisation sources has opened up many new important applications for TOF. One new application is in the field of non-covalent interactions of proteins, where the theoretically unlimited mass range of the TOF mass analyser is of great value [85]. Nowadays, TOF mass analysers provide a valuable second stage in hybrid tandem mass spectrometers [32, 33].

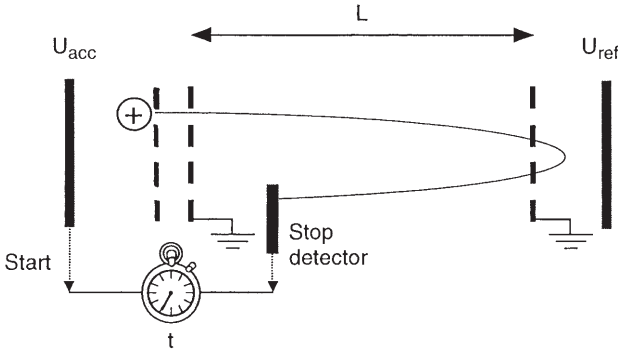


Figure 4. Schematic outline of a time-of-flight mass analyser with a two-stage accelerating region and a mass reflectron. [Modified from ref. 82.]

## Biological mass spectrometry

DNA sequencing is generally straightforward [86, 87] and gene sequence analysis using subsequent developments has led to the accumulation of gene structures, many of which have not been identified at the protein level [88–90]. The introduction of soft ionisation techniques has made it possible to transfer large molecules of biological origin to the gas phase for subsequent mass analysis. Consequently, the current goal for many groups using mass spectrometry in the post-genomic era is the identification and structure determination of proteins. It is important to be able to analyse the primary structure at the protein level and to determine post-translational modifications in order to correlate structure with biological function. However, mass spectrometry can also be used in DNA studies as described by Murray [91].

## Mass spectrometry in protein analysis

There are a number of ways to study primary, secondary, tertiary and quaternary structures of proteins using mass spectrometry, and different types of instruments can be used. The role of mass accuracy in these processes cannot be overestimated. Mass measurements of small biomolecules to high accuracies (a few ppm) can determine their elemental composition [92]. Monoisotopic masses of peptides and proteins with molecular weights in the range 5–8 kDa can be determined today, while for higher-mole-

cular-weight proteins, the average mass is generally determined (accurate mass measurements can be made even if the isotopic envelope remains unresolved [93]). Clarity when using such terms as monoisotopic mass and average mass is important. The monoisotopic mass of an element refers to its lightest stable isotope. Hence, the monoisotopic mass of a molecule is the sum of the monoisotopic masses of each element present. The chemical average mass of an element is the sum of the abundance-weighted masses of all its stable isotopes, e.g. 98.9% $^{12}C$ and 1.1% $^{13}C$ gives the isotope-weighted average mass of 12.011 for carbon. The average mass of a molecule, i.e. a protein, is then the sum of the chemical average masses of the elements present [5]. In mass spectrometry, the isotopic pattern of a compound is the best way to determine the charge state of the ions. A requirement for this determination is of course a sufficient resolving power of the mass spectrometer. The *m/z* separation between the individual isotopomers is 1/z, i.e. a doubly charged ion has 0.5 *m/z* units (or Th = Thomson) between its isotopomers, while a triply charged ion has isotopomers separated by 0.33 Th. The triply charged ion corresponding to the 19-residue N-terminal tryptic peptide of PYY is shown in figure 5 [94]. The abundance of the individual isotopomers is also of importance. The first isotopomer is generally the most abundant in peptides with a molecular weight less than about 1500–1700 Da. Above this, there is a higher probability that an individual ion will contain at least one $^{13}C$ and the second isotopomer becomes most abundant.

High-throughput protein identifications in proteomics projects are commonly performed by initial MALDI mass fingerprinting and database searches. The gel-separated proteins are in situ digested in the gel with a site-specific proteinase, often trypsin [95]. The resultant peptides are extracted and their masses determined by MALDI TOF. The generated peptide mass fingerprint is searched against protein or DNA sequence databases and compared to in silico digests of the corresponding sequences [42, 96–100]. Robotic systems that integrate everything from detection of gel spots, enzymatic digestion, mass spectrometric analysis and protein identification are currently being launched onto the market.

## Protein primary structure by mass spectrometry

The primary structure of proteins and peptides has traditionally been obtained by Edman degradation [101, 102]. A fully automated protein sequencer was developed in the late 1960s [103]. Edman sequencers can nowadays analyse peptides and proteins below the pmol level. Recently, efficient C-terminal sequencing has become available [104–106]. Especially for large proteins, N-terminal, C-terminal, or combined N- and C-terminal chemical analysis is still valuable in protein identification. Nevertheless, analysis of the amino acid sequence by tandem mass spectrometry is more sensitive and provides additional information concerning the identity and location of post-translational modifications. In addition, N-terminally blocked peptides are refractory to Edman degrada-

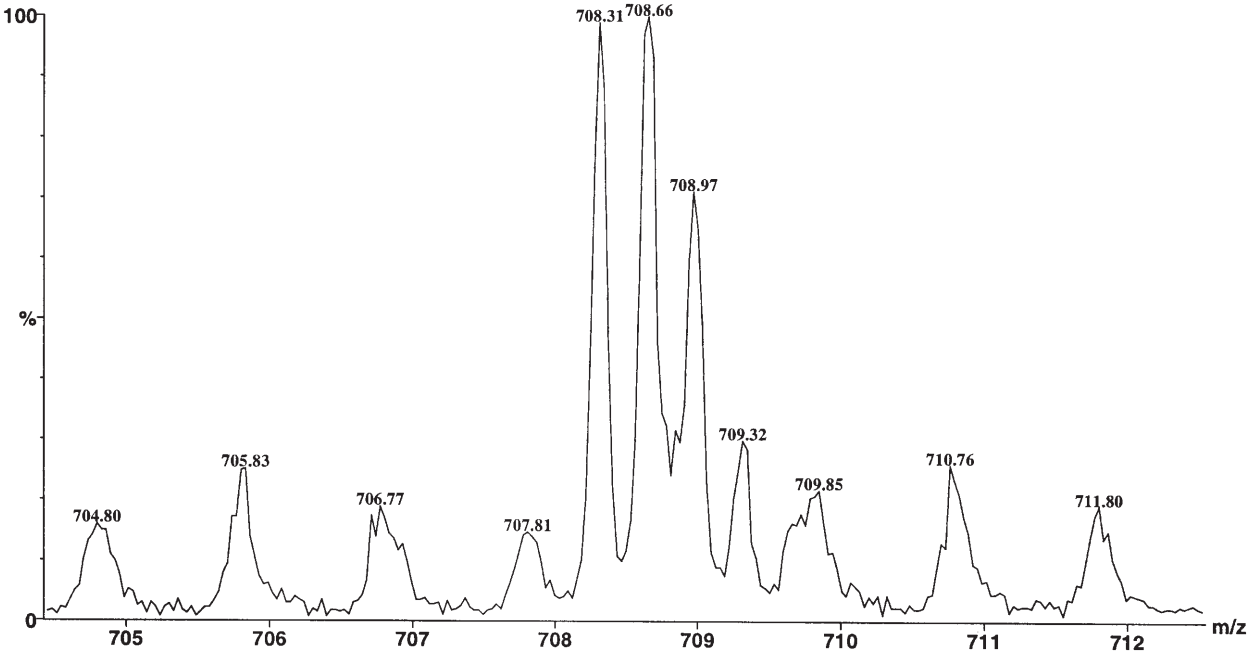


Figure 5. $[M+3H]^{3+}$ ion of the 19-residue tryptic peptide generated from PYY [94]. The mass separation (0.3) of the isotopomers indicates the charge state of the ion (3+).

tion but can be sequenced by MS/MS. In favourable cases, peptides can be analysed at the atto-to zeptomole level ($10^{-18}$–$10^{-21}$ mol) using a miniaturised electrospray ionisation source [64]. There are several approaches to obtain protein or peptide primary structures with mass spectrometry and these techniques all involve fragmentation of the protein/peptide in various manners. The most widely used method is collision-induced dissociation [107, 108] using a tandem mass spectrometer [38] and this technique will be discussed in detail below. Other approaches are ladder sequencing and post-source decay (PSD) using a MALDI TOF instrument. Peptide ladders, produced via enzymatic proteolysis (carboxypeptidase or aminopeptidase) or stepwise chemical degradation, can be mass measured by MALDI TOF [109, 110]. C-terminal sequences are generated by use of carboxypeptidases [111, 112], while a mixture of phenylisothiocyanate and phenylisocyanate can be used to produce N-terminal sequence ladders [109]. These methods can be employed directly on intact proteins but they usually require large sample amounts. PSD involves fragmentation of the peptide in the gas phase rather than fragmentation in the condensed phase as in ladder sequencing. The metastable decay in PSD is a fragmentation process that occurs after acceleration out of the ionisation source, but before the detector of the TOF instrument. In MALDI, the metastable ions are formed due to loss of small neutral molecules ($H_2O$ and $NH_3$) [113] or by peptide bond cleavages [114–116]. The products from the metastable decay can be detected in a reflectron TOF mass analyser where the reflectron voltage is stepwise decreased while the acceleration voltage is maintained constant. PSD is not generally employed for de novo sequencing since the product ion spectra are often incomplete. The resolution to select a parent ion in PSD is usually in the range of a few hundred at 'full-width at half-maximum height' (FWHM) and cannot be compared with conventional tandem mass spectrometry where a mass analyser can be operated at 1000–10,000 resolution (magnetic sector) or 1000–3000 resolution (Q-TOF). With an inefficient selection capability in combination with a poorly understood fragment ion chemistry, PSD is a difficult process to control, although an exciting new derivatisation chemistry may open up new applications [117].

CID using a tandem mass spectrometer has become the method of choice for sequence determination of peptides. Alternative methods, like surface-induced dissociation (SID) [118, 119] and electron capture dissociation (ECD) [120–122] are also available but they are not as widely used as CID. The idea behind MS/MS is to select a precursor ion with the first mass analyser (MS1) and to focus this ion into a collision cell. An inert gas is generally introduced into the cell and collisions occur between the inert gas and the precursor ion. During these collisions, the translational energy of the precursor ion will partly be converted into internal energy and ions become excited to an unstable state. Consequently, precursor ions dissociate (fragment) to product ions and this process is called CID. The masses of the product ions are determined in the second mass analyser (MS2), and the amino acid sequence of the peptide can be deduced.

The CID process has been employed for many decades, but the earliest CID studies for structural characterisation of peptides were carried out in the early 1980s [123, 124] and it is now considered a routine technique [38, 39]. CID can be performed at high or low collision energies. Dissociation of a peptide in a triple quadrupole or an ion trap is usually in the 10–100 eV collision energy range and considered as low-energy CID. Sector instruments are generally operated at keV energies and CID in this energy range is referred to as high-energy CID. Low-energy CID spectra of peptides tend to show product ions formed by small neutral losses and cleavage of the peptide bonds [39]. The high-energy CID process tends to dissociate other backbone bonds and also side chain bonds, and, consequently, generates more informative but also more complex product ion spectra [125, 126]. The appearance of CID spectra is very dependent on the charge state of the precursor, the number of basic amino acid residues and the presence or absence of acidic residues.

CID of a peptide results in a distribution of product ions formed by the dissociation of amide bonds in the peptide backbone. These product ions will either have the charge retained on the N-terminal side of the cleavage (B ions) or on the C-terminal side of the cleavage (Y″ ions) (fig. 6a). In the mid 1980s, Roepstorff and Fohlman [127] proposed a nomenclature for product ion formation and Biemann [128, 129] suggested modifications for positively charged product ions some years later. As mentioned above, low-energy CID of peptides tends to dissociate primarily the peptide bond, generating mainly Y″ and B ions. The general fragmentation mechanisms of peptides are now fairly well understood [130–135] but they are still a topic of investigation. Tryptic peptides (peptides generated after digestion of a protein with trypsin) have a basic residue (Arg or Lys) at the C terminus and formation of Y″ ions is favoured (fig. 6a). The general abundance of Y″ ions in product ion spectra can be a consequence of B ions tending to fragment further to A-type, lower B-type and internal immonium ions [130, 131, 136] (fig. 6a). Several strategies for differentiation between the different product ions have been employed. Derivatisation of the N and C terminus causes mass shifts of the B and Y″ ion series, respectively, but several amino acid side chains are also susceptible to derivatisation and the resultant spectra become complex. A successful approach is 'isotopic labelling' of the Y″ ions with $^{16}O/^{18}O$ [78]. The digestion of the protein is carried out in water with a 1:1 ratio of $^{16}O/^{18}O$ and the resulting proteolytic peptides (except the C-terminal one) become labelled by a 2 Da difference in their isotopic pattern.

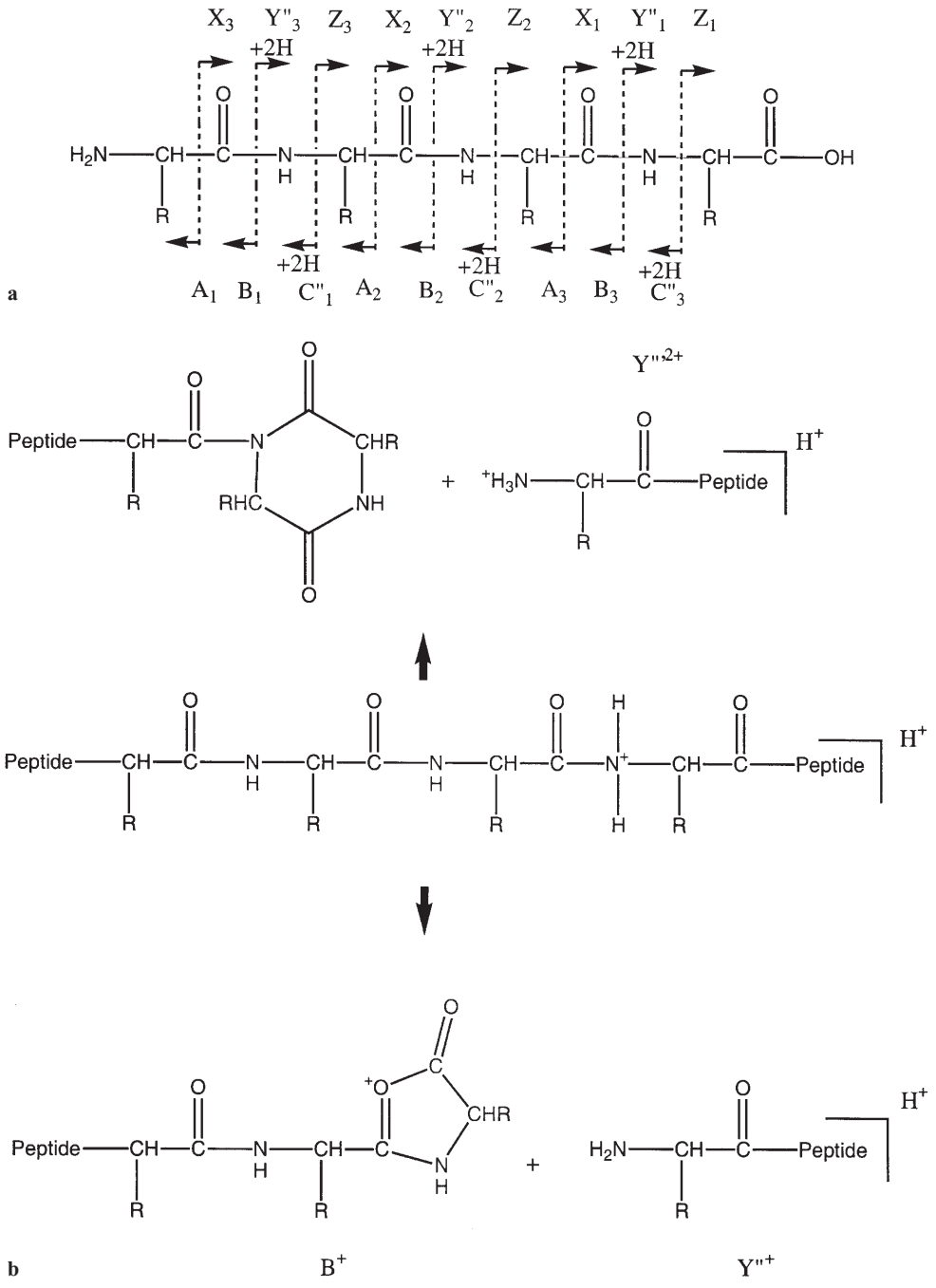


Figure 6. (*a*) Nomenclature for peptide fragmentation [127]. (*b*) Schematic representation of the cleavage of peptide bonds via the mobile-proton model.

Post-translational modifications, such as phosphorylation and glycosylation, can also be localised and determined by tandem mass spectrometry. The mass shift of the B or Y″ ion series following the modified amino acid residue is usually characteristic for a specific modification [40, 94].

The fragmentation of peptides at low-collision energy is believed to be charge mediated, meaning that a mobile proton is needed to initiate the fragmentation reaction. When one or more protons are added to a peptide, and non-selective cleavages occur across the entire peptide backbone, fragmentation is suggested to proceed via the 'mobile proton' or 'heterogeneously distributed proton' model. This model states that in the absence of a strongly basic residue (Arg) or when the number of added protons exceeds the number of strongly basic sites, the cleavages of various peptide bonds in protonated peptides occur following migration of a mobile proton to a carbonyl oxygen

or amide nitrogen of the peptide backbone. The low-energy pathways that form B or Y″ ions are then promoted by acid-catalysed charge-directed cleavage of the peptide bonds, initiated by a mobile proton. Electrospray ionisation generates multiply charged ions and, hence, the charges from a multiply charged precursor ion can be distributed among the product ions in a number of fashions. For example, a doubly charged precursor ion can dissociate to form a singly charged B ion and a singly charged Y″ ion, or a doubly charged B′ (or Y‴) ion and the corresponding neutral fragment (fig. 6b). Residues that tend to localise mobile protons, e.g. His, or donate protons (Asp, Glu) will have significant effects on peptide fragmentation [134, 135,137–139]. The localisation or donation of a mobile proton by these charged residues can give rise to a selective cleavage of the protonated peptide bond on their C-terminal side, generating an abundant B/Y″ ion pair. Pro residues are also known to promote facile cleavage of peptide bonds but on their N-terminal side, referred to as the proline effect [38, 39, 140]. This effect is due to the more basic characteristics of the amide bond present N-terminal in Pro, compared to the amide bonds contributed by the other amino acids. The peptide bond between Gln (or Glu) and Gly in a peptide has also been reported to generate abundant product ions upon CID [141, 142; W.J. Griffiths and A.P. Jonsson, unpublished data]. CID experiments of undigested polypeptides [120, 143, 144] have possible potential in high-throughput identification of proteins. Knowledge of the facile cleavage sites will make the interpretation of product ion spectra of intact proteins more efficient, since careful fragmentation leads to dissociation of the most labile peptide bonds (i.e. those mentioned above involving Pro residues, charged residues and Gln-Gly). The ability to predict the pattern of a product ion spectrum of a peptide or a protein is especially important for automated MS/MS experiments where computer programs are necessary to handle the spectral interpretation. Most algorithms available today do not consider the abundance of a particular product ion. Taking the relative abundance of product ions into account will add a new dimension to the interpretation programs.

As an alternative to peptide mass fingerprinting, identification of proteins can be performed by MS/MS. The amino acid sequence determined (or parts of the sequence) for a proteolytic peptide is submitted to a database search [36, 146–150]. The search engines are error tolerant but they demand precursor ion mass and product ion masses of good accuracy for correct positive identification.

## Studies of higher-order structures with mass spectrometry

Mass spectrometry can be used to study the primary structure as well as secondary, tertiary and quaternary structures of proteins. Hydrogen/deuterium (H/D) exchange has become a popular technique to determine which amino acid residues are located on the surface of large biomolecules [151–154]. Moreover, protein secondary structures (i.e. $\alpha$ helices and $\beta$ sheets) will exchange amide hydrogens at different rates. H/D experiments can be used to determine the secondary structure of parts of proteins and peptides. The H/D exchange is generally followed by pepsin digestion at low pH (to stop the exchange) and subsequent identification of the exchanged hydrogens in the proteolytic peptides by MS/MS.

Reduction of disulphide bonds (unfolding) and alkylation of the reduced Cys residues is commonly used to make proteins more accessible to proteolytic digestion. If the reduction step is omitted, alkylation only modifies the free Cys residues located on the surface of the protein. After digestion and determination of the labelled Cys residues using MS/MS, this type of experiment gives some information about the protein tertiary folding [4]. Similar methods for labelling of reactive amino acids (e.g. Tyr, Arg, His and Met) were carried out in the 1960s when the subsequent analysis was performed by paper chromatography [155].

With the advent of electrospray ionisation, determination of protein quaternary structure by mass spectrometry has become possible and the field of non-covalent interactions is growing rapidly. The electrospray interface must be optimised to preserve the non-covalent interactions between protein subunits, but still remove solvent adducts. Protein subunit interactions, receptor-ligand interactions and enzyme-substrate interactions can be analysed by mass spectrometry where the unlimited *m/z* range of the TOF analyser is of great benefit. Biomolecule complexes in the MDa range have been analysed by Robinson's group [22], many studies being carried out in parallel with nuclear magnetic resonance analysis.

## Recent developments in mass spectrometry

Mass spectrometry is currently developing very rapidly and its integration with biology is a major driving force. New instruments are being designed to fulfil the needs of automated analysis in the proteomic era. Both electrospray and MALDI are appropriate ionisation techniques for automatisation, and in combination with a tandem mass spectrometer they offer practical analytical tools for biological research. The following paragraphs will briefly describe three new instrument configurations, MALDI Q-TOF, MALDI TOF-TOF and a new-generation Q-TOF with both MALDI and nano-LC options, all with great potential for protein and peptide characterisation.

**MALDI Q-TOF**

In an attempt to combine the high speed and throughput of MALDI peptide mass fingerprinting with the specificity of tandem mass spectrometry, a hybrid MALDI Q-TOF has been constructed [156]. Generally, MALDI PMF is efficient when the number of matched peptide peaks is statistically significant, but the ability to characterise individual peptide ions has until recently been limited. With introduction of PSD, some MS/MS information can be obtained in a MALDI TOF experiment. However, the low precursor ion resolution, the poor mass accuracy, the low abundance of product ions together with the complexity of PSD spectra has made protein identification difficult using this method [157]. The abundance of product ions is often low in MALDI Q-TOF tandem mass spectra, due to the dissociation of singly charged tryptic peptides with preferences to maintain their charge at the basic C terminus (and thus not readily undergo charge-mediated reactions).

However, the high mass accuracy and the excellent resolving power in both MS and MS/MS modes of the MALDI Q-TOF allow protein identification either via PMF or from MS/MS spectra generated by CID of peptide precursor ions. The use of this kind of instrument eliminates the division of sample and the extra purification steps required for electrospray mass measurements, and both PMF and MS/MS experiments can be carried out on the same target. Another advantage is the time aspect: once the sample is on the target it will stay there and not be consumed within a few minutes. Nevertheless, the data quality obtained from this hybrid instrument is not yet as high as when using a separate nano-electrospray Q-TOF instrument in parallel with a MALDI TOF instrument.

**MALDI TOF-TOF**

This recently developed instrument employs TOF techniques for both precursor ion selection and product ion determination. The aim is to produce a tandem mass spectrometer with resolution and mass accuracy comparable to magnetic sector instruments, while maintaining the speed and the mass range of a MALDI-TOF [158]. A novel timed-ion selector has been developed which allows high-resolution selection of an ion of a particular *m/z* to be transmitted from the first TOF analyser while all other ions are rejected. This process employs two deflection gates that can be rapidly opened and closed when an ion of interest reaches the entrance. The first gate rejects low-mass ions and the second gate rejects high-mass ions. The resolving power of the selector is sufficient to select a single isotopomer of a peptide up to an *m/z* value of 2000, without significant transmission losses. Ions separated by the first TOF mass analyser can be selected for PSD or high-energy CID in the collision cell. Both TOF analysers are operated in delayed extraction mode and they allow complete acquisition of a product ion spectrum in a single run. The resolving power measured is in the range 2000–10,000 (FWHM).

The major drawback of this instrument is the poor sensitivity for high-mass product ions. This limits the ability to provide complete sequence information from an MS/MS

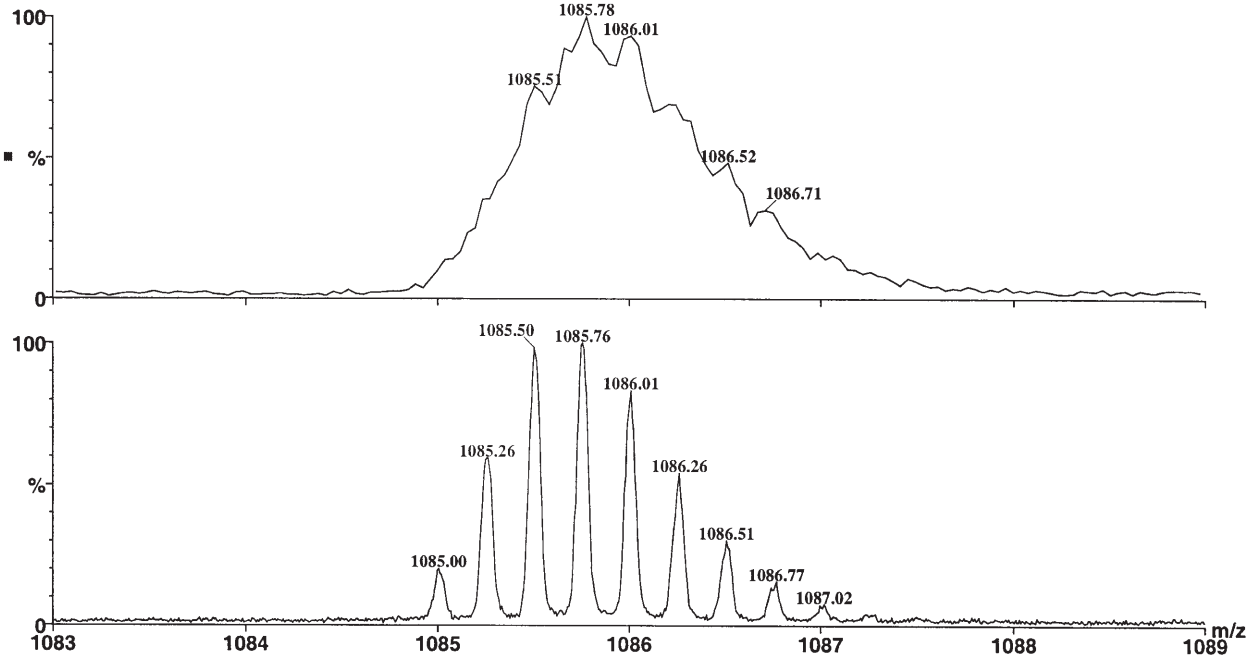


Figure 7. Comparison of the mass resolving power of the Q-TOF I (upper panel) and the Q-TOF Ultima 'W' (lower panel) for the quadruply protonated peptide $T_2$ from tryptic digestion of PRP-1 [40].

experiment of a peptide with a molecular weight above approximately 1800 Da. The sensitivity is also a problem when analysing protein digests at a level of 1 pmol or less using conventional sample preparation techniques. In this case, useful sequence ions are only generated from the most intense signals in the spectra. Immonium ions, on the other hand, can be detected with good sensitivity. The instrument design appears interesting, and optimisation of this prototype may well lead to an optimal tool for the routine characterisation of protein digests. Nevertheless, the singly charged peptide ions generally produced by a MALDI source decrease the possibility to achieve complete sequence information and, hence, an orthogonal electrospray source would be interesting to interface to this type of instrument.

## MALDI/LC-MS/MS–MS Q-TOF

A multi-inlet MS/MS-MS platform for both MALDI and electrospray ionisation has recently been launched onto the market (Q-TOF Ultima Global; Micromass, U.K.). This is the third generation of the Q-TOF family and the set-up of this new mass spectrometer features the ability to change between ionisation modes without venting the instrument. The mass analysers have an upper *m/z* mass range of 20,000 and resolution of 10,000 (FWHM) in the classical 'V' TOF mode. The instrument has an optional 'W' TOF technology for maximum resolution (>20,000 FWHM) but at some cost in sensitivity. The 'W' TOF analyser effectively doubles the ion flight path by the introduction of a small ion mirror that sends the ions on a second trip around the TOF analyser while maintaining spatial focus [159]. A comparison between the resolving power of Q-TOF I and Q-TOF Ultima 'W' for the $[M+4H]^{4+}$ ion of the tryptic peptide $T_2$ of PRP-1 [40] is presented in figure 7. The Q-TOF Ultima family also provides instruments with only MALDI or electrospray ionisation sources and these instruments have an extra degree of sensitivity. In addition, the Q-TOF Ultima instruments provide a new inlet system, with a larger opening and additional pumping, for optimal sample utilisation and efficiency. The possibility to switch between MALDI and electrospray ionisation on a single instrument with low-energy CID and outstanding resolution makes the instrument very attractive for research groups with interests in protein and peptide characterisation.


1 Jonsson A. P. (2001) Mass spectrometry in protein structure analysis. PhD thesis, Karolinska Institutet, Stockholm
2 Creighton T. E. (1993) Proteins: Structure and Molecular Properties, Freeman, New York
3 Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:** 680–685
4 Jonsson A. P., Carlquist M., Husman B., Ljunggren J., Jörnvall H., Bergman T. et al. (1999) Structural analysis of the thyroid hormone receptor ligand binding domain: studies using a quadrupole time-of-flight mass spectrometer. Rapid Commun. Mass Spectrom. **13:** 1782–1791
5 Carr S. A., Burlingame A. L. and Baldwin M. A. (2000) The meaning and usage of the terms monoisotopic mass, average mass, mass resolution, and mass accuracy for measurements of biological molecules. In: Mass Spectrometry in Biology and Medicine, pp. 553–561, Carr S. A., Burlingame A. L. and Baldwin M. A. (eds), Humana, Totowa N. J.
6 He F., Hendrickson C. L. and Marshall A. G. (2001) Baseline mass resolution of peptide isobars: a record for molecular mass resolution. Anal. Chem. **73:** 647–650
7 Thomson J. J. (1913) Rays of Positive Electricity and Their Applications to Chemical Analysis, Longmans Green, London
8 Griffiths I. W. (1997) J. J. Thomson – the centenary of his discovery of the electron and his invention of mass spectrometry. Rapid Commun. Mass Spectrom. **11:** 2–16
9 Aston F. W. (1919) A positive ray spectrograph. Philos. Mag. **38:** 707–714
10 Aston F. W. (1919) The distribution of intensity along the positive ray parabolas of atoms and molecules of hydrogen and its possible explanation. Proc. Camb. Phil. Soc. **19:** 317–323
11 Aston F. W. (1933) Mass spectra and isotopes, Arnold, London
12 Siuzdak G. (1994) The emergence of mass spectrometry in biochemical research. Proc. Natl. Acad. Sci. USA **91:** 11290–11297
13 Carr S. A., Hemling M. E., Bean M. F. and Roberts G. D (1991) Integration of mass spectrometry in analytical biotechnology. Anal. Chem. **63:** 2802–2824
14 Fenselau C. (1992) Tandem mass spectrometry: the competitive edge for pharmacology. Annu. Rev. Pharmacol. Toxicol. **32:** 555–578
15 Easterling M. L., Colangelo C. M., Scott R. A. and Amster I. J. (1998) Monitoring protein expression in whole bacterial cells with MALDI time-of-flight mass spectrometry. Anal. Chem. **70:** 2704–2709
16 Roepstorff P. (1997) Mass spectrometry in protein studies from genome to function. Curr. Opin. Biotech. **8:** 6–13
17 Pandey A. and Mann M. (2000) Proteomics to study genes and genomes. Nature **405:** 837–846
18 Yamashita M. and Fenn J. B. (1984) Electrospray ion source: another variation on the free-jet theme. J. Phys. Chem. **88:** 4451–4459
19 Yamashita M. and Fenn J. B. (1984) Negative ion production with the electrospray ion source. J. Phys. Chem. **88:** 4671–4675
20 Alexandrov M. L., Gall L. N., Krasnov N. V., Nikolaev V. I., Pavlenko V. A., Shkurov V. A. et al. (1984) Direct coupling of a microcolumn liquid chromatograph and a mass spectrometer. Bioorg. Khim. **10:** 710–712
21 Karas M., Bachmann D., Bahr U. and Hillenkamp F. (1987) Matrix-assisted ultraviolet laser desorption of non-volatile compounds. Int. J. Mass Spectrom. Ion Proces. **78:** 53–68
22 Rostom A. A. and Robinson C. V. (1999) Detection of the intact GroEL chaperonin assembly by mass spectrometry. J. Am. Chem. Soc. **121:** 4718-4719
23 Paul V. W. and Steinwedel H. (1953) Ein neues Massenspektrometer ohne Magnetfeld. Z. Naturforsch. **8a:** 448–450
24 Dempster A. J. (1918) A new method of positive ray analysis. Phys. Rev. **11:** 316–324
25 Johnson E. G. and Nier A. O. (1953) Angular aberrations in sector-shaped electromagnetic lenses for focusing beams of charged particles. Phys. Rev. **91:** 10–17
26 Wiley W. C. and McLaren I. H. (1955) Time-of-flight mass spectrometer with improved resolution. Rev. Sci. Instrum. **26:** 1150–1157
27 Guilhaus M., Mlynski V. and Selby D. (1997) Perfect timing: time of flight mass spectrometry. Rapid Commun. Mass Spectrom. **11:** 951–962

28 Paul W. and Steinwedel H. (1960) US Patent number 2939952
29 March R. E. (1997) An introduction to quadrupole ion trap mass spectrometry. J. Mass Spectrom. **32:** 351–369
30 Comisarow M. B. and Marshall A. G. (1974) Fourier transform ion cyclotron resonance spectroscopy. Chem. Phys. Lett. **25:** 282–283
31 Yost R. A. and Enke C. G. (1978) Selected ion fragmentation with a tandem quadrupole mass spectrometer. J. Am. Chem. Soc. **100:** 2274–2275
32 Morris H. R., Paxton T., Dell A., Langhorn B., Berg M., Bordoli R. S. et al. (1996) High sensitivity collisionally-activated decomposition tandem mass spectrometry on a novel quadrupole/orthogonal-acceleration time-of-flight mass spectrometer. Rapid Commun. Mass Spectrom. **10:** 889–896
33 Bateman R. H., Green M. R., Scott G. and Clayton E (1995) A combined magnetic sector-time-of-flight mass spectrometer for structural determination studies by tandem mass spectrometry. Rapid Commun. Mass Spectrom. **9:** 1227–1233
34 Hoffmann E. de (1996) Tandem mass spectrometry: a primer. J. Mass Spectrom. **31:** 129–137
35 McLafferty F. W. (1983) Tandem Mass Spectrometry. Wiley, New York
36 Mann M. and Wilm M. (1994) Error-tolerant identification of peptides in sequence databases by peptide sequence tags. Anal. Chem. **66:** 4390–4399
37 Yates J. R. III, Eng J. and McCormack A. L. (1995) Mining genomes: correlating tandem mass spectra of modified and unmodified peptides to sequences in nucleotide databases. Anal. Chem. **67:** 3202–3210
38 Hunt D. F., Yates J. R. III, Shabanowitz J., Winston S. and Hauer C. R. (1986) Protein sequencing by tandem mass spectrometry. Proc. Natl. Acad. Sci. USA **83:** 6233–6237
39 Papayannopoulos I. A. (1995) The interpretation of collision-induced dissociation tandem mass spectra of peptides. Mass Spectrom. Rev. **14:** 49–73
40 Jonsson A. P., Griffiths W. J., Bratt P., Johansson I., Strömberg, N., Jörnvall H. et al. (2000) A novel Ser *O*-glucuronidation in acidic proline-rich proteins identified by tandem mass spectrometry. FEBS Lett. **475:** 131–134
41 Annan R. S., Huddleston M. J., Verma R., Deshaies R. J. and Carr S. A. (2001) A multi-dimensional electrospray MS-based approach to phosphopeptide mapping. Anal. Chem. **73:** 393–404
42 Mann M., Højrup P. and Roepstorff P. (1993) Use of mass spectrometric molecular weight information to identify proteins in sequence databases. Biol. Mass Spectrom. **22:** 338–345
43 Roepstorff P. (2000) MALDI-TOF mass spectrometry in protein chemistry. In: Proteomics in Functional Genomics: Protein Structure Analysis, pp. 81–97, Jollès P. and Jörnvall H. (eds), Birkhäuser, Basel
44 Collins F., Patrinos A., Jordan E., Chakravarti A., Gesteland R. and Walters L. (1998) New goals for the U.S. human genome project: 1998–2003. Science **282:** 682–689
45 Munson M. S. B. and Field F. H. (1966) Chemical ionization mass spectrometry. I. General introduction. J. Am. Chem. Soc. **88:** 2621–2630
46 Barber M., Bordoli R. S., Sedgwick R. D. and Tyler A. N. (1981) Fast atom bombardment of solids (FAB): a new ion source for mass spectrometry. J. Chem. Soc. Chem. Commun. 325–327
47 Barber M., Bordoli R. S., Elliott G. J., Sedgwick R. N. and Tyler A. N. (1982) Fast atom bombardment mass spectrometry. Anal. Chem. **54:** 645A–657A
48 Dole M., Mach L. L., Hines R. L., Mobley R. C., Ferguson L. D. and Alice M. B. (1968) Molecular beams of macroions. J. Chem. Phys. **49:** 2240–2247
49 Meng C. K., Mann M. and Fenn J. B. (1988) Of protons or proteins. Z. Phys. D. **10:** 361–368
50 Fenn J. B., Mann M., Meng C. K., Wong S. F. and Whitehouse C. M. (1989) Electrospray ionization for mass spectrometry of large biomolecules. Science **246:** 64–71
51 Taylor G. (1964) Disintegration of water drops in an electric field. Proc. R. Soc. Lond A**280:** 383–397
52 Rayleigh, Lord (1882) On the equilibrium of liquid conducting masses charged with electricity. Philos. Mag. **14:** 184–186
53 Kebarle P. and Ho Y. (1997) On the mechanism of electrospray mass spectrometry. In: Electrospray Ionization Mass Spectrometry: Fundamentals, Instrumentation and Applications, pp. 3–63, Cole R. B. (ed), Wiley, New York
54 Iribarne J. V. and Thomson B. A. (1976) On the evaporation of small ions from charged droplets. J. Chem. Phys. **64:** 2287–2294
55 Fenn J. B. (1993) Ion formation from charged droplets: roles of geometry, energy, and time. J. Am. Soc. Mass Spectrom. **4:** 524–535
56 Kebarle P. and Tang L. (1993) From ions in solution to ions in the gas phase. Anal. Chem. **65:** 972A–986A
57 Cole R. B. (ed) (1997) Electrospray Ionization Mass Spectrometry: Fundamentals, Instrumentation and Applications, Wiley, New York
58 Kebarle J. (2000) A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry. J. Mass Spectrom. **35:** 804–817
59 Cole R. B. (2000) Some tenets pertaining to electrospray ionization mass spectrometry. J. Mass Spectrom. **35:** 763–772
60 Bruins A. P. (1991) Liquid chromatography-mass spectrometry with ionspray and electrospray interfaces in pharmaceutical and biomedical research. J. Chromatogr. **554:** 39–46
61 Barnidge D. R., Nilsson S. and Markides K. E. (1999) A design for low-flow sheathless electrospray emitters. Anal. Chem. **71:** 4115–4118
62 Gale D. C. and Smith R. D. (1993) Small volume and low flow-rate electrospray ionization mass spectrometry of aqueous samples. Rapid Commun. Mass Spectrom. **7:** 1017–1021
63 Emmett M. R. and Caprioli R. M. (1994) Micro-electrospray mass spectrometry: ultra-high-sensitivity analysis of peptides and proteins. J. Am. Soc. Mass Spectrom. **5:** 605–613
64 Andren P. E., Emmett M. R. and Caprioli R. M. (1994) Micro-electrospray: zeptomole/attomole per microliter sensitivity for peptides. J. Am. Soc. Mass Spectrom. **5:** 867–869
65 Wilm M. S. and Mann M. (1994) Electrospray and Taylor-cone theory: Dole's beam of macromolecules at last? Int. J. Mass Spectrom. Ion Proces. **136:** 167–180
66 Wilm M. and Mann M. (1996) Analytical properties of the nano-electrospray ion source. Anal. Chem. **68:** 1–8
67 Richards D., Verrier H., Major H. and Rontree J. (1998) The direct LC-MS analysis of drug candidates in cell culture media using a dual orthogonal sampling API interface, p. 75, Proc. 16th IMMS, 4–6 May, Budapest
68 Ferrige A. G., Seddon M. J., Green B. N., Jarvis S. A. and Skilling J. (1992) Disentangling electrospray spectra with maximum entropy. Rapid Commun. Mass Spectrom. **6:** 707–711
69 Karas M. and Hillenkamp F. (1988) Ultraviolet laser desorption of ions above 10 kDa. In: Abstracts, 11th Int. Mass Spectrom. Conf., Bordeaux, 29 August – 2 September
70 Karas M. and Hillenkamp F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10 000 Daltons. Anal. Chem. **60:** 1299–2301
71 Cotter R. J. (1997) Time-of-Flight Mass Spectrometry: Instrumentation and Applications in Biological Research, pp. 131–134, Am. Chemical Society, Washington, D.C.
72 Vorm O. and Mann M. (1994) Improved mass accuracy in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of peptides. J. Am. Soc. Mass Spectrom. **5:** 955–958

73 Gobom J., Nordhoff E., Mirgorodskaya E., Ekman R. and Roepstorff P. (1999) Sample purification and preparation technique based on nano-scale reversed phase columns for the sensitive analysis of complex peptide mixtures by matrix-assisted laser desorption/ionization mass spectrometry. J. Mass Spectrom. **34:** 105–116

74 Colby S. M., King T. B. and Reilly J. P. (1994) Improving the resolution of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry by exploiting the correlation between ion position and velocity. Rapid Commun. Mass Spectrom. **8:** 865–868

75 Vestal M. L., Juhasz P. and Martin S. A. (1995) Delayed extraction matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Commun. Mass Spectrom. **9:** 1044–1050

76 Brown R. S. and Lennon J. J. (1995) Mass resolution improvement by incorporation of pulsed ion extraction/ionization linear time-of-flight mass spectrometry. Anal. Chem. **67:** 1998–2003

77 Mamyrin B. A., Karataev V. I., Shmikk D. V. and Zagualin V. A. (1973) The mass-reflectron, a new non-magnetic time-of-flight mass spectrometer with high resolution. Sov. Phys. JETP. **37:** 45–48

78 Shevchenko A., Chernushevich I., Ens W., Standing K. G., Thomson B., Wilm M. et al. (1997) Rapid 'de novo' peptide sequencing by a combination of nano electrospray, isotopic labelling and a quadrupole/time-of-flight mass spectrometer. Rapid Commun. Mass Spectrom. **11:** 1015–1024

79 Winger B. E., Light-Wahl K. J., Ogorzalek Loo R. R., Udseth H. R. and Smith R. D. (1993) Observation and implications of high mass-to-charge ratio ions from electrospray ionisation mass spectrometry. J. Am. Soc. Mass Spectrom. **4:** 536–545

80 Tyler A. N., Clayton E. and Green B. N. (1996) Exact mass measurements of polar organic molecules at low resolution using electrospray ionization and a quadrupole mass spectrometer. Anal. Chem. **68:** 3561–3569

81 Miller P. E. and Bonner Denton M. (1986) The quadrupole mass filter: basic operating concepts. J. Chem. Educ. **63:** 617–622

82 Håkansson K. (2000) Method and technique development in peptide and protein mass spectrometry. PhD thesis, Uppsala University

83 Dodonov A. F., Chernushevich I. V., Dodonona T. F., Raznikov, V. V. and Talrose V. L. (1987) USSR patent number 1681340A1

84 Guilhaus M., Selby D. and Mlynski V. (2000) Orthogonal acceleration time-of-flight mass spectrometry. Mass Spectrom. Rev. **19:** 65–107

85 Tito M. A., Tars K., Valegard K., Hajdu J. and Robinson C.V. (2000) Electrospray time-of-flight mass spectrometry of the intact MS2 virus capsid. J. Am. Chem. Soc. **122:** 3550–3551

86 Maxam A. M. and Gilbert W. (1977) A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA **74:** 560–564

87 Sanger F., Nicklen S. and Coulson A. R. (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:** 5463–5467

88 Hattori M., Fujiyama A., Taylor T. D., Watanabe H., Yada T., Park H.-S. et al. (2000) The DNA sequence of human chromosome 21. Nature **405:** 311–319

89 Lander E. S., Linton L. M., Birren B., Nusbaum C., Zody M. C., Baldwin J. et al. (2001) Initial sequencing and analysis of the human genome. Nature **409:** 860–921

90 Venter J. C., Adams M. D., Myers E. W., Li P. W., Mural R. J., Sutton G. G. et al. (2001) The sequence of the human genome. Science **291:** 1304–1351

91 Murray K. K. (1996) DNA sequencing by mass spectrometry. J. Mass Spectrom. **31:** 1203–1215

92 Mata de Urquiza A., Liu S., Sjöberg M., Zetterström R. H., Griffiths W., Sjövall J. et al. (2000) Docosahexaenoic acid, a ligand for the retinoid X receptor in mouse brain. Science **290:** 2140–2144

93 Yergey J., Heller D., Hansen G., Cotter R. J. and Fenselau C. (1983) Isotopic distribution in mass spectra of large molecules. Anal. Chem. **55:** 353–356

94 Chen Z., Eriste E., Jonsson A. P., Norberg Å., Nepomuceno D., Lovenberg T. W. et al. (2001) $Ser^{13}$-phosphorylated PYY from porcine intestine with a potent biological activity. FEBS Lett. **492:** 119–122

95 Hellman U., Wernstedt C., Gonez J. and Heldin C. H. (1995) Improvement of an "in-gel" digestion procedure for the micropreparation of internal protein fragments for amino acid sequencing. Anal. Biochem. **224:** 451–455

96 Henzel W. J., Billeci T. M., Stultz J. T., Wong S. C., Grimley C. and Wantanabe C. (1993) Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. Proc. Natl. Acad. Sci. USA **90:** 5011–5015

97 Henzel W. J., Stultz J. T. and Wantanabe C. (1989) Paper presented at the 37th ASMS conference on mass spectrometry and allied topics, Seattle, Wash.

98 Pappin D. J. C., Højrup P. and Bleasby A. J. (1993) Protein identification by peptide mass fingerprinting. Curr. Biol. **3:** 327–332

99 James P., Quadroni M., Carafoli E. and Gonnet G. (1993) Protein identification by mass profile finger printing. Biochem. Biophys. Res. Commun. **195:** 58–64

100 Yates J. R., Speicher S., Griffin P. R. and Hunkapiller T. (1993) Peptide mass maps: a highly informative approach to protein identification. Anal. Biochem. **214:** 397–408

101 Edman P. (1949) A method for the determination of the amino acid sequence in peptides. Arch. Biochem. **22:** 475–476

102 Edman P. (1950) Method for determination of the amino acid sequence in peptides. Acta Chem. Scand. **4:** 283–293

103 Edman P. and Begg G. (1967) A protein sequenator. Eur. J. Biochem. **1:** 80–91

104 Boyd V. L., Bozzini M., Guga P. J., DeFranco R. J. and Yuan P.-M. (1995) Activation of the carboxy terminus of a peptide for carboxy-terminal sequencing. J. Org. Chem. **60:** 2581–2587

105 Boyd V. L., Bozzini M., Zon G., Noble R. L. and Mattaliano R. J. (1992) Sequencing of peptides and proteins from the carboxy terminus. Anal. Biochem. **206:** 344–352

106 Bergman T., Cederlund E. and Jörnvall H. (2001) Chemical C-terminal protein sequence analysis: improved sensitivity, length of degradation, proline passage, and combination with Edman degradation. Anal. Biochem. **290:** 74–82

107 Jennings K. R. (1968) Collision-induced decomposition of aromatic molecular ions. Int. J. Mass Spectrom. Ion Phys. **1:** 227–235

108 McLafferty F. W., Bente P. F. III, Kornfeld R., Tsai S.-C. and Howe I. (1973) Collision activation spectra of organic ions. J. Am. Chem. Soc. **95:** 2120–2129

109 Chait B. T., Wang R., Beavis R. C. and Kent S. B. H. (1993) Protein ladder sequencing. Science **262:** 89–92

110 Bergman T. (2000) Ladder sequencing. In: Proteomics in Functional Genomics: Protein Structure Analysis, pp. 133–144, Jollès P. and Jörnvall H. (eds), Birkhäuser, Basel

111 Patterson D. H., Tarr G. E., Regnier F. E. and Martin S. A. (1995) C-terminal ladder sequencing via matrix assisted laser desorption/ionization mass spectrometry coupled with carboxypeptidase Y time-dependent and concentration-dependent digestions. Anal. Chem. **67:** 3971–3978

112 Bonetto V., Bergman A.-C., Jörnvall H. and Sillard R. (1997) C-terminal sequence analysis of peptides and proteins using carboxypeptidases and mass spectrometry after derivatization of Lys and Cys residues. Anal. Chem. **69:** 1315–1319

113 Hill J. A., Annan R. S. and Biemann K. (1991) Matrix-assisted laser desorption ionization with a magnetic mass spectrometer. Rapid Commun. Mass Spectrom. **5:** 395–399

114 Spengler B., Kirsch D. and Kauffmann R. (1991) Metastable decay of peptides and proteins in matrix-assisted laser desorption mass spectrometry. Rapid Commun. Mass Spectrom. **5:** 198–202

115 Kauffmann R., Kirsch D. and Spengler B. (1994) Sequencing of peptides in a time-of-flight mass spectrometer: evaluation of postsource decay following matrix assisted laser desorption ionisation (MALDI). Int. J. Mass Spectrom. Ion Proces. **131:** 355

116 Gevaert K. and Vandekerckhove J. (2000) Protein identification methods in proteomics. Electrophoresis **21:** 1145–1154

117 Keough T., Youngquist R. S. and Lacey M. P. (1999) A method for high-sensitivity peptide sequencing using postsource decay matrix-assisted laser desorption ionization mass spectrometry. Proc. Natl. Acad. Sci. USA **96:** 7131–7136

118 Cooks R. G., Ast T. and Mabud M.-D. A. (1990) Collisions of polyatomic ions with surfaces. Int. J. Mass Spectrom. Ion Proces. **100:** 209–265

119 Dongré A. R., Somogyi Á. and Wysocki V. H. (1996) Surface-induced dissociation: an effective tool to probe structure, energetics and fragmentation mechanisms of protonated peptides. J. Mass Spectrom. **31:** 339–350

120 Zubarev R. A., Kelleher N. L. and McLafferty F. W. (1998) Electron capture dissociation of multiply charged protein cations: a nonergodic process. J. Am. Chem. Soc. **120:** 3265–3266

121 Zubarev, R. A., Krunger, N. A., Fridriksson, E. K., Lewis, M. A., Horn, D. M., Carpenter, B. K. et al. (1999) Electron capture dissociation of gaseous multiply-charged proteins is favoured at disulfide bonds and other sites of high hydrogen atom affinity. J. Am. Chem. Soc. **121:** 2857–2862

122 Kelleher N. L., Zubarev R. A., Bush K., Furie B., Furie B. C., McLafferty F. W. et al. (1999) Location of labile posttranslational modifications by electron capture dissociation: the case of γ-carboxyglutamic acid. Anal. Chem. **71:** 4250–4253

123 Hunt D. F., Buko A. M., Ballard J. M., Shabanowitz J. and Giordani A. B. (1981) Sequence analysis of polypeptides by collision activated dissociation on a triple quadrupole mass spectrometer. Biomed. Mass Spectrom. **8:** 397–408

124 Hunt D. F., Bone W. M., Shabanowitz J., Rhodes J. and Ballard J. M. (1981) Sequence analysis of oligopeptides by secondary ion/collision activated dissociation mass spectrometry. Anal. Chem. **53:** 1704–1706

125 Johnson R. S., Martin S. A. and Biemann K. (1988) Collision-induced fragmentation of (M+H)+ ions of peptides: side chain specific sequence ions. Int. J. Mass Spectrom. Ion Proces. **86:** 137–154

126 Johnson R. S., Martin S. A., Biemann K., Stultz J. T. and Watson J. T. (1987) Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. Anal. Chem. **59:** 2621–2625

127 Roepstorff P. and Fohlman J. (1984) Proposal for a common nomenclature of sequence ions in mass spectra of peptides. Biomed. Mass Spectrom. **11:** 601

128 Biemann K. (1988) Contributions of mass spectrometry to peptide and protein structure. Biomed. Environ. Mass Spectrom. **16:** 99–111

129 Biemann K. (1990) Nomenclature for peptide fragment ions (positive ions). Methods Enzymol. **193:** 886–887

130 Yalcin T., Csizmadia I. G., Peterson M. R. and Harrison A. G. (1996) The structure and fragmentation of $B_n$ (n>3) ions in peptide spectra. J. Am. Soc. Mass Spectrom. **7:** 233–242

131 Yalcin T., Khouw C., Csizmadia I. G., Peterson M. R. and Harrison A. G. (1995) Why are B ions stable species in peptide spectra? J. Am. Soc. Mass Spectrom. **6:** 1165–1174

132 Nold M. J., Wesdemiotis C., Yalcin T. and Harrison A. G. (1996) Amide bond dissociation in protonated peptides: structures of the N-terminal ionic and neutral fragments. Int. J. Mass Spectrom. Ion Proces. **164:** 137–153

133 Nair H., Somogyi Á. and Wysocki V. H. (1996) Effect of alkyl substitution at the amide nitrogen on amide bond cleavage: electrospray ionization/surface-induced dissociation fragmentation of substance P and two alkylated analogs. J. Mass Spectrom. **31:** 1141–1148

134 Summerfield S. G., Whiting A. and Gaskell S. J. (1997) Intra-ionic interactions in electrosprayed peptide ions. Int. J. Mass Spectrom. Ion Proces. **162:** 149–161

135 Tsaprailis G., Hair H., Somogyi Á., Wysocki V. H., Zhong W., Futrell J. H. et al. (1999) Influence of secondary structure on the fragmentation of protonated peptides. J. Am. Chem. Soc. **121:** 5142–5154

136 Ambihapathy K., Yalcin T., Leung H.-W. and Harrsison A. G. (1997) Pathways to immonium ions in the fragmentation of protonated peptides. J. Mass Spectrom. **32:** 209–215

137 Tsaprailis, G., Somogyi, Á., Nikolaev, E. N. and Wysocki, V. H. (2000) Refining the model for selective cleavage at acidic residues in arginine-containing protonated peptides. Int. J. Mass Spectrom. **195/196:** 467–479

138 Tsaprailis G. and Wysocki V. H. (2000) Another selective cleavage in peptides: a common mechanism for the formation of complementary $b^+/y^+$ or $b_2^+$ ions at protonated histidine. Proc. 48th ASMS Conf. on Mass Spectrometry and Allied Topics, Long Beach, Calif.

139 Gu C., Tsaprailis G., Breci L. and Wysocki V. H. (2000) Selective gas-phase cleavage at the peptide bond C-terminal to aspartic acid in fixed-charge derivatives of Asp-containing peptides. Anal. Chem. **72:** 5804–5813

140 Wang Y., Johansson J. and Griffiths W. J. (2000) Characterisation of variant forms of prophenin: mechanistic aspects of the fragmentation of proline-rich peptides. Rapid Commun. Mass Spectrom. **14:** 2182–2202

141 Jonsson A. P., Bergman T., Bratt P., Strömberg N., Jörnvall H. and Griffiths W. J. (2001) Gln-Gly cleavage: correlation between collision-induced dissociation and biological degradation. J. Am. Soc. Mass Spectrom. **12:** 337–342

142 Jonsson A. P., Bergman T., Jörnvall H. and Griffiths W. J. (2001) Gln-Gly cleavage: a dominant dissociation site in the fragmentation of protonated peptides. Rapid Commun. Mass Spectrom. **15:** 713–720

143 Loo J. A., Edmonds C. G. and Smith R. D. (1993) Tandem mass spectrometry of very large molecules. 2. Dissociation of multiply charged proline-containing proteins from electrospray ionization. Anal. Chem. **65:** 425–438

144 Witkowska H. E., Green B. N., Morris M. and Shackleton C. H. L. (1995) Intact protein electrospray ionization tandem mass spectrometry can be the sole technique used for confirming the structure of a variant hemoglobin. J. Mass Spectrom. Rapid Commun. Mass Spectrom. S111–S115

145 Eng J. K., McCormack A. L. and Yates J. R. (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in protein database. J. Am. Soc. Mass Spectrom. **5:** 976–989

146 Patterson S. D. and Aebersold R. (1995) Mass spectrometric approaches for the identification of gel separated proteins. Electrophoresis **16:** 1791–1814

147 Wilm M., Shevchenko A., Houthaeve T., Breit S., Schweigerer L., Fotsis T. et al. (1996) Femtomole sequencing of proteins from polyacrylamide gels by nano electrospray mass spectrometry. Nature **379:** 466–469

148 Mørtz E., O'Connor P. B., Roepstorff P., Kelleher N. L., Wood, T. D. and McLafferty F. W. (1996) Sequence tag identification of intact proteins by matching tandem mass spectral data against sequence data bases. Proc. Natl. Acad. Sci. USA **93:** 8264–8267

149 Shevchenko A., Jensen O. N., Podtelejnikov A. V., Sagliocco F., Wilm M., Vorm O. et al. (1996) Linking genome and pro-

teome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. Proc. Natl. Acad. Sci. USA **93:** 14440–14445
150 Yates J. R. (1998) Mass spectrometry and the age of the proteome. J. Mass Spectrom. **33:** 1–19
151 Zhang Z. and Smith D. L. (1993) Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. Protein Sci. **2:** 522–531
152 Smith D. L., Deng Y. and Zhang Z. (1997) Probing the noncovalent structure of proteins by amide hydrogen exchange and mass spectrometry. J. Mass Spectrom. **32:** 135–146
153 Mandell J. G., Falick A. M. and Komvies E. A. (1998) Measurement of amide hydrogen exchange by MALDI-TOF mass spectrometry. Anal. Chem. **70:** 3987–3995
154 Deng Y., Zhang Z. and Smith D. L. (1999) Comparison of continuous and pulsed labeling amide hydrogen exchange/mass spectrometry for studies of protein dynamics. J. Am. Soc. Mass Spectrom. **10:** 675–684
155 Bailey J. L. (1967) Techniques in Protein Chemistry, 2nd edn, Elsevier, Amsterdam
156 Shevchenko A., Loboda A., Shevchenko A., Ens W. and Standing K. G. (2000) MALDI quadrupole time-of-flight mass spectrometry: a powerful tool for proteomic research. Anal. Chem. **72:** 2132–2141
157 Griffin P. R., MacCoss M. J., Eng J. K., Blevins R. A., Aaronson J. S. and Yates J. R. (1995) Direct database searching with MALDI-PSD spectra of peptides. Rapid Commun. Mass Spectrom. **9:** 1546–1551
158 Vestal M., Juhasz P., Hines W. and Martin S. (2000) A new delayed extraction MALDI-TOF MS-MS for characterization of protein digests. In: Mass Spectrometry in Biology and Medicine, pp. 1–16, Carr S. A., Burlingame A. L. and Baldwin M. A. (eds), Humana Press, Totowa, N. J.
159 Hoyes J. B., Bateman R. H. and Wildgoose J. L. (2000) A high resolution orthogonal TOF with selectable drift length. Proc. 48th ASMS Conf. on Mass Spectrometry and Allied Topics, Long Beach, Calif.