REVIEW ARTICLE

Electrospray and tandem mass spectrometry in biochemistry

William J. GRIFFITHS[1], Andreas P. JONSSON, Suya LIU, Dilip K. RAI and Yuqin WANG

Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm SE-17177, Sweden

Over the last 20 years, biological MS has changed out of all recognition. This is primarily due to the development in the 1980s of 'soft ionization' methods that permit the ionization and vaporization of large, polar, and thermally labile biomolecules. These developments in ionization mode have driven the design and manufacture of smaller and cheaper mass analysers, making the mass spectrometer a routine instrument in the biochemistry laboratory today. In the present review the revolutionary 'soft ionization' methods will be discussed with particular reference to electrospray. The mass analysis of ions will be described, and the concept of tandem MS introduced. Where appropriate, examples of the application of MS in biochemistry will be provided. Although the present review will concentrate on the MS of peptides/proteins and lipids, all classes of biomolecules can be analysed, and much excellent work has been done in the fields of carbohydrate and nucleic acid biochemistry.

Key words: lipids, MS, peptides, proteins.

INTRODUCTION

In MS the mass-to-charge ratios (m/z) of gas-phase ions are determined. The technique has its origins in the studies performed by J. J. Thomson [1] and his student F. W. Aston [2] around the turn of the last century. Today MS is the most sensitive method for the structural characterization of biomolecules. Steroids can be detected and characterized at the picogram level [3,4], proteins separated by PAGE identified at the picomole level and below [5,6], and, in the best case, neuropeptides can be detected at the zeptomole (10^{-21} mol) level [7]. Many of the improvements in biochemical MS are a direct consequence of the introduction of the 'soft ionization' methods of matrix-assisted laser-desorption–ionization (MALDI) [8,9] and electrospray (ES) [10–13]. These ionization methods have allowed the direct analysis of polar and thermally labile biomolecules without prior derivatization. Furthermore, the mass of biomolecules amenable to these methods is now within the megadalton range [14,15]. Developments in ionization modes have driven MS manufacturers to build instruments with increased mass range, higher routine resolution and mass accuracy, and of smaller dimension and lower cost. The tandem MS (MS/MS) technique [16], in which a precursor ion is mass-selected and induced to fragment to give structurally significant product ions, was until recently the preserve of the specialist mass spectrometrist [17]. Today, instrument development is such that MS/MS spectra can be recorded under the control of automated software, and, at least in the area of peptide chemistry, unknown precursor ions identified by computer algorithms.

Over the last 10 years, studies in many MS laboratories have reflected these developments. Both MALDI and ES have been widely used as methods of sample ionization/vaporization, and MS/MS has been thoroughly exploited in the characterization of unknowns. The importance of high-resolution (HR) and accurate mass measurement have been in evidence in many studies. Although the present review will focus on MS in the areas of lipid and protein biochemistry, the techniques employed are of similar value in other areas of biochemistry [18,19]. Today, as the characterization of protein and peptide primary structure using MS becomes more and more routine [20,21], interest has shifted to the identification of post-translational modifications [22,23] and also into the area of non-covalent interactions [14,15,24,25]. It seems likely that many of the important new developments in MS will be in these areas.

SOFT-IONIZATION METHODS

In MS we measure the m/z ratio of gas-phase ions. Thus, to analyse a sample by MS, it must first be ionized and vaporized. Until the early 1980s, biomolecules were usually ionized by electron-impact (EI) or chemical-ionization (CI) methods. A requirement for these two ionization methods is a vaporized sample. This is of no great concern in the analysis of small organic molecules or those amenable to gas chromatography. However, polar thermally labile samples cannot be analysed by EI- or CI-MS without prior derivatization.

Fast atom bombardment

In 1981, Barber and co-workers introduced the fast-atom-bombardment (FAB) method of ionization [26]. This, for the first time, allowed the routine mass-spectrometric analysis of polar thermally labile molecules of masses of up to a few thousand daltons [27]. FAB is most suitable for the analysis of samples

Abbreviations used: MALDI, matrix-assisted laser-desorption–ionization; ES, electrospray; MS/MS, tandem MS; HR, high resolution; EI, electron impact; CI, chemical ionization; FAB, fast atom bombardment; TOF, time-of-flight; PEG, poly(ethylene glycol); FWHM, full width at half-maximum; PF, prophenin; FMN-bp, flavin mononucleotide-binding protein; ADH, alcohol dehydrogenase; CID, collision-induced dissociation; BIRD, black-body infrared radiative dissociation; ECD, electron-capture-induced dissociation; OATOF, orthogonal acceleration time-of-flight; DHA, docosahexaenoic acid; CRFs, charge-remote fragmentations; FTICR, Fourier-transform ion cyclotron resonance.

[1] To whom correspondence should be addressed (e-mail william.griffiths@mbb.ki.se).

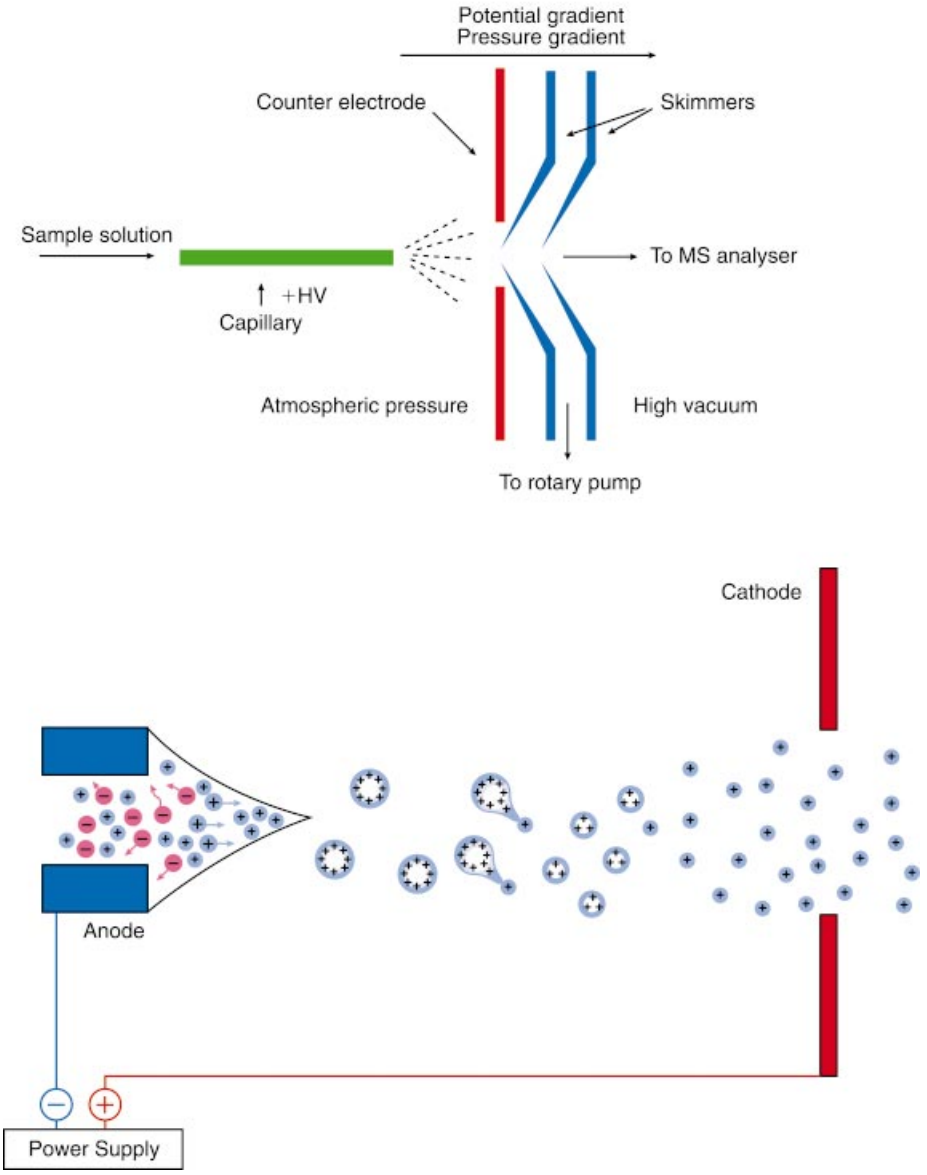


Figure 1 Features of the ES interface (upper panel) and schematic representation of the ES process (lower panel)

A high positive potential is applied to the capillary (anode), causing positive ions in solution to drift towards the meniscus. Destabilization of the meniscus occurs, leading to the formation of a cone [47] and a fine jet emitting droplets with excess positive charge. Gas-phase ions are formed from charged droplets in a series of solvent evaporation–Coulomb fission cycles. With the continual emission of positively charged droplets from the capillary (green), to maintain charge balance, oxidation occurs within the capillary. If the capillary is metal, oxidation of the metal may occur at the liquid/metal interface:

$M(s) \rightarrow M^{2+}(aq.) + 2e^{-}$ (in metal)

or, alternatively negative ions may be removed from solution by electrochemical oxidation:

$4OH^{-}(aq.) \rightarrow O_2(g) + 2H_2O + 4e^{-}$ (in metal)

where 'aq.' is aqueous phase and 'g' is gas phase. The upper panel is modified from [39] and the lower panel is modified from [40]; both with permission

which exist as preformed ions in solution, i.e. protonated or sodiated molecules (or deprotonated molecules). In simplistic terms, a solution (1 μl) of sample, containing perhaps $[M+H]^+$ ions and complementary negatively charged counter-ions, is dissolved in the FAB matrix (2 μl), which is very often glycerol, and introduced into the mass spectrometer vacuum system. The matrix-sample solution is then bombarded by fast atoms (8 keV, Xe) or ions (20 keV, Cs^+) and energy is transferred to the matrix-sample solution with the result that sample $[M+H]^+$ ions are vaporized along with protonated matrix clusters. These ions are then mass-analysed. While Barber et al. analysed protonated molecules as large as insulin [27], the method works best for smaller species of mass below about 1000 Da. A further limitation of the FAB method is the formation of matrix-associated chemical noise, which in the low-mass range may mask the presence of sample ions. Continuous-flow FAB [28] has been developed to minimize this problem and also make FAB compatible with on-line liquid chromatography.

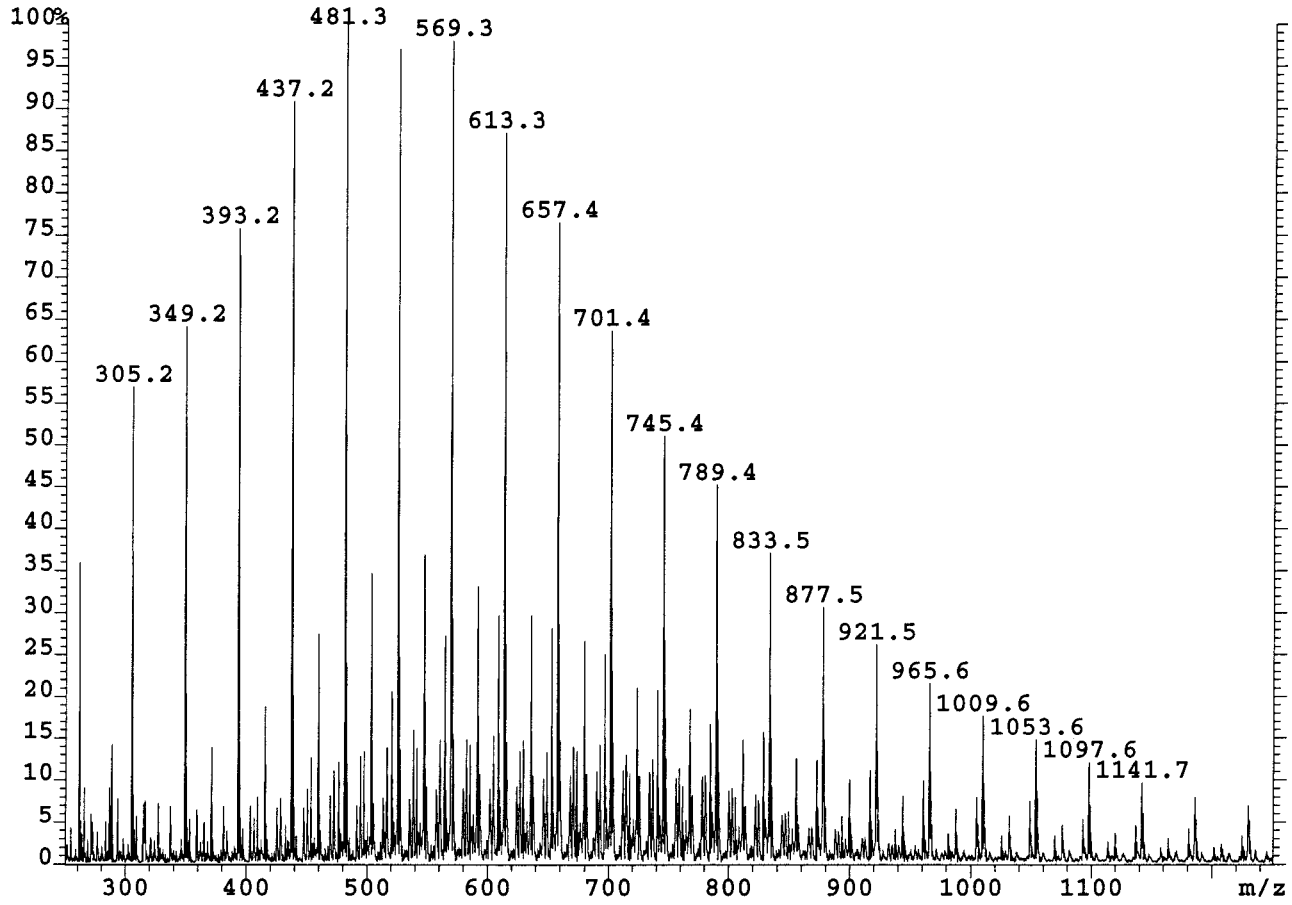


Figure 2 Positive-ion ES mass spectrum of a PEG mixture

The spectrum was recorded on an AutoSpec double-focusing magnetic-sector instrument. Peaks corresponding to $[H(C_2H_4O)_nOH + Na]^+$ are annotated.

MALDI

The MALDI technique attracted the attention of the MS community during the 11th International MS Conference in Bordeaux (1988), when Karas and Hillenkamp described MALDI of proteins with molecular masses exceeding 10 kDa [8]. Since this time MALDI has become an integral part of the mass spectrometrist's armoury, particularly in the field of protein chemistry, where it is the most efficient method of ionizing peptides. MALDI is most often coupled to time-of-flight (TOF) mass analysers [29]. As both MALDI and TOF mass analysis are pulsed events, this is an ideal arrangement. MALDI is quite similar to FAB; however, instead of FAB of a sample dissolved in a glycerol matrix, laser light is directed at the co-crystallization product of sample and a light-absorbing matrix. Light energy is absorbed by the crystals and dissipated, with the result that protonated (or deprotonated) sample molecules and matrix become vaporized. Very large proteins have been ionized by MALDI [30] and mass-analysed by TOF, which theoretically has an unlimited mass range.

In biochemistry, MALDI-TOF is perhaps most used in peptide mass mapping, where proteins, often separated by PAGE, are digested by a suitable enzyme and the resulting peptides mass-analysed [29]. The peptide mass map can be searched against a protein or DNA database and the parent protein identified [31–36]. Like FAB, MALDI is limited in the low-mass range by matrix-associated chemical noise; however, this is not a problem for the analysis of many proteolytic peptides. Both FAB and MALDI, being 'soft-ionization' methods, tend to give abundant $[M+H]^+$ (or $[M-H]^-$) ions, and although some fragmentation does occur [37], the fragment ion intensity is generally low.

ES process

While MALDI is the most efficient method for ionizing peptides, ES is the optimum method of ionization/vaporization for the widest range of polar biomolecules. The ES process for forming gas-phase ions originates from the work of Dole et al. [38], but it was Fenn's group at Yale University that coupled ES with MS [10,11]. Electrospray is conceptually a rather simple method for the ionization/vaporization of polar molecules (Figure 1). Initially the sample of interest is dissolved in a solvent where, to a certain extent, it will exist in an ionized form, e.g. $[M+H]^+$ ($[M-H^-]$). In conventional ES the solution is then pumped through a thin capillary (internal diameter approx. 0.1 mm) which is raised to a high potential (4 kV). Small charged droplets are sprayed from the ES capillary into a bath gas at atmospheric pressure and travel down a pressure and potential gradient towards an orifice in the mass-spectrometer high-vacuum system. As the droplets traverse this path they become desolvated and reduced in size to such an extent that surface-coulombic forces

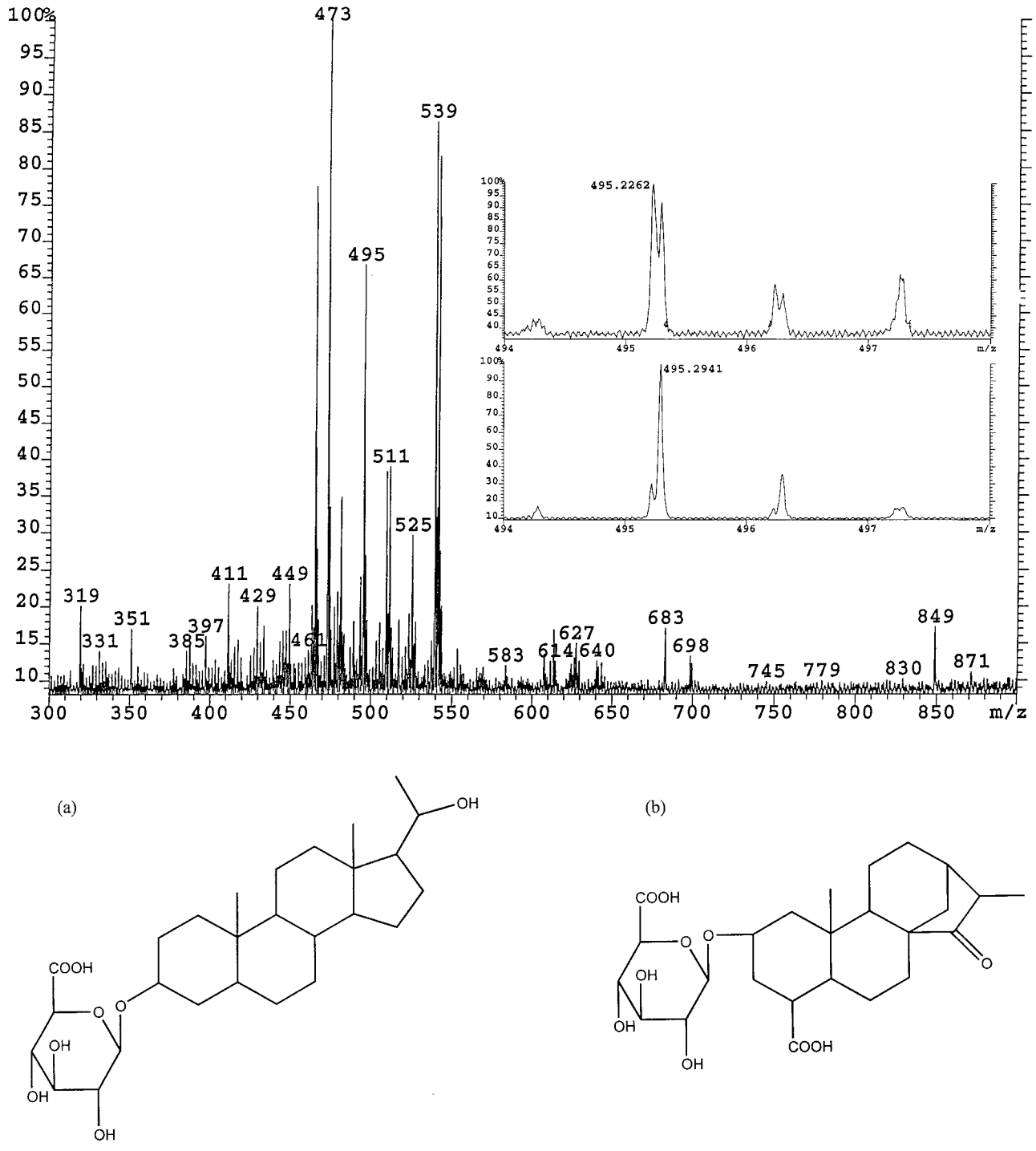


Figure 3 Negative-ion ES mass spectrum of the weakly acidic lipid fraction extracted from the urine of a female coffee drinker suffering from intrahepatic cholestasis of pregnancy

The inset in the upper panel shows expanded views of the region *m/z* 494–498 from two different samples. The spectra were recorded on an AutoSpec double-focusing mass spectrometer. In the lower panel the structures of the glucuronides of (**a**) pregnandiol, and (**b**) 2-β-hydroxy-15-oxo-atractylan-4-carboxylic acid are also shown.

overcome surface-tension forces and the droplets break up into smaller droplets [39,40]. This process continues until the point is reached that either an ion desorbs from a droplet [41,42] or solvent is completely removed [38]. The exact mechanism of ion formation, whether it be by ion evaporation (ion-evaporation model) or by complete solvent removal (charge-residue model), from the charged droplet is under debate, and in all probability different mechanisms apply in different situations [43–46]. Whatever the mechanism, the result is a beam of ions, which are sampled by the mass spectrometer. For a more detailed description of the ES process the interested reader is directed to the excellent volume *Electrospray Ionization Mass Spectrometry:*

Table 1 Tryptic peptide *m/z* values submitted to a database search using the MS-Fit program [68]

$[M+H]^+$ (m/z)*				
Submitted	Matched	δ (p.p.m.)	Sequence†	PTM‡
262.1555§	–	–	–	–
263.1395¶	263.1396	0.2567	(R)FP(−)	–
1077.5746‖	1077.5845	9.1865	(R)FPPPNVPGPR(F)	–
1125.5838‖¶**	1125.5845	0.6568	(R)FPPPNFPGPR(F)	–
1148.6216¶	1148.6217	0.0462	(−)AFPPPNVPGPR(F)	–
1259.6518‖**	1259.6537	1.4937	(R)QAFPPPNVPGPR(F)	PyroGlu
3983.0703‖¶**	3983.0782	1.9951	(R)FPPPNFPGPPFPPPIFPGPWFPPPPPFRPPPFGPPR(F)	–

* Monoisotopic mass.
† (X) residues preceding and following peptide.
‡ Post-translational modification.
§ Unatributed mass.
‖ Matched to PR-2 protein.
¶ Matched to prophenin-1.
** Matched to homologue of proline/arginine-rich antibacterial protein.

Table 2 MS-Fit results* using the *m/z* values in Table 1

Rank‡§	Masses matched	NCBI† number	Species	Protein molecular mass (Da)	Protein name
1	4/7	1730501	Wild boar (*Sus scrofa*)	25855.1	PR-2 protein
2	4/7	945173	Domestic pig (*Sus domestica*)	8681.2	Prophenin-1
3	3/7	1730500	Wild boar (*Sus scrofa*)	23955.9	Homologue of proline/arginine-rich antibacterial protein

* 10 p.p.m. mass tolerance, monoisotopic mass, all species, 1–100 kDa, trypsin digest, one missed cleavage, Cys unmodified, three masses matched, post-translational modification considered N-terminal Gln to pyroGlu, oxidized Met.
† National Center for Biotechnology Information.

Fundamentals, Instrumentation and Applications edited by Cole [48], and to a series of papers [43,49–52] in a special feature on electrospray ionization in the *Journal of Mass Spectrometry*.

ES is a concentration- rather than a mass-dependent process, and improved sensitivity is obtained for high-concentration low-volume samples [13]. This has led to the development of low-flow-rate ES (10–100 nl/min). Initial reports of the application of low-flow-rate ES came from Smith's laboratory [53], soon followed by Caprioli and co-workers' [7,54] description of micro-ES, and Wilm and Mann's illustration of the practical use of nano-ES [55]. Both nano- and micro-ES are now widely used in biological MS [3–7,14,15].

ES SPECTRA

Singly charged ions

The earliest ES mass spectra recorded were of mixtures of poly(ethylene glycol) (PEG) polymers [10]. An ES spectrum of a PEG mixture is shown in Figure 2. A major and a minor series of peaks, corresponding to $[H(C_2H_4O)_nOH+Na]^+$ and $[H(C_2H_4O)_nOH+H]^+$ respectively are observed. Formation of sodium adducts is a common process in positive-ion ES, particularly for species which do not possess a strongly basic group. Gas-phase $[H(C_2H_4O)_nOH+Na]^+$ ions are efficiently formed by the ES process as a consequence of their high surface activity. The use of PEG in many biochemical experiments can provide a major obstacle to the recording of ES spectra, in that PEG ions can completely dominate spectra, obscuring the ion of interest. The fact that PEG polymers are heterogeneous with respect to carbon number, and give two ion series in ES, makes them ideal calibrants of the instrument m/z scale. The exact mass for any given member of an ion series can be readily calculated, e.g. $[H(C_2H_4O)_{10}OH+Na]^+ = 481.2625$ Da, and successive members differ in mass by 44.0262 Da (C_2H_4O).

Accurate mass measurements using ES

In many applications the determination of the nominal mass of an ion is sufficient to satisfy the investigator; however, for the characterization of unknowns and for the exact confirmation of elemental formula, accurate mass measurements are required [56,57]. In the 1950s, Beynon reported the use of accurate mass measurements to determine the elemental composition of ions [58,59]. The fact that the masses of the isotopes of the elements have non-integer values (with the exception of ^{12}C), results in polyatomic ions having non-integer masses. In the mass range below about 500 Da, unique elemental compositions are separated by at least 1 p.p.m., and in many cases determination of exact mass to an accuracy of 5 p.p.m., i.e. $\pm 0.0005\%$, allows elemental composition to be determined when valence rules and complementary chemical information is considered [60]. HR has often been considered a requirement for accurate mass measurements, but this criterion is non-essential, at least for the analysis of pure samples [58,60,61]. However, to determine the elemental composition of an unknown, it is desirable to record the mass spectrum at HR to minimize the possibility of overlapping peaks [4,62–66].

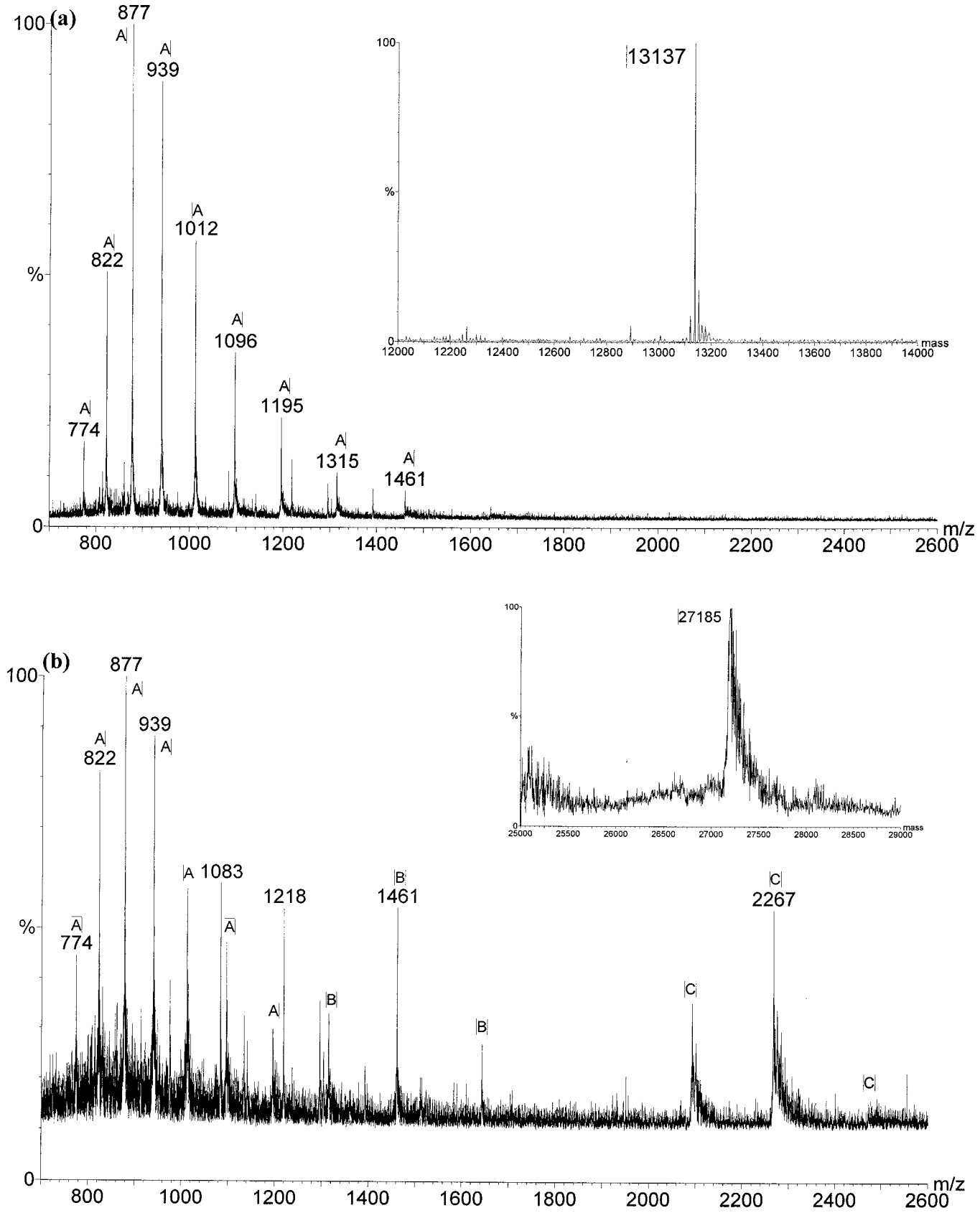


Figure 4 Positive-ion ES spectra of FMN-bp recorded under (a) denaturing conditions and (b) non-denaturing conditions

Peaks labelled 'A' and 'B' correspond to multiply protonated molecules of the FMN-bp monomer in different conformations, and those labelled 'C' correspond to the homodimeric complex of FMN and FMN-bp. The insets show the deconvoluted mass spectra obtained by application of the maximum entropy algorithm [72]. Spectra were recorded on an AutoSpec double-focusing mass spectrometer.

HR in combination with accurate mass measurements

Traditionally HR mass spectrometers [58,67] have been capable of attaining resolutions of at least 10000 (10% valley definition). However, the term HR is now loosely used, particularly in connection with TOF mass analysers, where the full-width-at-half-maximum (FWHM) height definition is employed. Resolution on the FWHM scale has a value 2–10 times higher than on the 10% valley scale. Today, manufacturers often describe their instruments as being capable of HR if they can achieve unit resolution on the isotopic peaks of oxidized bovine insulin, which has a mass of only 3.5 kDa ($[M+H]^+$ at 3493.643 Th).

HR is particularly important when analysing complex mixtures extracted from biological matrixes. Figure 3 is the negative-ion ES spectrum of the weakly acidic lipid fraction extracted from the urine of a female patient suffering from intrahepatic cholestasis of pregnancy [66] who also happened to be a coffee drinker. Intense ions are observed at m/z 473, 539 and 541, which correspond to deprotonated molecules of steroid glucuronides. An intense ion is also observed at m/z 495 and less intense ions at m/z 509 and 511, corresponding in mass to the deprotonated molecules of the glucuronides of pregnandiol, pregnandiolone and pregnantriol respectively. HR analysis of the peak at m/z 495 showed it to be a doublet (inset to Figure 3). Accurate mass measurement indicates that the ion at m/z 495 is a composite of deprotonated pregnandiol glucuronide (calc. 495.2958 Th), which has an elemental composition of $C_{27}H_{43}O_8$ and a second component at m/z 495.2262 Th, giving an elemental composition of $C_{25}H_{35}O_{10}$ (calc. 495.2230), which upon further analysis was characterized to be the deprotonated glucuronide of 2-β-hydroxy-15-oxo-atractylan-4-carboxylic acid (Figure 3). 2-β-Hydroxy-15-oxo-atractylan-4-carboxylic acid is an isomer of atractyligenine, which is the aglycone of atractylosides isolated from *Cofea arabica* (coffee) beans. The patient was a coffee drinker, and the glucuronide of 2-β-hydroxy-15-oxo-atractylan-4-carboxylic acid found in her urine is a metabolic product of a coffee atractyloside. This example provides a cautionary note to those involved in quantitative metabolic studies. In this instance the use of the ion at m/z 495 as a reporter for pregnandiol glucuronide levels would lead to incorrect results.

Accurate mass measurements in protein identification

With the ever-increasing amount of protein sequence information available in the databases, protein identification using database search routines [31–36] is becoming more and more common [68]. The general procedure to identify a protein is to subject it to proteolysis, usually using trypsin, and to record the mass spectrum of the resulting peptides. The masses of the proteolytic peptides are then searched against a protein or DNA database [29]. Programs have been written that are accessible over the World Wide Web and which allow the investigator to feed in experimentally determined peptide masses together with an estimate of their accuracy, the proteolytic enzyme used, the proteins expected mass range and origin, e.g.

http://pepsea.protana.com/

http://www.mann.embl-heidelberg.de/Services/PeptideSearch/PeptideSearchIntro.html

http://falcon.ludwig.ucl.ac.uk/mshome3.2.htm

http://www.proteometrics.com/

The search algorithm then theoretically digests all appropriate proteins in the database with the specified enzyme and matches the theoretical and experimental peptide masses. The best matches are then displayed. Clearly the more accurate the experimentally determined peptide masses the more likely the protein will be correctly identified. This is of particular importance when the peptide mass map is a composite for a number of proteins.

An example of protein identification using a database search routine is the characterization of prophenin (PF), an antibacterial protein isolated from porcine pulmonary tissue [69,70]. The protein was digested with trypsin and a mass spectrum recorded. Tables 1 and 2 list the measured m/z values of the tryptic peptides and the database search results using the MS-Fit program [68]. The peptide ranked 1 in Table 2 is a PF-2 precursor; that ranked 2 is PF-1, an antibacterial protein isolated from porcine leucocytes [71], and that ranked 3 is a PF-1 precursor. These results suggest that the antibacterial protein isolated from porcine pulmonary tissue is an isoform of PF.

Further studies [69,70] showed porcine pulmonary PF to exist predominantly in two isoforms, one of which is C-terminally amidated, i.e. PF-2, while the other is N-terminally modified with pyroglutamic acid, but not C-terminally amidated, i.e. PF-2-Pyr. A third minor isoform is also present which is N-terminally truncated, i.e. PF-1. None of these three isoforms correspond to the protein originally isolated from porcine leucocytes, which shows a Val-to-Phe isomorphism and is N-terminally truncated [71]. The database search did not provide a perfect match to the submitted masses, as the porcine pulmonary PF isolated had yet to be incorporated into a database.

Proteins and multiply charged ions

While small organic molecules tend to appear as monoprotonated (or monodeprotonated) molecules in ES spectra, larger molecules with multiple basic (or acidic) sites tend to give multiply protonated (or multiply deprotonated) molecules [12,13]. This phenomenon of multiple charge formation has the effect of reducing the m/z. Thus, although a mass analyser may have an m/z range of 1–2000 Th, a protein of, for example, 40 kDa, can be readily mass-analysed. At first sight a protein ES mass spectrum might appear rather complex (Figure 4a); however, peaks from a given protein differ only by the degree of protonation where nearest neighbours differ by a single charge. Each peak gives a measure of the protein mass, and thus, in each spectrum, multiple mass measurements are made, thereby improving the precision of the measurement. For two successive peaks at positions x_1 and x_2 on the m/z scale and corresponding to members of the same ion series then, if

$$x_1 = (M+n)/n$$

where M is molecular mass, and n is the number of charges which is equivalent to the number of added protons, and

$$x_2 = (M+n+1)/(n+1)$$

Solving these equations gives:

$$n = (x_2-1)/(x_1-x_2)$$

In reality the mass-spectrometer data system performs such calculations, and plots a molecular mass spectrum on a zero-charge-state x-axis. A refinement of this approach is to use a maximum entropy processing [72], where a probabilistic method is applied which functions by repeatedly processing trial-and-error spectra of different mass and charge and comparing them with the experimental data (Figure 4).

When ES mass spectra are recorded at HR, it is possible to determine monoisotopic mass values for small proteins. For

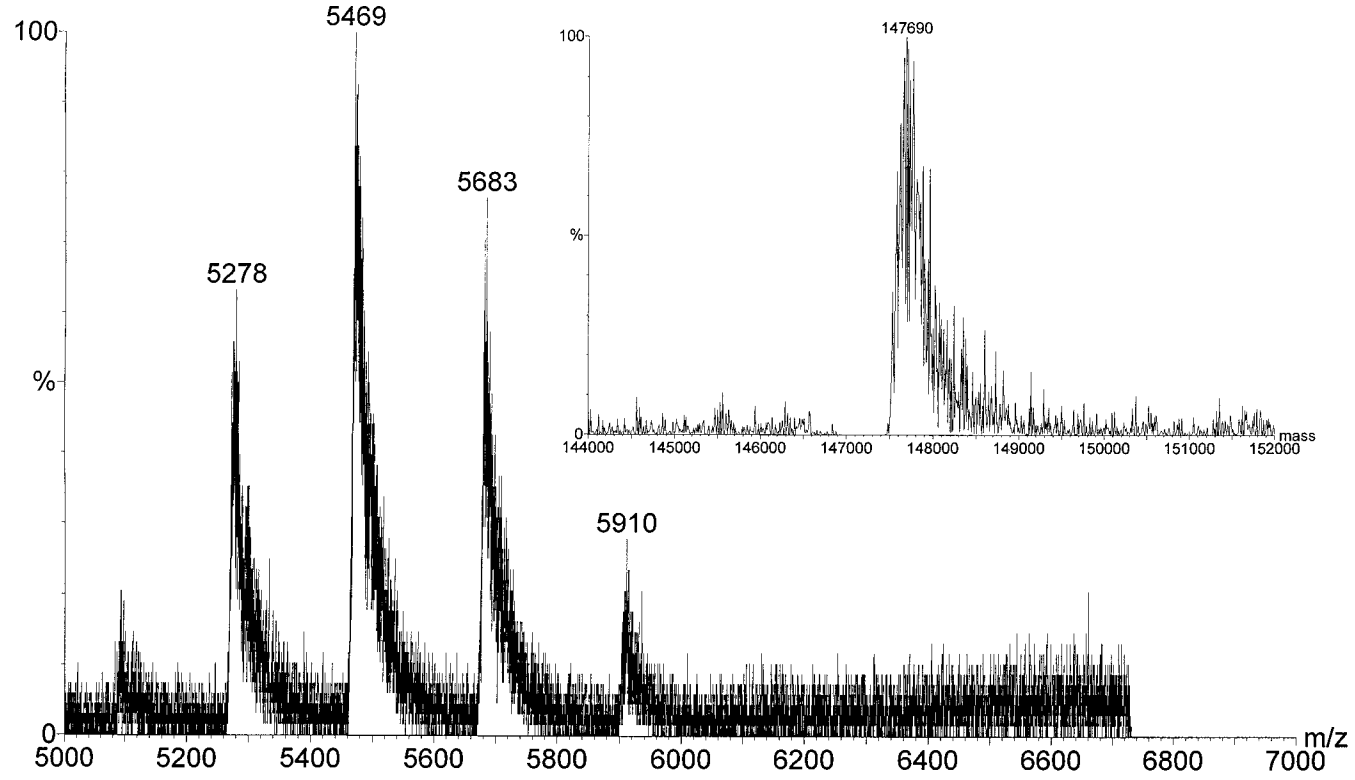


Figure 5 Positive-ion ES spectrum of the tetrameric complex of yeast ADH recorded on the Q-TOF mass spectrometer

The inset shows the deconvoluted mass spectrum obtained by application of the maximum entropy algorithm.

example, the monoisotopic mass of porcine pulmonary PF-2 was to be determined with an accuracy of better than 0.1 Da (calc. 8754.5676, exp. 8754.5447 Da). For proteins of larger mass, determination of monoisotopic mass becomes more difficult, and average mass is measured instead (the monoisotopic mass of an element refers specifically to the lightest stable isotope of that element). 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, to give the isotope weighted average mass of 12.011 for carbon. The average mass of a molecule is then the sum of the chemical average masses of the elements present [56]. Protein mass measurements are typically of average mass, but accurate mass measurements can be made even when the isotopic envelope remains unresolved, provided that the peak shape is Gaussian [73]. The accuracy of mass measurements made using average mass values are ultimately limited to 10 p.p.m., owing to variations in the natural abundance of ^{13}C, which causes the average mass of carbon to vary between 12.0107 and 12.0111.

ES in the study of non-covalent interactions

The major function of the ES interface is the disruption of non-covalent interactions between analyte and solvent molecules. However, non-covalently bound species can be observed in ES spectra provided that the interactions between analytes are sufficiently strong to survive disruption. ES-MS has been used to study interactions between (a) identical components which lead to multimer formation [74,75] and (b) structurally dissimilar species which are related by biological function [25,74,76–78]. The observation of multimeric species in ES spectra may be assumed to represent forms present in the condensed phase, although possibly only in the electrosprayed droplets rather than the bulk solution [39]. The formation of non-covalent multimers is often favoured by conditions which do not maximize the charge state of the sprayed ions and are thus are best observed on extended-mass-range instruments.

An interesting example of a protein which forms multimers and also binds to dissimilar species is the flavin-mononucleotide-binding protein (FMN-bp) from *Desulfovibrio vulgaris* Miyaki F. This protein is related structurally to a hypothetical single domain precursor of chymotrypsin and related proteases [79,80]. FMN-bp has been structurally characterized by both NMR [79] and X-ray crystalography [80], the X-ray data indicating the existence of a homodimer. It has been shown that FMN is bound non-covalently in a shallow groove at the surface of the protein, with the negatively charged phosphoester group interacting with positively charged lysine and histidine residues. Using ES-MS it was possible to confirm the formation of homodimers of FMN-bp, and to demonstrate the non-covalent binding of FMN to FMN-bp [81]. An initial experiment was performed in which FMN-bp was dissolved in a denaturing solvent of aq. 50 % methanol/1 % acetic acid. The ES spectrum (Figure 4a) shows a classical charge-state distribution envelope, which upon maximum entropy treatment gave the protein molecular mass to be 13137 Da, in agreement with the theoretically calculated average mass. There was no evidence of any protein–cofactor interactions being maintained in this solvent or of any homodimers, despite the concentration of protein being 4 pmol/μl. The ES spectrum of FMN-bp sprayed from a solution of water/acetic acid (99:1, v/v) (Figure 4b) shows a different distribution of peaks. In fact

there are three discrete charge-state distributions. The first maximizes at 877 Th ($z = 15$) as in the spectrum obtained from electrospraying the protein from the aq. 50% methanol/1% acetic acid solvent (Figure 4a). A second distribution consisting of just three charged states is centred around 1461 Th ($z = 9$). Both these distributions correspond to charged states of the monomeric FMN-bp. Maximum entropy treatment of the raw data gives a mass of 13137 Da. The third charge-state distribution is centred at 2267 Th ($z = 12$). This distribution consists of three peaks which correspond to the 13+, 12+ and 11+ charge states of a homodimeric complex between FMN and FMN-bp. Maximum entropy processing of this series gives a mass of 27185 Da, which is in exact agreement with the theoretical mass of the cofactor bound homodimer, consisting of two units of FMN and two units of FMN-bp. It is suggested that the three charge-state distributions, maximizing at 877, 1461 and 2267 Th respectively correspond to the protein in three different conformations. The charge-state distribution maximizing at 877 Th resembles the spectrum of FMN-bp sprayed from the denaturing aqueous acid/methanol solvent. This charge-state distribution can be regarded as characteristic of the denatured unbound monomeric protein. In contrast, the charge-state distribution centred at 2267 corresponds to the homodimer of the protein–cofactor complex, and it can be assumed that this distribution is characteristic of the cofactor-bound homodimer. The other charge-state distribution, centring at 1461 Th, corresponds to a free monomer, but one with a conformation different from that found in the denaturing solvent. It is significant that in the 2267-Th-centred charge-state distribution, no cofactor-free homodimers were observed.

Much larger protein complexes are also readily observed by ES-MS. Yeast alcohol dehydrogenase (ADH) is an example of a large multimeric protein complex where metal binding is also important. The non-covalently bound tetrameric complex of yeast ADH ($M_r \approx 147000$) has been observed by recording ES spectra under non-denaturing conditions [75]. Under these conditions, dimeric yeast ADH complexes are not observed, consistent with the solution-phase characteristics of the protein. Figure 5 shows an ES mass spectrum of yeast ADH recorded under non-denaturing conditions. The tetrameric species consists not only of four protein subunits, but also eight non-covalently bound Zn^{2+} ions. The measured mass of the tetramer is consistent with this proposal; however, under the mild ES conditions required to maintain the protein complex, complete solvent declustering does not occur.

MS/MS

In an MS/MS experiment a precursor ion is mass-selected by mass analyser 1 (MS_1) and focused into a collision region preceding a second mass analyser (MS_2) (see Figure 6). The mass analysers are arranged in series either in space (sector, triple quadrupole and hybrid instruments) or in time (trapping instruments). Inert gas is generally introduced into the collision region and collisions occur between the precursor ion and inert-gas atoms (molecules). In these collisions part of the precursor ion's translational energy can be converted into internal energy, and as a result of single or multiple collisions an unstable excited state is populated. Excited precursor ions decompose to product ions in a process termed 'collision-induced dissociation' (CID). Product ions are mass-analysed by MS_2 (Figure 6). Alternative methods of precursor-ion dissociation include surface-induced dissociation ('SID') [82], black-body infrared radiative dissociation (BIRD) [83] and electron-capture-induced dissociation (ECD) [84–86].

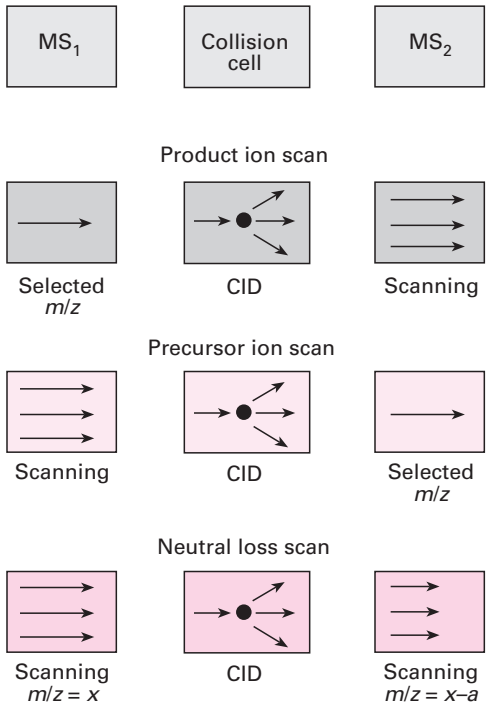


Figure 6 MS/MS experiments on spatially separated mass analysers

Adapted from [16] with permission ©John Wiley & Sons Inc., New York.

The application of MS/MS for the structural characterization of organic molecules was pioneered by Beynon and co-workers at Purdue University [87]. Early MS/MS experiments were performed on multiple sector and hybrid instruments, e.g. BE, EB, BEEB, EBQ (B = magnetic sector, E = electric sector and Q = quadrupole) [16,88–90]. However, such instruments are loosing popularity on account of their high complexity, cost and size. In the mid 1970s, Yost and Enke [91] built the first triple-quadrupole MS/MS instrument, and in the 1990s a third type of MS/MS instrument became popular – the quadrupole ion trap [92]. These instruments have since been commercialized and manufactured at considerably lower cost than sector instruments. Fourier-transform ion cyclotron resonance (FTICR) instruments [93,94] are also a form of trapping mass spectrometer. With these instruments it is possible to mass analyse ions while they are trapped. FTICR instruments show great promise in that they also provide exceptional HR and mass accuracy. The introduction of ECD [84] and BIRD [83], in addition to sustained off-resonance irradiation ('SORI') [94], have furthered their MS/MS capabilities, and no doubt these instruments will appear in mainstream biochemistry laboratories in the near future. However, high cost may yet prove to be a limiting factor.

While each type of MS/MS instrument has its own strengths and weaknesses, by combining analysers in a hybrid conformation, an attempt is made to accentuate positive features while minimizing the negative ones. The first commercial hybrid MS/MS instrument using orthogonal acceleration (OA) TOF technology was the Micromass AutoSpec-OATOF, which employs a double-focusing EBE arrangement as MS_1, and an orthogonally arranged linear TOF as MS_2 [95,96]. Instruments using a quadrupole as MS_1 and an orthogonally arranged reflectron TOF as MS_2 have subsequently been introduced, e.g. Micromass Q-TOF, Sciex Q-Star [97–99].

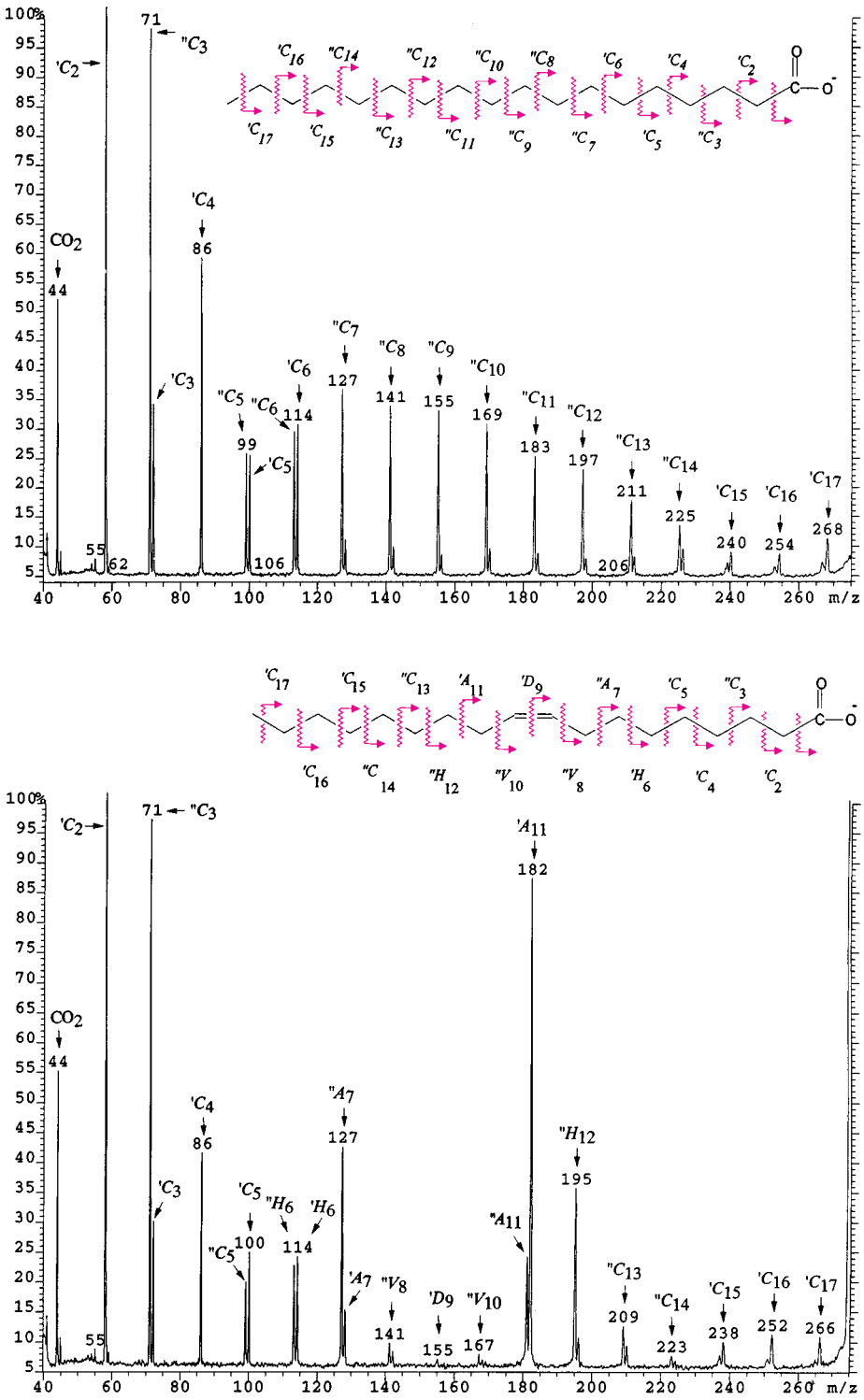
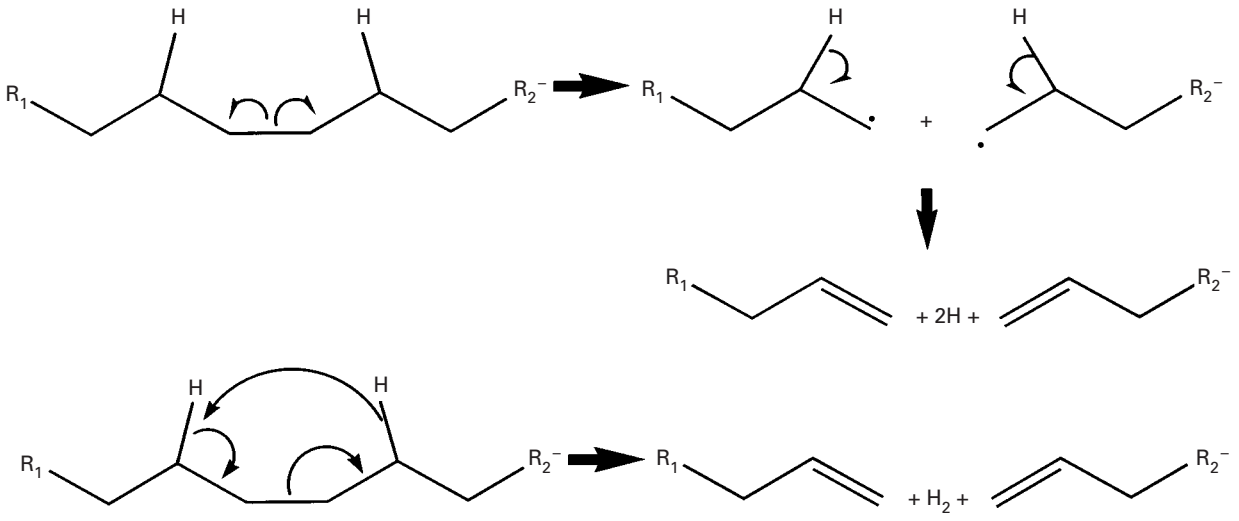


Figure 7 400 eV CID spectrum of the $[M-H]^-$ ion of stearic acid (283.3 Th) (upper panel) and oleic acid (281.2 Th) (lower panel)

The spectra were recorded on an AutoSpec-OATOF tandem mass spectrometer. Xe was the collision gas. See [106] for a description of the nomenclature employed. The lower panel is taken from [106] with permission © John Wiley and Sons Ltd., Chichester.

High- and low-energy CID

With the proliferation of multiple-sector and triple-quadrupole instruments during the 1980s, it was noted that the appearance of a CID spectrum was dependent on the projectile's collision energy. Sector instruments are traditionally operated at high accelerating potentials, giving projectile ions keV translational energies. Triple-quadrupole instruments, by contrast, are operated with accelerating potentials of a few tens of electron-volts, and thus provide projectile translational energies in the



Scheme 1 CRF mechanisms to account for the formation of distonic radical anions and terminally unsaturated anions

tens-to-hundreds-of-electronvolts range. High energy CID spectra (keV) tend to be dominated by fragment ions formed by charge-remote mechanisms [100,101], while low-energy CID spectra (< 100 eV) often show intense fragment ions formed by abundant neutral losses. The appearance of peptide CID spectra, especially those of $[M+H]^+$ ions, are highly dependent on the collision energy [21]. The CID spectra of peptides recorded at low collision energy tend to show abundant fragment ions formed by cleavage of the peptide bonds, while high-energy CID spectra often contain fragment ions formed by other backbone and also side-chain cleavages [102,103]. In fact CID spectra tend to be classified as of the high energy-type if fragment ions formed by peptide side-chain cleavage are observed. Bordoli and Bateman [104] have shown that, to achieve side-chain fragmentation, a centre of mass collision energy of 20–25 eV is required. Centre of mass collision energy (E_{cm}) being the maximum amount of the projectile's translational energy (E_{lab}), which can be converted into internal energy for fragmentation. Centre of mass energies above 25 eV are defined as high-collision energies and those below 20 eV as low-collision energies. Centre of mass collision energies are calculate from the equation:

$$E_{cm} = E_{lab} \cdot M_{cg}/(M_p + M_{cg})$$

Where M_p corresponds to the mass of the projectile ion, and M_{cg} the mass of the collision gas.

CID of fatty acid $[M-H]^-$ ions

Non-esterified and conjugated fatty acids tend to have molecular masses below 700 Da, e.g. $[M-H]^-$ stearic acid, 283.3; oleic acid, 281.2; docosahexaenoic acid (DHA), 327.2; leukotriene C_4, 624.3. The upper panel of Figure 7 shows the high-energy CID spectrum of the $[M-H]^-$ ion of stearic acid. This spectrum is typical of straight-chain saturated fatty acids. Fragmentation is believed to occur via mechanisms that do not involve the charged group directly and are thus referred to as charge-remote fragmentations (CRFs) [105]. As can be seen from the spectrum, a mixture of odd ($''C_n$) and even ($'C_n$) mass fragment ions are formed, where successive pairs of ions are separated by 14 Th (CH_2). (See [106] for a description of the nomenclature employed in Figure 7.) Ions of even mass ($'C_n$) are believed to be distonic radical anions, while those of odd mass ($''C_n$) are thought to be terminally unsaturated anions. Mechanisms have been proposed to account for these fragment ions. Even mass fragment ions ($'C_n$) are believed to be formed in a simple homolytic fragmentation reaction (Scheme 1), while odd mass fragment ions ($''C_n$) may be formed via two successive homolytic fragmentations [107] or via a 1,4-H_2 elimination reaction [105]. In contrast with spectra recorded at high collision energies, spectra of $[M-H]^-$ ions of straight-chain fatty acids recorded at low collision energies, on, for example, a triple-quadrupole instrument, are dominated by ions formed by simple neutral losses from the precursor ion and are far less structurally informative [108]. The informative nature of high-energy CID spectra is further illustrated in the lower panel of Figure 7, which shows the high-energy CID spectrum of the $[M-H]^-$ ion of oleic acid. The double bond is immediately located by the disruption in the regular spacing of fragment ions. Similarly, the presence of a side-chain substituent would be readily identified. In this way, and with additional information from accurate mass measurements, DHA was identified as an endogenous ligand for the retinoid X receptor [109].

High-energy CID of conjugated bile acids and steroids

The MS/MS of steroids and bile acids has been an area of considerable study [3,4,110–115]. When analysed by negative-ion FAB or ES-MS, bile acids, steroid sulphates and steroid glucuronides give abundant $[M-H]^-$ ions and few fragment ions. To gain structural information, CID spectra are recorded. High-energy CID spectra of $[M-H]^-$ ions are most informative. Bile acid, steroid sulphate and steroid glucuronide $[M-H]^-$ ions fragment in high-energy CID reactions via series of CRFs. Figure 8 shows the 400 eV CID spectrum of 3β-hydroxy-5α-pregnane-20-one sulphate. Any modification of the steroid skeleton or side chain would result in a change in the pattern of fragment ions.

Siuzdak and co-workers have analysed steroid sulphates on a triple-quadrupole instrument under low-collision-energy conditions [3]. They found that the major fragment ions were at m/z 97, corresponding to HSO_4^- ions. Although their spectra provided little structural information, the high abundance of the ion at m/z 97 encouraged them to perform precursor ion scans for this ion.

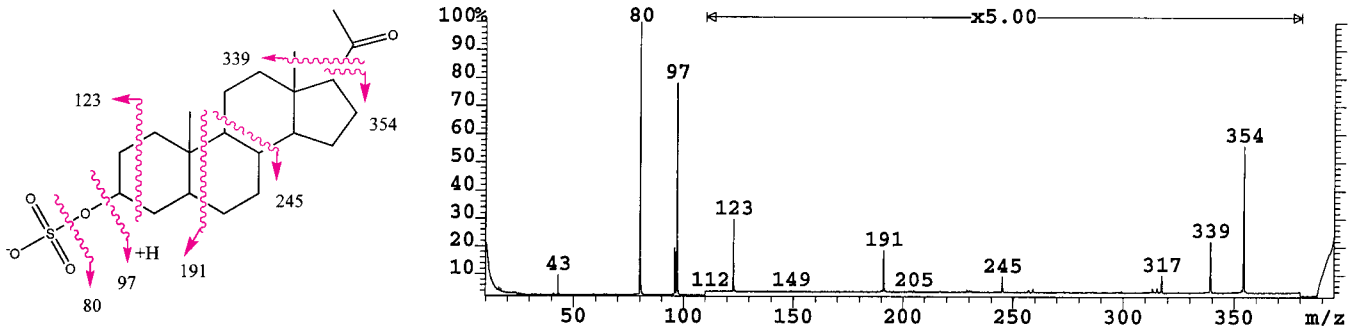


Figure 8 400 eV CID spectrum of the $[M-H]^-$ ion from 3β-hydroxy-5α-pregnane-20-one sulphate (397.2 Th)

The spectrum was recorded on a AutoSpec-OATOF instrument with Xe as the collision gas.

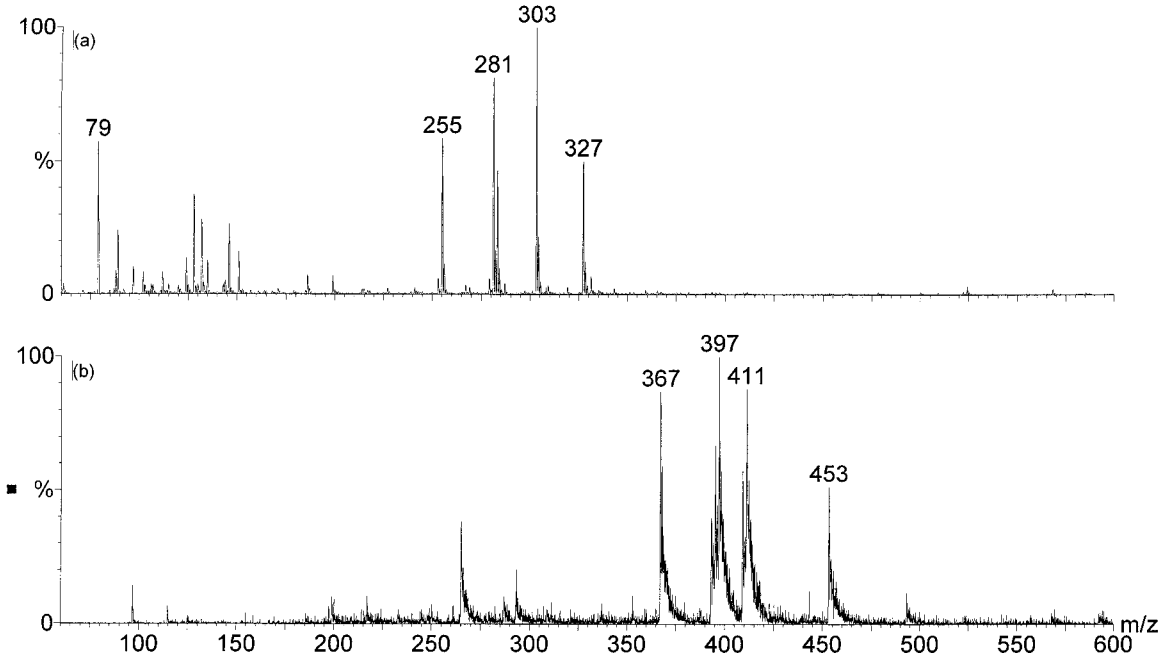


Figure 9 Negative ion ES mass spectrum and precursor ion spectrum

(**a**) Mass spectrum of the lipid fraction extracted from rat brain containing added steroid sulphates of $[M-H]^-$ *m/z* 367, 393, 395, 397, 409, 411 and 453. The peaks at *m/z* 255, 281, 283, 303 and 327 correspond to fatty acid $[M-H]^-$ ions. (**b**) Precursor ion scan for *m/z* 97. Spectra were recorded on a Quattro Ultima mass spectrometer.

Precursor-ion and constant-neutral-loss scans

In a precursor ion scan, MS_2 is set to transmit a specific product ion (e.g. *m/z* 97), and MS_1 is scanned (Figure 6). The resultant spectrum shows the precursors of the defined product ion. Siuzdak and co-workers [3] applied the precursor-ion scan to determine steroid sulphates in biological fluids. They were able to detect steroid sulphated at a level of 80 fg/μl from a biological matrix. The triple-quadrupole instrument is most sensitive for the recording of precursor-ion scans (Figure 9), although similar scans can be recorded on sector instruments.

The constant-neutral-loss scan is also most effectively carried out on triple-quadrupole instruments. In a constant-neutral-loss scan, MS_1 and MS_2 are scanned in parallel with an offset between them equivalent to the neutral loss to be observed (Figure 6). Constant-neutral-loss scans are effective when fragmentation proceeds with the facile loss of a neutral. For example, Libert et al. [111] showed how sulphated bile acids could be selectively monitored in a biological mixture by scanning MS_1 and MS_2 in parallel but with an offset of 80. Only ions which fragment with this neutral loss are detected. The effect is the removal of background chemical noise.

CID spectra of peptides and proteins

It is in the area of protein chemistry that MS/MS has become most widely popularized [5–7,97,102,103,116–134]. Multiply charged ions are the dominant products of the ES ionization of peptides and proteins, and fragment predominantly through peptide bond cleavage to give series of Y- and B-type ions (Figure 10a). The mechanism of peptide-bond cleavage is now fairly well understood [117–126] (Figure 11, lower panel). Considering, for example, a doubly protonated tryptic peptide, one proton will be localized on the C-terminal Arg or Lys side chain

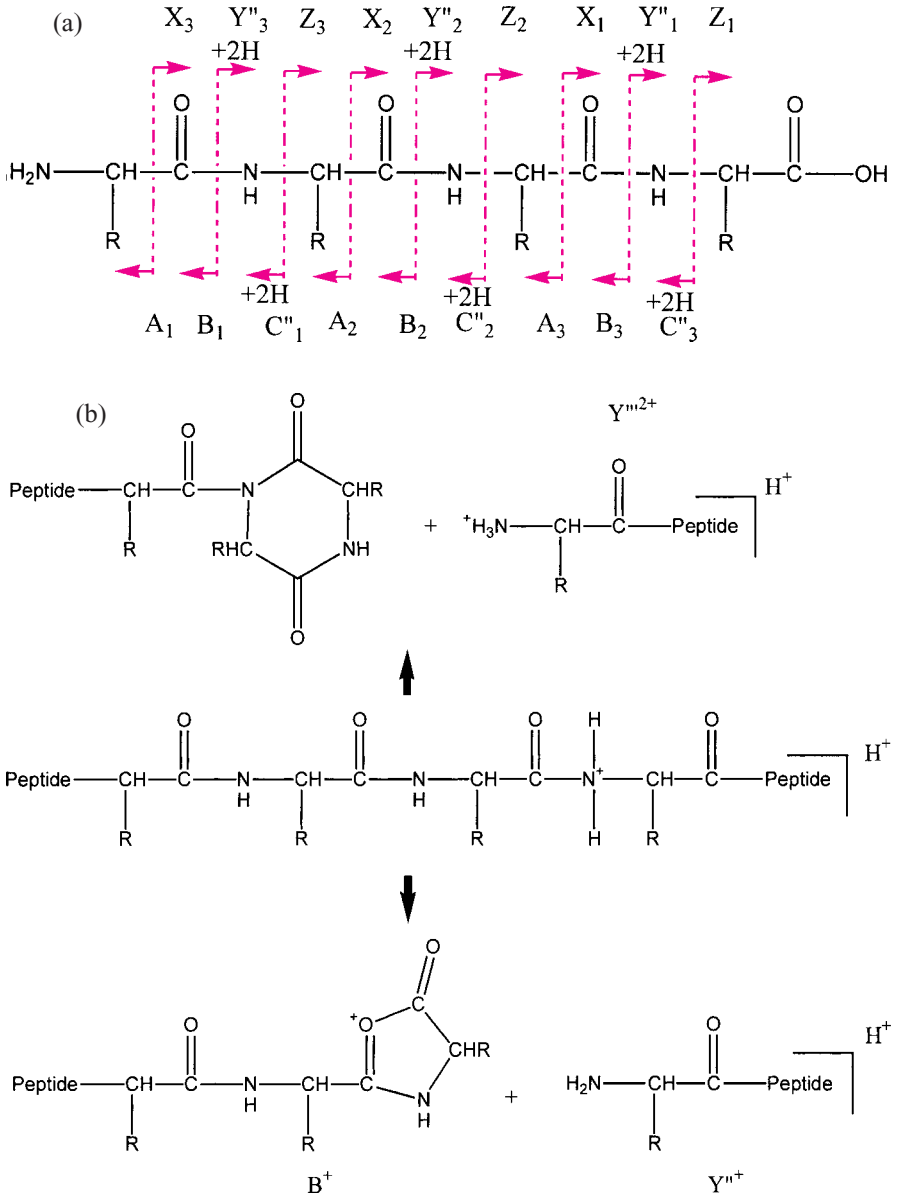


Figure 10 (a) Nomenclature for peptide fragmentation [116] and (b) schematic representation of the cleavage of peptide bonds via the mobile proton model

(i.e. the residues with high gas-phase basisity) and the second may be localized at one of the amide bonds (or the N-terminus). In the collision process this second proton can be mobilized, giving a heterogeneous population of ions in which the second proton may reside at any one of the peptide bonds (mobile proton model). Protonation of an amide nitrogen will weaken the amide bond and lead to cleavage, generating fragment ions. As a result of the heterogeneous population of precursor ions, a series of peptide bonds are cleaved, giving complementary series of Y- and B-type ions. B-type ions tend to fragment further to A-type, smaller B ions and internal immonium ions [117,118,123]. Fragmentations of peptide bonds N-terminal to Pro and C-terminal to both Asp and Glu are particularly prevalent and can readily be explained by application of the mobile proton model. Proline residues contain a tertiary nitrogen atom and are thus more basic than other residues, and hence more likely to be protonated by a mobile proton, thereby leading to enhanced peptide-bond cleavage [124]. The acidic residues Asp, Glu and cysteic acid also catalyse peptide-bond cleavage. The acidic side chains provide a local 'mobile' proton which catalyses cleavage at the peptide bonds C-terminal to these residues [125,126].

With the ever-growing amount of sequence information available in protein and nucleotide databases, MS is becoming ever more important in protein-identification protocols. As discussed above, proteins separated by two-dimensional gels can be identified from a peptide mass map of their tryptic peptides. Alternatively, a protein can be identified by performing MS/MS on one or more of its tryptic peptides and submitting the determined amino acid sequence to a database search [127–132]. The complete amino acid sequence of the peptide is not required – it may even contain errors – and yet a correct identification can be made. This is illustrated in the following example.

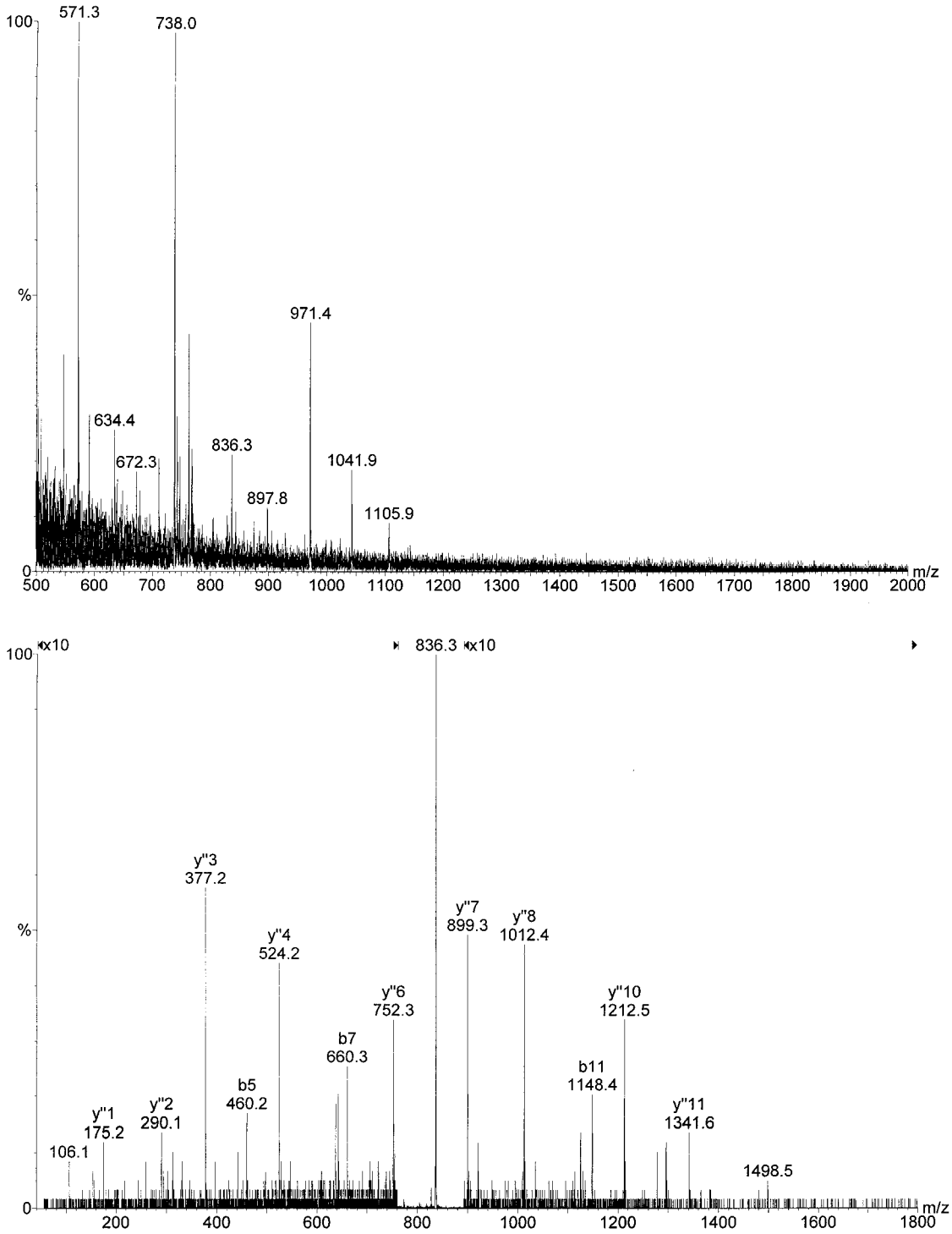


Figure 11 ES mass spectrum of the tryptic digest of a two-dimensional-gel spot (upper panel) and positive-ion CID spectrum of the $[M+2H]^{2+}$ ion at *m/z* 836.3 (lower panel)

Spectra were recorded on a Q-TOF mass spectrometer.

Application of peptide sequence tags in protein identification

The upper panel of Figure 11 shows an ES mass spectrum of the tryptic digest of a spot from a two-dimensional gel. The doubly charge ion at m/z 836.3 was selected for fragmentation. The lower panel of Figure 11 is the resultant CID spectrum. The sequence tag FLDFLLS was readily obtained by manual inspection of the spectrum. The separation of the successive peaks, i.e. 377.2–524.2 = F, 524.2–637.3 = L/I, 637.3–752.3 = D, 752.3–899.3 = F, 899.3–1012.4 = L/I, 1012.4–1125.5 = L/I, 1125.5–1212.5 = S, gives the sequence tag. By submission of the sequence tag (FLDFLLS) along with its starting (377.2) and end

(1212.5) m/z, plus the mass of the neutral parent (1670.7) to the PeptideSearch algorithm, (http://www.mann.embl-heidelberg.de/Services/PeptideSearch/PeptideSearchIntro.html), the following sequence was found GGSEESLLM*DLFSDR, where M* is oxidized methionine. This corresponds to a tryptic peptide of a protein suspected to be in the sample. Peaks in the lower panel of Figure 11 are annotated according to the proposed sequence. It should be noted that the database search is error-tolerant. Although manual interpretation gave the sequence tag to be FLDFLLS, the correct sequence is FLDM*LLS. Oxidized methionine (147.0354 Da) and phenylalanine (147.06841 Da) residues are isobaric (i.e. have the same molecular mass), and unless accurate mass measurements are made they are indistinguishable in low-energy CID spectra.

CID in the analysis of proteins

Early studies on the MS/MS of intact proteins were performed by Smith's group [133,134]. In collaboration with Shackleton's group [134] they studied the CID of Hb (haemoglobin) in its normal, and variant forms, and from the resultant spectra were able to identify the Hb variants. More recently, with the introduction of ECD, McLafferty's group have been applying FTICR MS to study the dissociation of large proteins [84].

PERSPECTIVES

The introduction of ES as a method of desorption of ions from solution into the gas phase for subsequent mass analysis revolutionized biological MS. Using ES, peptides can be analysed routinely at the subpicomole level, and multimeric proteins with molecular masses in the MDa range investigated. ES, while popularized in the study of peptides/proteins, can be used in the analysis of a wide range of biomolecules, including carbohydrates, lipids and nucleic acids. The combination of ES with MS/MS is immensely powerful for the structural characterization of biomolecules. Peptides extending up to whole proteins can now be sequenced. Post-translational modifications can be identified, ligand for receptors can be characterized and all classes of organic molecules investigated. With the ever-increasing improvements in mass-spectrometric technology, together with complementary improvements in computer software and hardware, MS is certain to play a pivotal role in biochemistry in the future.

We thank the Swedish Medical Research Council (grant no. 03X-12551) and the Karolinska Institutet for financial support.

REFERENCES

1 Thomson, J. J. (1913) Rays of Positive Electricity and their Applications to Chemical Analysis, Longmans Green, London
2 Aston, F. W. (1933) Mass Spectra and Isotopes, Edward Arnold, London
3 Chatman, K., Hollenbeck, T., Hagey, L., Vallee, M., Purdy, R., Weiss, F. and Siuzdak, G. (1999) Nanoelectrospray mass spectrometry and precursor ion monitoring for quantitative steroid analysis and attomole sensitivity. Anal. Chem. **71**, 2358–2363
4 Griffiths, W. J., Liu, S., Yang, Y., Purdy, R. H. and Sjövall, J. (1999) Nano-electrospray tandem mass spectrometry for the analysis of neurosteroid sulphates. Rapid Commun. Mass Spectrom. **13**, 1595–1610
5 Shevchenko, A., Wilm, M., Vorm, O. and Mann, M. (1996) Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. Anal. Chem. **68**, 850–858
6 Wilm, M., Shevchenko, A., Houthaeve, T., Brelt, S., Schweigerer, L., Fotsis, T. and Mann, M. (1996) Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. Nature (London) **379**, 466–469
7 Andrén, P. E., Emmett, M. R. and Caprioli, R. M. (1994) Micro-electrospray: zeptomole/attomole per microlitre sensitivity for peptides. J. Am. Soc. Mass Spectrom. **5**, 867–869
8 Karas, M. and Hillenkamp, F. (1988) Ultraviolet laser desorption of ions above 10 kDa. In Abstracts of the 11th International Mass Spectrometry Conference, Bordeaux, abstract MOA OR, 29 August–2 September 1988
9 Karas, M. and Hillenkamp, F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10000 daltons. Anal. Chem. **60**, 1299–2301
10 Yamashita, M. and Fenn, J. B. (1984) Electrospray ion source. Another variation on the free-jet theme. J. Phys. Chem. **88**, 4451–4459
11 Yamashita, M. and Fenn, J. B. (1984) Negative ion production with the electrospray ion source. J. Phys. Chem. **88**, 4671–4675
12 Meng, C. K., Mann, M. and Fenn, J. B. (1988) Electrospray ionization of some polypeptides and small proteins. In Proceedings of the 36th American Society for Mass Spectrometry Conference on Mass Spectrometry and Allied Topics, San Francisco, CA, June 1988, pp. 711–772, ASMS, Sante Fe, NM
13 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
14 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
15 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
16 de Hoffmann, E. (1996) Tandem Mass Spectrometry: a Primer. J. Mass Spectrom. **31**, 129–137
17 McLafferty, F. W. (1983) Tandem Mass Spectrometry, John Wiley and Sons, New York
18 Farmer, P. B. and Sweetman, G. M. A. (1995) Mass spectrometric detection of carcinogen adducts. J. Mass Spectrom. **30**, 1369–1379
19 Murray, K. K. (1996) DNA sequencing by mass spectrometry. J. Mass Spectrom. **31**, 1203–1215
20 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. U.S.A. **83**, 6233–6237
21 Papayannopoulos, I. A. (1995) The interpretation of collision-induced dissociation tandem mass spectra of peptides. Mass Spectrom. Rev. **14**, 49–73
22 Teng-umnauy, P., Morris, H. R., Dell, A., Panico, M., Paxton, T. and West, C. M. (1998) The cytoplasmic F-box binding protein SKP1 contains a novel pentasaccharide linked to hydroxyproline in *Dictyostelium*. J. Biol. Chem. **273**, 18242–18249
23 Lacko, A. G., Reason, A. J., Nuckolls, C., Kudchodkar, B. J., Nair, M. P., Suundarrajan, G., Pritchard, P. H., Morris, H. R. and Dell, A. (1998) Characterization of recombinant human plasma lecithin:cholesterol acetyltransferase (LCAT): N-linked carbohydrate structures and catylytic properties. J. Lipid Res. **39**, 807–820
24 Smith, R. D. and Light-Wahl, K. J. (1993) The observation of non-covalent interactions in solution by electrospray ionization mass spectrometry: promise, pitfalls and prognosis. Biol. Mass Spectrom. **22**, 493–501
25 Pramanik, B. N., Bartner, P. L., Mirza, U. A., Liu, Y.-H. and Ganguly, A. K. (1998) Electrospray ionization mass spectrometry for the study of non-covalent complexes: an emerging technology. J. Mass Spectrom. **33**, 911–920
26 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
27 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
28 Caprioli, R. M. (1990) Continuous-Flow Fast Atom Bombardment mass spectrometry, John Wiley and Sons, Chichester
29 Roepstorff, P. (2000) MALDI-TOF mass spectrometry in protein chemistry. In Proteomics in Functional Genomics: Protein Structure Analysis (Jollès, P. and Jörnvall, H., eds.), pp. 81–97, Birkhäuser Verlag, Basel
30 Nelson, R. W., Dogruel, D. and Williams, P. (1994) Mass determination of human immunoglobulin IgM using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Commun. Mass Spectrom. **8**, 627–631
31 Henzel, W. J., Stultz, J. T. and Wananabe, C. (1989) American Society for Mass Spectrometry Conference, Seattle, WA, July 29–August 2 1989, ASMS, Santa Fe, NM
32 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. U.S.A. **90**, 5011–5015
33 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
34 Pappin, D. J. C., Højrup, P. and Bleasby, A. J. (1993) Protein identification by peptide mass fingerprinting. Curr. Biol. **3**, 327–332

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

36 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

37 Spengler, B. (1997) Post-source decay analysis in matrix-assisted laser desorption/ionization mass spectrometry of biomolecules. J. Mass Spectrom. **32**, 1019–1036

38 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

39 Gaskell, S. J. (1997) Electrospray: principles and practice. J. Mass Spectrom. **32**, 677–688

40 Kebarle, P. and Ho, Y. (1997) On the mechanism of electrospray mass spectrometry. In Electrospray Ionization Mass Spectrometry: Fundamentals, Instrumentation and Applications (Cole, R. B., ed.), pp. 3–63, John Wiley and Sons, New York

41 Iribarne, J. V. and Thomson, B. A. (1976) On the evaporation of small ions from charged droplets. J. Chem. Phys. **64**, 2287–2294

42 Thomson, B. A. and Iribarne, J. V. (1979) Field induced ion evaporation from liquid surfaces at atmospheric pressure. J. Chem. Phys. **71**, 4451–4463

43 Kebarle, J. (2000) A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry. J. Mass Spectrom. **35**, 804–817

44 Tang, L. and Kebarle, P. (1991) Effect of the conductivity of the electrosprayed solution on the electrospray current. Factors determining analyte sensitivity in electrospray mass spectrometry. Anal. Chem. **63**, 2709–2715

45 Tang, L. and Kebarle, P. (1993) Dependence of ion intensity in electrospray mass spectrometry on the concentration of the analyte in the electrosprayed solution. Anal. Chem. **65**, 3654–3668

46 Kebarle, P. and Tang, L. (1993) From ions in solution to ions in the gas phase. Anal. Chem. **65**, 972A–986A

47 Taylor, G. I. (1964) Disintegration of water drops in an electric field. Proc. R. Soc. London Ser. A. **280**, 383–397

48 Electrospray Ionization Mass Spectrometry: Fundamentals, Instrumentation and Applications (Cole, R. B., ed.), John Wiley and Sons, New York

49 Cole, R. B. (2000) Some tenets pertaining to electrospray ionization mass spectrometry. J. Mass Spectrom. **35**, 763–772

50 Van Berkel, G. J. (2000) Electrolytic deposition of metals on to the high-voltage contact in an electrospray emitter: implications for gas-phase ion formation. J. Mass Spectrom. **35**, 773–783

51 Amad, M. H., Cech, N. B., Jackson, G. S. and Enke, C. G. (2000) Importance of gas-phase proton affinities in determining the electrospray ionization response for analytes and solvents. J. Mass Spectrom. **35**, 784–789

52 Gamero-Castaño, M. and Fernandez de la Mora, J. (2000) Kinetics of small ion evaporation from the charge and mass distribution of multiply charged clusters in electrosprays. J. Mass Spectrom. **35**, 790–803

53 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

54 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

55 Wilm, M. S. and Mann, M. (1994) Electrospray and Taylor-cone theory: Dole's beam of macromolecules at last? Int. J. Mass Spectrom. Ion Processes **136**, 167–180

56 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 (Carr, S. A., Burlingame, A. L. and Baldwin, M. A., eds.), pp. 553–561, Humana Press, Totowa

57 Gross, M. L. (1994) Accurate masses for structure conformation. J. Am. Soc. Mass Spectrom. **5**, 57

58 Beynon, J. H. (1954) Qualitative analysis of organic compounds by mass spectrometry. Nature (London) **174**, 735–737

59 Beynon, J. H. (1960) mass spectrometry and its Applications to Organic Chemistry, Elsevier, Amsterdam

60 Russell, D. H. and Edmondson, R. D. (1997) High-resolution mass spectrometry and accurate mass measurements with emphasis on the characterization of peptides and proteins by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. J. Mass Spectrom. **32**, 263–276

61 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

62 Griffiths, W. J., Hjertman, M., Wejde, J., Larsson, O. and Sjövall, J. (1997) Tandem mass spectrometry of geranylgeranylcysteine. J. Mass Spectrom. **32**, 899–903

63 Hamberg, M., Liepinsh, E., Otting, G. and Griffiths, W. J. (1998) Isolation and structure of a new galactolipid from oat seeds. Lipids, **33**, 355–363

64 Yang, Y., Griffiths, W. J., Nordling, M., Nygren, J., Möller, L., Liepinsh, E., Otting, G., Gustafsson, J., Rafter, J. and Sjövall, J. (2000) Ring opening of benzo[*a*]pyrene in the germfree rat is a novel pathway for formation of potentially genotoxic metabolites. Biochemistry **39**, 15585–15591

65 Griffiths, W. J., Gustafsson, M., Yang, Y., Curstedt, T., Sjövall, J. and Johansson, J. (1989) Analysis of variant forms of porcine surfactant polypeptide-C by nano-electrospray mass spectrometry. Rapid Commun. Mass Spectrom. **12**, 1104–1114

66 Meng, L. J., Griffiths, W. J. and Sjövall, J. (1996) Tandem mass spectrometry in the identification of a novel steroid-like metabolite: the beneficial effects of coffee on liver disease. In Abstracts of the 22nd Annual Conference of the British Mass Spectrometry Society, p. 22, University College of Wales, Swansea, 8–11 September 1996

67 Morgan, R. P., Beynon, J. H., Bateman, R. H. and Green, B. N. (1978) The MM-ZAB-2F double focusing mass spectrometer and mike spectrometer. Int. J. Mass Spectrom. Ion Phys. **28**, 171–191

68 Clauser, K. R., Baker, P. and Burlingame, A. L. (1999) Role of accurate mass measurements ($\pm$10 ppm) in protein identification strategies employing MS or MS/MS and data base searching. Anal. Chem. **71**, 2871–2882

69 Johansson, J., Wang, Y., Griffiths, W. J. and Curstedt, T. (1999) Porcine pulmonary surfactant preparations contain the antibacterial peptide prophenin and a C-terminal 18-residue fragment thereof. FEBS Lett. **460**, 257–262

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

71 Harwig, S. S. L., Kokryakov, V. N., Swiderek, K. M., Aleshina, G. M., Zhao, C. and Lehrer, R. L. (1995) Prophenin-1, an exceptionally proline-rich antimicrobial peptide from porcine leukocytes. FEBS Lett. **362**, 65–69

72 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

73 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

74 Nettleton, E. J., Sunde, M., Lai, Z., Kelly, J. W., Dobson, C. M. and Robinson, C. V. (1998) Protein subunit interactions and structural integrity of amyloidogenic transthyretins: evidence from electrospray mass spectrometry. J. Mol. Biol. **281**, 553–564

75 Loo, J. A. (1995) Observation of large subunit protein complexes by electrospray ionization mass spectrometry. J. Mass Spectrom. **30**, 180–183

76 Witkowska, H. E., Green, B. N., Carlquist, M. and Shackleton, C. H. (1996) Intact noncovalent dimer of estrogen receptor ligand-binding domain can be detected by electrospray ionization mass spectrometry. Steroids. **61**, 433–438

77 Ganem, B., Li, Y.-T. and Henion, J. D. (1991) Observation of noncovalent enzyme–substrate and enzyme–product complexes by ion-spray mass spectrometry. J. Am. Chem. Soc. **113**, 7818–7819

78 Ishigai, M., Langridge, J., Bordoli, R. S. and Gaskell, S. J. (2000) Noncovalent associations of glutathione S-transferase and ligands: a study using electrospray quadrupole/time-of-flight mass spectrometry. J. Am. Soc. Mass Spectrom. **11**, 606–614

79 Liepinsh, E., Kitamura, M., Murakami, T., Nakaya, T. and Otting, G. (1997) Pathway of chymotrypsin evolution suggested by the structure of the FMN-binding protein from *Desulfovibrio vulgaris* (Miyazaki F). Nat. Struct. Biol. **4**, 975–979

80 Suto, K., Kawagoe, K., Shibata, N., Morimoto, Y., Higuchi, Y., Kitamura, M., Nakaya, T. and Yasuoka, N. (1999) Crystallization and preliminary crystallographic studies of FMN-binding protein from *Desulfovibrio vulgaris* Miyazaki F. Acta Crystallogr. **D55**, 1089–1090

81 Griffiths, W. J., Otting, G. and Sjövall, J. (1998) Analysis of the FMN-binding protein from *Desulfovibrio vulgaris* by electrospray mass spectrometry. In Proceedings of the 46th American Society for Mass Spectrometry Conference on mass spectrometry and Allied Topics, Orlando, FL, 31 May–4 June 1998, p. 211, ASMS, Santa Fe, NM

82 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

83 Schnier, P. D., Price, W. D., Jockush, R. A. and Williams, E. R. (1996) Blackbody infrared radiative dissociation of bradykinin and its analogues: energetics, dynamics, and evidence for salt-bridge structures in the gas phase. J. Am. Chem. Soc. **118**, 7178–7189

84 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

85 Zubarev, R. A., Krunger, N. A., Fridriksson, E. K., Lewis, M. A., Horn, D. M., Carpenter, B. K. and McLafferty, F. W. (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

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

87 Cooks, R. G., Beynon, J. H., Caprioli, R. M. and Lester, G. R. (1973) Metastable Ions, Elsevier, Amsterdam

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

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

90 Amy, J. W., Baitinger, W. E. and Cooks, R. G. (1990) Building mass spectrometers and a philosophy of research. J. Am. Soc. Mass Spectrom. **1**, 119–128

91 Yost, R. A. and Enke, C. G. (1978) Selected ion fragmentation with a tandem quadrupole mass spectrometer. J. Am. Chem. Soc. **100**, 2274–2275

92 March, R. E. (1997) An introduction to quadrupole ion trap mass spectrometry. J. Mass Spectrom. **32**, 351–369

93 Comisarow, M. B. and Marshall, A. G. (1974) Fourier transform ion cyclotron resonance spectroscopy. Chem. Phys. Lett. **25**, 282–283

94 Amster, I. J. (1996) Fourier transform MS. J. Mass Spectrom. **31**, 1325–1337

95 Clayton, E. and Bateman, R. H. (1992) Time-of-flight mass analysis of high energy collision-induced dissociation fragment ions. Rapid Commun. Mass Spectrom. **6**, 719–720

96 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**, 1277–1233

97 Morris, H. R., Paxton, T., Dell, A., Langhorn, B., Berg, M., Bordoli, R. S., Hoyes, J. and Bateman, R. H. (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

98 Hopfgartner, G., Chernushevich, I. V., Covey, T., Plomley, J. B. and Bonner, R. (1999) Exact mass measurement of product ions for the structural elucidation of drug metabolites with a tandem quadrupole orthogonal-acceleration time-of-flight mass spectrometer. J. Am. Soc. Mass Spectrom. **10**, 1305–1314

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

100 Gross, M. L. (1992) Charge-remote fragmentations: method, mechanism and applications. Int. J. Mass Spectrom. Ion Processes **118/119**, 137–165

101 Adams, J. (1990) Charge-remote fragmentations: analytical applications and fundamental studies. Mass Spectrom. Rev. **9**, 141–186

102 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 Processes **86**, 137–154

103 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

104 Bordoli, R. S. and Bateman, R. H. (1992) Int. J. Mass Spectrom. Ion Processes **122**, 243

105 Tomer, K. B., Crow, F. W. and Gross, M. L. (1983) Location of double bond position in unsaturated fatty acids by negative ion MS/MS. J. Am. Chem. Soc. **105**, 5487–5488

106 Griffiths, W. J., Yang, Y., Lindgren, J. Å. and Sjövall, J. (1996) Charge remote fragmentation of fatty acid anions in 400 eV collisions with xenon atoms. Rapid Commun. Mass Spectrom. **10**, 21–28

107 Wysocki, V. H. and Ross, M. M. (1991) Charge-remote fragmentation of gas-phase ions: mechanistic and energetic considerations in the dissociation of long-chain functionalized alkanes and alkenes. Int. J. Mass Spectrom. Ion Processes. **104**, 179–211

108 Kerwin, J. L., Wiens, A. M. and Ericsson, L. H. (1996) Identification of fatty acids by electrospray mass spectrometry and tandem mass spectrometry. J. Mass Spectrom. **31**, 184–192

109 Mata de Urquiza, A., Liu, S., Sjöberg, M., Zetterström, R., Griffiths, W. J., Sjövall, J. and Perlmann, T. (2000) Docosahaxaenoic acid, a ligand for the retinoid X receptor in mouse brain. Science **290**, 2140–2144

110 Tomer, K. B., Jensen, N. J., Gross, M. L. and Whitney, J. (1986) Fast atom bombardment with tandem mass spectrometry for determination of bile salts and their conjugates. Biomed. Environ. Mass Spectrom. **13**, 265–272

111 Libert, R., Hermans, D., Draye, J. P., Van Hoof, F., Sokal, E. and de Hoffmann, E. (1991) Bile acids and conjugates identified in metabolic disorders by fast atom bombardment and tandem mass spectrometry. Clin. Chem. **37**, 2102–2110

112 Griffiths, W. J., Brown, A., Reimendal, R., Yang, Y., Zhang, J. and Sjövall, J. (1996) A comparison of fast atom bombardment and ES as methods of ionization in the study of sulphated- and sulphonated-lipids by tandem mass spectrometry. Rapid Commun. Mass Spectrom. **10**, 1196–1174

113 Yang, Y., Griffiths, W. J., Nazer, H. and Sjövall, J. (1997) Analysis of bile acids and bile alcohols in urine by capillary column liquid chromatography–mass spectrometry using fast atom bombardment or electrospray ionization and collision induced dissociation. Biomed. Chromatgr. **11**, 240–255

114 Stroobant, V. and de Hoffmann, E. (1995) Fast-atom bombardment mass spectrometry and low energy collision-induced dissociation tandem mass spectrometry of tauroconjugated bile acid anions. J. Am. Soc. Mass Spectrom. **6**, 588–596

115 Tomer, K. B. and Gross, M. L. (1988) Fast atom bombardment and tandem mass spectrometry for structural determination: remote site fragmentation of steroid conjugates and bile salts. Biomed. Environ. Mass Spectrom. **15**, 89–98

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

117 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

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

119 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 Processes, **164**, 137–153

120 Arnott, D., Kottmeier, D., Yates, N., Shabanowitz, J. and Hunt, D. F. (1994) Fragmentation of multiply protonated peptides under low energy conditions. In Proceedings of the 42nd American Society for Mass Spectrometry Conference on mass spectrometry and Allied Topics, Chicago, IL, 29 May–3 June 1994, p. 470, ASMS, Santa Fe, NM

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

122 Dongré, A. R., Jones, J. L., Somogyi, Á. and Wysocki, V. H. (1996) Influence of peptide composition, gas-phase basicity, and chemical modification on fragmentation efficiency: evidence for the mobile proton model. J. Am. Soc. Mass Spectrom. **118**, 8365–8374

123 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

124 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

125 Tsaprailis, G., Hair, H., Somogyi, Á., Wysocki, V. H., Zhong, W., Futrell, J. H., Sommerfield, S. G. and Gaskell, S. J. (1999) Influence of secondary structure on the fragmentation of protonated peptides. J. Am. Chem. Soc. Spectrom. **121**, 5142–5154

126 Harrison, A. G. and Tu, Y.-P. (1998) Ion chemistry of protonated aspartic acid derivatives. J. Mass Spectrom. **33**, 532–542

127 Mann, M. and Wilm, M. (1994) Error-tolerant identification of peptides in sequence databases by peptide sequence tags. Anal. Chem. **66**, 4390–4399

128 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. U.S.A. **93**, 8264–8267

129 Shevchenko, A., Jenson, O. N., Podtelejnikov, A. V., Sagliocco, F., Wilm, M., Vorm, O., Mortensen, P., Shevchenko, A., Boucherie, H. and Mann, M. (1996) Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. Proc. Natl. Acad. Sci. U.S.A. **93**, 14440–14445

130 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

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

132 Yates, J. R. (1998) Mass spectrometry and the age of the proteome. J. Mass Spectrom. **33**, 1–19

133 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

134 Light-Wahl, K. J., Loo, J. A., Smith, R. D., Witkowska, H. E., Shackleton, C. H. and Wu, C. S. (1993) Collisionally activated dissociation and tandem mass spectrometry of intact hemoglobin beta-chain variant proteins with electrospray ionization. Biol. Mass Spectrom. **22**, 112–120