

Mass spectrometry technologies for proteomics

Benito Cañas, Daniel López-Ferrer, Antonio Ramos-Fernández, Emilio Camafeita and Enrique Calvo

Advance Access publication date 3 February 2006

Abstract

In the late 1980s, the advent of soft ionization techniques capable of generating stable gas phase ions from thermally unstable biomolecules, namely matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI), laid the way for the development of a set of powerful alternatives to the traditional Edman chemistry for the structural characterization of peptides and proteins. The rapid protein identification capabilities that, coupled with two-dimensional gel electrophoresis, provided insights into all sorts of biological systems since the dawn of proteomics and have been exploited in the last few years for the development of more powerful and automatable gel-free strategies, mainly based on multidimensional chromatographic separations of peptides from proteolytic digests. In parallel to the evolution of ion sources, mass analysers and scan modes, the invention of new elegant biochemical strategies to fractionate or simplify highly complex mixtures, or to introduce isotopic labels in peptides in a variety of ways now makes also possible large-scale, high-coverage quantitative studies in a wide dynamic range. In this review, we provide the fundamental concepts of mass spectrometry (MS) and describe the technological progress of MS-based proteomics since its earliest days. Representative literature examples of their true power, either when employed as exploratory or as targeted techniques, is provided as well.

Keywords: proteomics; mass spectrometry; post-translational modifications; second generation proteomics

INTRODUCTION

Mass spectrometry (MS) has been widely used as an analytical technique since its establishment more than one hundred years ago [1]. Its sensitivity and selectivity are excellent and, either the molecular weight (MW) or structural information of a compound can be obtained in a short time. Despite these advantages, until the early 1990s this technique was seldom applied to the study of peptides and proteins. The methodology, based in automated Edman degradation, was used for the identification and sequencing of proteins [2] and sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) was the method of choice for MW determination [3].

MS instrumentation separates and detects the ions in gas phase and therefore, prior to any separation by MS, molecules must be ionized and converted into gas using different techniques. Most of the ionization techniques in use at this time were not applicable to peptide and protein molecules because it was not possible to convert them into intact ions. It can be intuitively realized that it is not easy to ionize and take into gas phase large and polar molecules so that any fragmentation or decomposition occurs.

Fast atom bombardment (FAB) [4], plasma desorption (PD) [5] and thermospray [6] were the unique ionization techniques allowing the work

Corresponding author. Benito Cañas Montalvo, Dpto. de Química Analítica, Facultad de C. Químicas, Universidad Complutens, E-28040, Madrid, Spain. Tel: +34 91 394 43 68; Fax: +34 91 394 43 29; E-mail: bcanasmo@quim.ucm.es

Benito Cañas is senior research scientist at the Universidad Complutense de Madrid, where he completed his PhD. His research interests are focused on the development of methodologies for the study of seleno-proteins.

Daniel López-Ferrer and **Antonio Ramos-Fernández** work in the Proteomics and Protein Chemistry laboratory at the CBM-SO in Madrid. They are engaged in the development of new methodologies for second generation proteomics.

Enrique Calvo and **Emilio Camafeita** are senior technicians in the Proteomics Unit at the Centro Nacional de Investigaciones Cardiovasculares (CNIC) in Madrid. They are focused in the development of methodologies for protein post-translational modification characterization.

with proteins. They were called 'soft' ionization techniques because they were able to convert large molecules into gas phase without affecting their integrity and constituted a very important development. For the first time, it was possible to calculate easily the MW of proteins with certain exactitude, as well as to study peptide fragmentation behaviour under a collision-induced dissociation (CID) [7] regimen. Sensitivity was the main problem: the quantities needed to perform an analysis, much higher than those needed for Edman sequencing, made these ionization techniques unsuitable for routine protein identification.

Only a few laboratories were experimenting with the application of MS as an instrumental tool for analytical and structural studies of proteins. In the 1980s, using FAB as ionization technique coupled with a four sector instrument, Biemann intensively studied the fragmentation patterns of peptides [8], acquiring a knowledge that allowed the classification of these fragments [9]. At that time, Hunt and collaborators were able to sequence the first protein using MS exclusively [10]. Again an FAB ionization source was employed.

The most important breakthrough, making possible the complete application of MS to the study and analysis of proteins, was the development of two new ionization techniques in the late 1980s. Electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) allowed for the first time the acquisition of mass spectra with minute quantities of peptides and proteins. In 1988, Fenn *et al.* demonstrated that it was possible to obtain the mass spectrum of proteins with only a few picomoles [11]. The same year, Karas and Hillenkamp published spectra of proteins as large as 10 kDa, obtained coupling their newly developed MALDI source to a time-of-flight (TOF) mass spectrometer [12].

At the protein society meeting, held in 1989, Henzel *et al.* presented the germinal work of the methodology that, 4 years later, was to be called peptide mass fingerprinting (PMF) [13, 14]. These authors used an FAB source coupled with a tandem sector instrument to obtain spectra of protein digests using sample quantities between 100 pmols and 1 nmol. Despite the huge protein amounts employed as compared with the femtomole quantities used presently, they demonstrated that using the mass spectrum of a protein digest and the available rudimentary protein databases (20 000 entries in 1989 versus 2.5 million entries in 2005), it was

possible to identify a protein in a few minutes. The third piece of the puzzle was a search software capable of comparing the experimental mass spectrum taken from a protein digest with the theoretical spectra produced by the proteins in the database if subjected to the same digestion conditions. The foundations for the development of proteomics had been laid.

In 1993 five groups published their results in protein identification using MALDI-TOF instruments and their own search programs [15–19]. By this time the exponential growth in database entries situated them in 100 000, establishing the possibility of identification of thousands of proteins. In principle, any protein can be identified with this procedure provided its sequence is annotated in any protein database.

Nevertheless, for different reasons proteins included in databases were not always identified by PMF. In such cases, Edman chemical degradation or mass spectrometric CID fragmentation were the alternatives. Although the expertise in the first technique was widespread, not many experts in interpreting peptide fragmentation spectra existed at this time. In 1994, two special search programs, SEQUEST [20, 21] and PeptideSearch [22], were developed to identify proteins using peptide fragmentation spectra after chemical or enzymatic digestion. Using any of these programs, the interpretation of fragmentation spectra was unnecessary, as they enabled the automated comparison of measured fragmentation spectra with those expected under the same collision conditions for every peptide obtained by the digestion of each protein present in the databases interrogated. One advantage of this methodology for protein identification was the possibility of using the emerging EST databases [23, 24], built up with fast partial cDNA sequencing and containing large amounts of information not useful for PMF.

These two procedures, PMF and searches with uninterpreted peptide fragmentation spectra, constitute the methodology used in most proteomics laboratories during the last 10 years. Many proteins have been identified in this way and a number of publications show the usefulness of this approach, now renamed as 'classical proteomics'.

However, modern biology needs an even faster production of data and proteomics techniques are in the process of adaptation to these demands [25]. High throughput protein identification, quantitative differences in protein expression and the global study of post-translational modifications are among others,

increasing demand by many researchers. New methodologies are under development, being used routinely by specialized laboratories and in the process of establishing by others. Multidimensional chromatography protein identification (MudPit) [26] and isotope labelling [27] are techniques that may solve new research demands and many groups are now moving on to this methodology that has been called 'Second generation Proteomics', in contrast to the well known and widely used Classical Proteomics.

The present review is focused on the description and functionalities of the MS instrumentation that has made possible the development of proteomics methodologies, as well as of the development of new MS instruments that run in parallel with the advances in the discipline of proteomics. Publications of the major achievements in the last 10 years will be reviewed with emphasis on the instrumentation used.

MASS SPECTROMETRY

General aspects

Mass spectrometers are instruments capable of producing and separating ions according to their mass-to-charge ratio (m/z). To make separations possible, electric or magnetic fields are generated inside the instrument. These fields separate the ions influencing their spatial trajectories, velocity and/or direction. The effects of an electromagnetic field on the ion movement are inversely proportional to the mass of the ion and directly proportional to its electrical charge. Then, a mass spectrum of a compound consists in a plot of ion abundance versus its m/z ratio, from which the MW can be deduced.

To be effective, ion separation must be done under vacuum to avoid the collision between accelerated ions and air molecules. Ion trajectories decrease as pressure increases and the mean distance run by an ion after acceleration may decrease from more than one metre to less than a millimetre by decreasing the vacuum from 10^{-6} to 10^{-3} torr.

Ionization source, analyser, detector, data processor and vacuum pumps are the main components in a mass spectrometer (Figure 1). As previously stated, the analyser and the detector must be under vacuum, although ion production can also be done at atmospheric pressure depending on the type of ionization source.

To be separated by electromagnetic fields, neutral molecules must be converted into ions and, if necessary, transferred into gas phase, a process

that takes place in an ionization source. Given the high MW and polarity of peptides and proteins, the ionization source not only has to ionize them, but also has to produce desolvated ions in gaseous phase (desorption). Among the diverse ionization sources available for desorption, only ESI and MALDI produce peptide and protein ions efficiently (Figure 1). In ESI, ions are formed at atmospheric pressure, while ions in MALDI they may be generated either at atmospheric pressure [28] or under vacuum conditions, although the best performance is obtained when working at low pressure.

The analyser is the instrumental part that separates the ions obtained at the ion source. Either electric or magnetic fields can be applied for the separation. All the existing analysers can be used for working with peptides and proteins, if coupled with the adequate ion source. Due to their ease of use and low price, the analysers mostly used are quadrupoles, TOF instruments and ion traps, either alone or combined in hybrid instruments (Figure 1a–f).

As an analytical instrument, one of the most important characteristics of mass analysers is their resolution, i.e. their capacity to differentiate two close signals (Figure 2). In MS, resolution is given by the equation $R = \Delta m/M$, where M is the mass of a compound and Δm is the width of the mass peak and can be taken at different peak heights. Frequently, resolution is measured taking Δm as the full peak width at half maximum (FWHM). As mass resolution increases, there is an improvement on the exactitude with which the m/z of an ion can be taken.

The mass spectra of peptides, as well as those of every organic substance, are composed of several peaks separated by a constant m/z value. These peaks represent the m/z values of the different isotopic peaks of a given molecule and are originated by the presence of $\sim 1\%$ of ^{13}C in all the organic molecules (Figure 2a). An instrument with enough resolution will be able to resolve the isotope peaks for a given substance. As the MW of a compound increases, an instrument with higher resolution will be needed to have isotope peaks completely separated, i.e. to obtain isotopic resolution.

Ionization sources

Matrix-assisted laser desorption/ionization

MALDI [12] relies on the utilization of a matrix compound capable of absorbing ultraviolet (UV) light. The matrix and the sample (e.g. a peptide

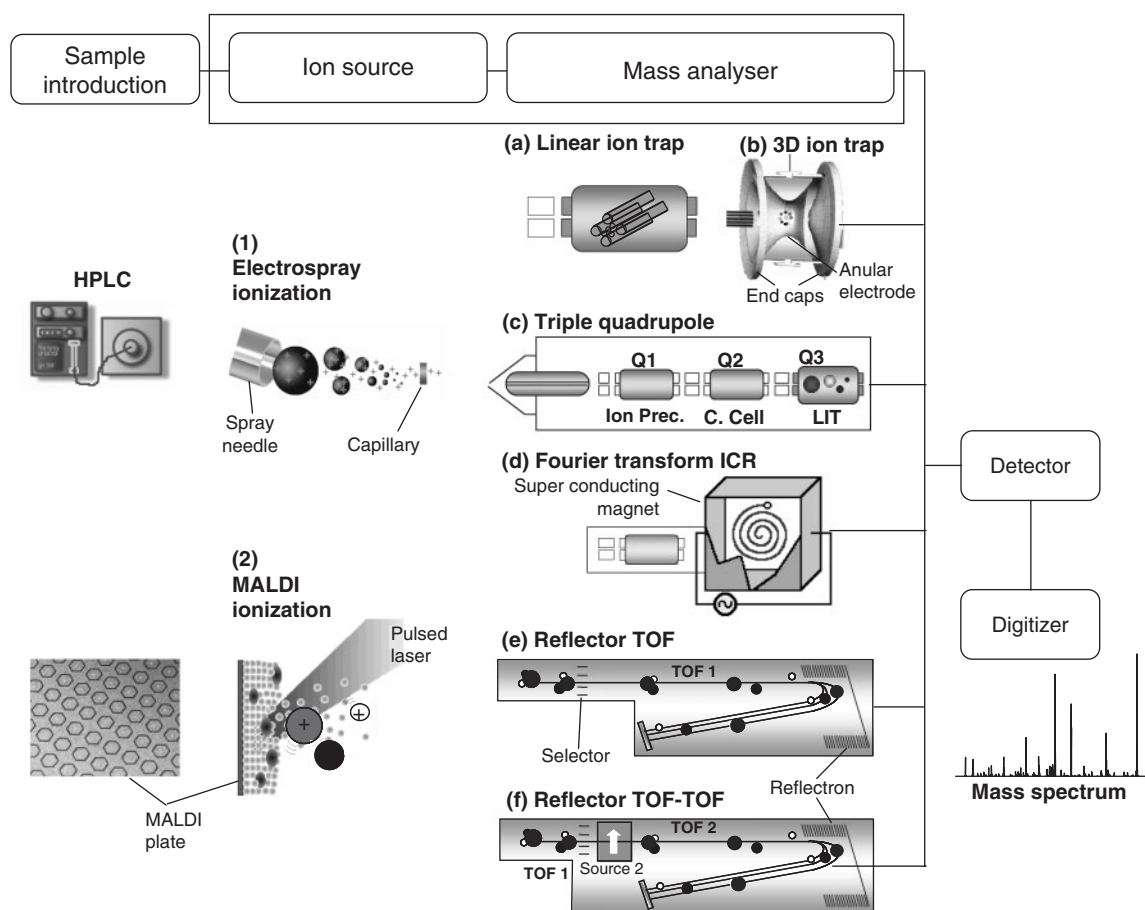


Figure 1: Main components of a mass spectrometer. Sample introduction device, ionization source for ion generation, mass analyser for ion separation and ion detector to transform analogue signals into digital signals and record a mass spectrum. Common ionization sources for proteomic research are electrospray (ESI, 1) and matrix-assisted laser desorption/ionization (MALDI, 2). Widespread mass analysers are ion traps (a) linear, and (b) three-dimensional; (c) triple quadrupoles; (d) Fourier transform cyclotrons; (e) and (f) time-of-flight (TOF). Ion trap and quadrupole analysers are normally coupled to ESI ion sources, whereas TOF analysers are usually employed with MALDI ion sources. Nevertheless, hybrid systems are available with alternative ion source/mass analyser combinations (e.g. ESI–TOF and MALDI–QqTOF).

mixture) are mixed in the appropriate solvent with approximately 10 000-fold molar excess of matrix and deposited onto a sample probe that normally allows the measurement of dozens to hundreds of samples. The solvent is allowed to evaporate and co-crystallized analyte molecules embedded in matrix crystals are obtained. The sample probe is then placed into the MS at high vacuum. The MALDI process is depicted in Figure 1(2). When the probe is hit by the pulsed UV laser beam, the energy is absorbed by the matrix, which is partially vaporized and carries intact analyte molecules into the gas phase [29]. During the expansion of the MALDI plume, protons are exchanged between analytes and matrix molecules, resulting in the formation of

positively and negatively charged analyte molecules in a process that is not yet fully understood [30].

MALDI is a competitive process in which the ionization of an analyte may be inhibited dramatically by the presence of others [31]. Thus, in tryptic peptide mixtures, arginine-containing peptides ionize preferentially due to the strong gas phase basicity of this amino acid [32, 33]. Chemical procedures have been developed to modify lysine in order to increase its basicity [34]. Conversely, phosphopeptides, due to the acidic properties of the phosphate group, ionize poorly with MALDI [35].

Sinapinic and α -cyano-4-hydroxycinnamic acids are, respectively, the most used matrices for protein and peptide analysis, although 2,5-dihydroxybenzoic

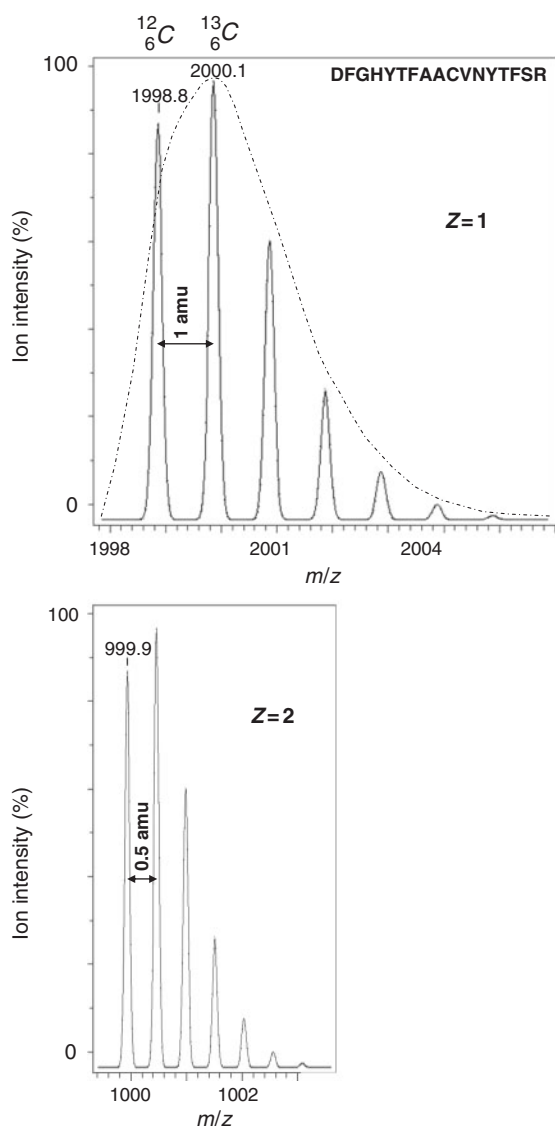


Figure 2: Isotopic resolution and determination of charge state. Mass measurement of the high-resolution protonated molecular mass of a peptide with the sequence DFGHYTFAACVLYTFSR yields an isotopic pattern reflecting the natural $^{12}\text{C}/^{13}\text{C}$ ratio. At lower resolution, an envelope centered at a higher m/z ratio value is obtained as depicted by a dashed line (Top). The ion charge state can be directly inferred by measuring the exact distance between adjacent isotopic components (1 amu for $z = 1$, top; 0.5 amu for $z = 2$, bottom).

acid is frequently used as well for peptide mapping [36]. Chemical additives to matrices are claimed to perform better for phosphopeptides [37], or to enhance the MALDI process in general [38]. Several neutral detergents, like octyl-beta-glucoside [39, 40], are highly compatible with peptide ionization and may enhance it, while charged detergents as SDS [41] have a deleterious effect. In any case, MALDI

has the advantage over ESI that it is more tolerant to sample contaminants like buffers and salts, when these are present at low concentrations. The spectra in the presence of chaotropes, like urea, may be taken even at concentrations as high as 2M.

Although nitrogen UV lasers are the most commonly used, infrared lasers have demonstrated their utility for desorption and ionization of peptides from membranes after blotting [42, 43]. Matrices such as ice can be used for peptides and proteins [44].

Electrospray

When a high voltage is applied to a liquid flowing through a narrow capillary, an electrical spray, composed of charged small drops ($<10\ \mu\text{m}$ in diameter), is formed. These micro-drops evaporate very fast until the number of charges on their surface becomes very high and surpasses the Rayleigh limit, then they explode forming smaller micro-drops. This process is repeated several times until ionizable analytes present in the solution escape from micro-drops (Figure 1(1)). Analytes are further desolvated in the interphase with the mass spectrometer, passing through a heated capillary or with warm nitrogen counter current, depending on type of the ESI source [45]. To prevent adduct formation with salts or detergents, samples to be ionized by ESI are previously purified by reverse phase cartridges.

One of the advantages of electrospray is that ions, depending on their molecular mass and structure, may acquire multiple charges. Peptides and proteins can be successively protonated, acquiring different m/z ratios suitable for good ion transmission in a quadrupole analyser. The multiple peak spectra of proteins are very useful for molecular mass determination in a low resolution instrument. Tryptic peptides usually become doubly or triply charged, fragmenting easily with less activation energy, giving rich information patterns for database searches and are also amenable to manual interpretation.

A dramatic improvement of sensitivity is achieved with miniaturized ESI sources. The spraying capillary opening may be reduced to several microns, even to $<1\ \mu\text{m}$ and flow rates may become as small as $10\ \text{nl}/\text{min}$ [46]. In the static form of nanospray, $1\text{--}2\ \mu\text{l}$ of the sample may last for hours, facilitating the acquisition of dozens of spectra [45]. Dynamically, coupling these miniaturized ESI sources to a micro-capillary HPLC and using columns of $<100\ \mu\text{m}$ diameter allows high sensitivity protein digest analysis, even for mixtures [47].

ESI is a very soft ionization technique as it is proven by the possibility of producing macromolecular complexes in the gas phase maintaining intact their non-covalent interactions [48, 49]. The MW of an intact virus was measured using ESI coupled to FTICR [50].

Analysers

Electric and/or magnetic fields are used for separating ions in gas phase. The analyser is the part of the mass spectrometer where this separation takes place (Figure 1a–f). Mass analysers presently used were originally designed decades ago, but some of them were not widely available commercially until recently, and now they are in widespread use thanks to the technical developments in the last two decades. Mass spectrometers may be constructed with one or more analysers, depending on the task they will be used for. Instruments composed of two or more mass analysers coupled together are known as tandem mass spectrometers.

The construction of the first mass spectrometer dates back to the beginning of the last century. For more than 60 years, sector (magnetic or electric) and quadrupole analysers have been almost the only instrument used for analytical purposes. Until the spurt in interest in proteomics during the 1990s, analysers like TOF and ion trap were in use only in highly specialized labs [51]. This review will focus on quadrupoles, TOFs and ion traps (ITs), which, alone or in tandem, are commonly used in proteomics applications. The particularities and usefulness of the most used tandem devices such as triple quadrupole (TQ), double quadrupole–time of flight (QqTOF), double quadrupole–ion trap (QTRAP) and tandem TOF (TOF/TOF), will also be discussed [52].

Time of flight. TOF (Figure 1e) is the simplest mass analyser. It consists essentially of a flight tube in high vacuum. Ions, accelerated with equal energies, fly along the tube with different velocities, which are inversely proportional to their masses [53]. A great majority of the TOF instruments uses MALDI as a source of ion formation. Given that TOF is a discontinuous separation technique, its coupling to a pulsed source such as MALDI is straightforward. With a high frequency laser, the sample throughput can be really high.

The so-formed ions are accelerated in a strong electric field (typically 20 kV), and then allowed to drift freely over the field-free region, usually

between 0.5 and 3 m long. Ions of different masses are separated since lighter ions arrive at the detector faster than those of higher mass. The TOF spectrum is a recording of the signal produced by an ion detector at the end of the flight tube upon the impact of each ion group. A conventional mass spectrum, displaying intensity over m/z ratio is achieved by taking into account the relationship of the time of arrival to the detector (t) with the square root of m/z ratio value of the ion. However, the complex MALDI process gives rise to relatively broad spatial and kinetic energy distributions, which degrade resolution. Short (100–200 ns) delay in the application of a gradient extraction voltage (delayed extraction, DE) [54] reduces peak width and increases resolution, which may be enhanced further by using an electrostatic mirror called a reflectron. The pathway for the ions in a TOF is reversed and enlarged using an electrostatic mirror to reflect ions at the end of the field-free region (Figure 1) [55]. The electrostatic mirror, allowing a deeper penetration of faster ions, compensates for small kinetic energy differences in a given ion population. Peak mass resolution in the order of 25 000 FWHM can be achieved with the present MALDI-TOF instruments. The technique is applicable not only for peptides and proteins, but also for different types of bio-molecules, such as polysaccharides, lipids and polynucleotides [56–62].

Quadrupoles. Quadrupole ions are separated due to the electrical fields created by four parallel rods, which have a hyperbolic section, although cylindrical shape is widely used. Opposite rods are electrically connected and two voltages, direct current (DC) and a radio frequency (RF) alternate current (AC) are applied. Contiguous rods have opposite DC polarity and the alternate current is 180° out of phase between them. Inside this electrically oscillating field, ions describe complex trajectories and only those with stable trajectories will travel along the quadrupole and reach the detector. When these DC and RF voltages are increased, maintaining their ratios constant, ions with increasing values of m/z ratio are selected [63, 64].

Voltages in a quadrupole may be manipulated in three different ways depending on analytical purposes. If DC is zero and only RF is applied, every ion trajectory is stable and all ions pass through the quadrupole. In this case, it is said that the quadrupole is set in RF mode. In the scanning mode,

DC and RF are varied simultaneously, allowing the ions with different m/z ratio to pass sequentially. In this instance, the duty cycle, i.e. the percentage of time in which a particular ion is getting through the quadrupole and reaching the detector, is low, usually below 0.1%; thus as most of the ions are lost and sensitivity is not high. The duty cycle may be improved at the expense of resolution; therefore sensitivity and resolution are opposed using quadrupole analysers. The third mode to use a quadrupole is to fix the voltages, allowing stable trajectories only for ions having a predetermined m/z ratio. The duty cycle now is high, close to 100%.

Ion transmission in quadrupoles decreases as m/z ratio increases and mass range is very limited when compared with TOFs. Most quadrupoles in analytical instruments have an m/z ratio limit of 4000, although the sensitivity begins to decrease well below this level. Nevertheless, due to multiple charge ion production in ESI sources, quadrupoles may be used to determine the MW of macromolecules, including proteins and nucleic acids [65].

Trap instrumentation. This group includes ion traps: both 3D (Figure 1a) and linear (Figure 1b), and ion cyclotron resonance (ICR; Figure 1d) mass spectrometers. These types of instruments have in common the particularity that ions are retained inside them and may remain trapped during the time needed to perform the usual operations in MS, i.e. full scan, precursor selection, fragmentation and product ion analysis. Furthermore, the process of ion selection, decomposition and fragment analysis can be repeated several times in a process known as MSⁿ. Successive information from fragment ions is produced by applying MSⁿ. In trapping instruments, analyser electrodes and RF voltages are designed to induce stable and close trajectories to ions, which may last entrapped from fractions of a second to hours [66].

The spherical or three-dimensional IT (QIT) [67] is, along with MALDI-TOF, the most common mass spectrometer used in proteomics. This is an affordable, compact and easy to manipulate instrument, whose popularity is based on its high sensitivity for peptide analysis when coupled to a nanospray source, as well as on its capability of producing the data quickly. ITs are formed by three electrodes (Figure 1b), one ring-shaped, named annular or ring electrode, and two end-cap electrodes at both sides of the ring. Inside the small

cavity formed by these electrodes, around 1–2 cm³, the ion trapping and analysing processes take place. Once inside the trap, increasing the ring RF voltage causes ion trajectories to destabilize successively and therefore ions are ejected from the trap. The ejected ions reach the detector sequentially and their m/z ratios are registered. Resolution in ITs is inversely related to the scanning speed of choice. Generally, both full and product scans are performed at low resolution and high scan speed, enabling the compatibility with HPLC separations but at the same time maintaining a high throughput regimen for mass analysis. Nevertheless, medium resolution, needed for ion charge assignment, may be obtained at lower scan speeds when only a short m/z range is considered (e.g. 10 Da). These short, slow scans are compatible with the high performance of the instrument.

Although the technique is based on early experiments by W. Paul in the 1950s, ion trap commercial burst began in 1996. Since then, proteomics laboratories have been replacing their TQ instrumentation with ITs for the aforementioned reason. The sensitivity is not especially good in the full scan mode due to charge–space effects because the charge repulsion phenomena can cause a limitation in the number of ions that may be trapped inside without compromising the resolution and mass exactitude of the measurement. Nevertheless, sensitivity in daughter ion scan mode is much better than on any other instrument. The peptides that are not detectable in full scan mode may be fragmented in the trap in less than a minute, producing very clear spectra useful either for database searching or for *de novo* sequencing. The capability of performing successive fragmentation steps is the basis of a successful methodology for *de novo* sequence assignment [68].

Spherical traps are normally connected to static or dynamic miniaturized ESI sources, the latter coupled to micro HPLC separation, depending on the complexity of the sample. In both cases, a highly sensitive analysis of peptide digests can be achieved as shown by hundreds of publications in the last 10 years [69]. Although commercially available, MALDI sources for ITs have not been widely used. Frequently, only poor fragmentation patterns are produced for long (15–20 amino acid residues) tryptic peptides with 1⁺ charge, and despite several publications reflecting the feasibility of this methodology, it is not routinely used in proteomic research [70].

In the recently introduced linear or 2D ITs (LITs) [71], ions are trapped in a cavity inside a quadrupole of a combination of RF voltages applied to the quadrupole rods and a DC voltage applied to the end lenses situated at both sides of the trap. Trapping efficiency is one order of magnitude higher than for 3D traps and the capacity, i.e. number of ions that may be trapped, is roughly 50 times higher. Both the magnitudes contribute to a sensible increase in instrument sensitivity.

In ion cyclotron resonance, ICR [72], ions are trapped in a cell composed of four electrodes situated in a strong magnetic field. Once trapped, the ions oscillate with a frequency (cyclotron frequency) inversely related to their m/z ratio. The cyclotron frequencies are directly related to the intensity of the magnetic field. The trapped ions may be excited by an electric RF with a frequency in resonance with their cyclotron frequency. Though the ion oscillation ratio increases, its frequency is maintained. Nondestructive detection is possible due to an image current, created in the detector, having the same frequency as the cyclotron frequency for the ion. The frequencies can be measured with great precision. As cyclotron frequency increases with the strength of the magnetic field, higher resolution and precision in mass measurement will be possible with instruments working under higher magnetic fields. When ions with different m/z ratios must be measured simultaneously, fourier transform (FT), a complex mathematical procedure is necessary. The FTICR is at present the mass spectrometer providing the highest mass resolution (values $>10^6$ have been reported [73]), enabling proteomics studies with intact undigested proteins [74]. This methodology is known as 'top-down' proteomics and will be discussed later [75].

Tandem instruments. Schemes of the tandem mass spectrometers mostly used in proteomics are shown in Figure 1. The arrangement of three quadrupoles (TQ; Figure 1c) in tandem presents very interesting scanning possibilities and has been widely used [76]. The third quadrupole can be substituted by a different analyser, e.g. a TOF (QqTOF) [77] or a linear IT (QqLIT; Figure 1c) [78].

The TQ coupled to an ESI source was the most used instrument during the early 1990s to perform peptide fragmentation. Numerous publications show the results (impressive for that time), which could be produced when nanospray coupled to TQ

was used to identify proteins by peptide fragmentation, including *de novo* sequencing [79]. The methodology used consists in the initial mass isolation of an ion by adequately setting the voltages in the first quadrupole, followed by their fragmentation in the second quadrupole, and scanning of the fragment masses in the third quadrupole. The second quadrupole is set in the RF mode and filled with gas molecules, usually argon to a pressure of around 10^{-5} torr. The ions leaving the first quadrupole collide with gas molecules and dissociate into fragment ions. These fragment ions are further mass separated in the third quadrupole (daughter ion scanning). Although one order of magnitude less sensitive than the first generation of ITs, its particular scanning modes make the TQ a very versatile instrument. As it can be seen in Figure 3, tandem MSs, which have two scanning analysers, as TQ and QqLIT, can be used as well, in the parent ion scan mode and in the neutral loss mode. In the parent ion scan mode, the first quadrupole is set to scan the full mass range, while the third is set to allow the isolation of ions with a specific m/z ratio. Using the parent scan mode, ions producing a particular fragment may be detected even in the presence of high chemical noise. The parent ion scanning for amino acid immonium ions compensates when full scan sensitivity in TQs is not good enough. The peptide ions, detected by the parent ion, are further fragmented with high sensitivity, because chemical noise is substantially removed during daughter ion scans [80].

The usefulness of neutral loss and parent ion scan modes will be discussed further, when dealing with post-translational modifications. QTRAP [78], a new type of mass spectrometer, can be used to perform parent ion and neutral loss scans with better mass resolution and much higher sensitivity than TQs and is increasingly used, especially in the study of post-translational modifications.

Instruments in which two quadrupoles are set in tandem with an orthogonal TOF (QqTOFs) [77] constitute nowadays, together with ITs, the most extended choice to analyse and fragment ESI-produced peptide ions. The high resolution (*ca.* 20 000) and sensitivity (one order of magnitude higher as compared with TQ) of QqTOFs instruments is due to the substitution of the last quadrupole in the TQ configuration by a TOF analyser. When working with the parent ion scanning mode, QqTOFs have the advantage that

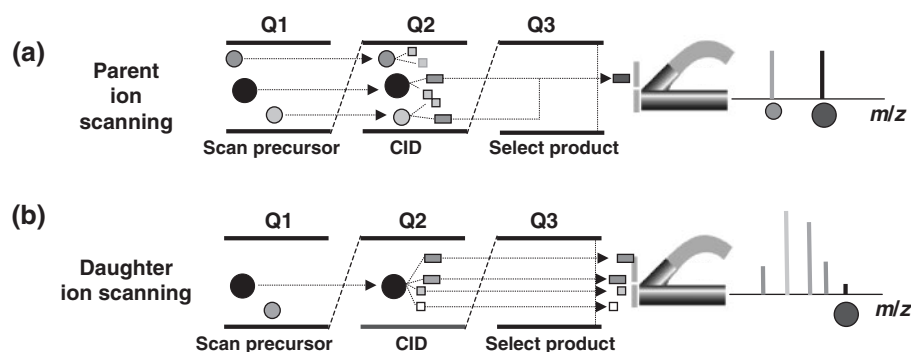


Figure 3: Combined scanning modes in a TQ. (a) Parent ion scanning: the first quadrupole is making a regular full scan, while the second one is used as collision cell fragmenting at any time the ions allowed to pass by the first quadrupole. The third quadrupole is fixed, allowing the pass of only one selected ion. (b) Daughter ion scanning: the first quadrupole is fixed for precursor selection, the second one is used as previously as collision cell, working in the RF mode. The third one is used to scan the fragments produced in the second quadrupole.

several ions may be checked simultaneously, in contrast with the single ion checking of TQs. Furthermore, higher resolution allows checking the ions with m/z ratio very close to contaminants [81]. Nevertheless, due to the poor transmission of ions between the quadrupoles and the orthogonal TOF, TQs and QqLIT are generally preferred when performing parent ion experiments [81]. A methodology recently developed facilitates one to obtain comparable sensitivity in parent ion scanning with QqTOFs, using a software that makes it unnecessary to scan their first quadrupole [82].

The coupling of a MALDI source with instruments capable of an efficient peptide isolation and fragmentation has proved to be very useful. When a protein cannot be identified by PMF, it would be very interesting to obtain peptide fragmentation data using the same sample without further purification. In this way, ambiguous PMF matches could be transformed into significant hits when enriched with the fragmentation data from at least one peptide. QqTOFs [83], TOF/TOFs [84], QITs [70] and recently LIT, have been coupled to MALDI sources. The good quality of fragmentation spectra has been extensively proven with QqTOFs and TOF/TOFs instruments (Figure 1f). MALDI-TOF/TOF constitutes today workhorses for high throughput proteomics [85]. MALDI-LIT is a promising instrument, although very new and not enough studies on its use have been published.

Recently, FTICR instruments have been coupled to quadrupole (Figure 1d) and linear IT analysers. The advantages of this configuration lie in the

possibility of performing ion selection and fragmentation outside the ICR cell [86]. High vacuum preservation and a more efficient usage of the cell are the benefits expected.

PROTEIN IDENTIFICATION

Peptide mass fingerprinting

The most frequently used strategy for studying protein expression patterns and identifying proteins involved starts with SDS-PAGE separations, either in 1D or 2D. In 2D electrophoresis, proteins are separated according to both, their isoelectric point and their molecular weight. The separated proteins are then visualized by staining with silver nitrate, Coomassie Blue or fluorescent dyes, and usually digested in-gel with specific proteases (e.g. trypsin and LysC). The resulting proteolytic peptides are extracted from the gel piece and analysed by MALDI-TOF MS [12]. The set of measured peptide masses is the peptide mass fingerprint. Proteins can be identified in this way with good high throughput compatibility [87] and a high sensitivity even below the femtomole range [88].

This experimental mass profile is matched against the theoretical masses obtained from the *in silico* digestion at the same enzyme cleavage sites of all protein amino acid sequences in the database. The proteins in the database are then ranked according to the number of peptide masses matching their sequence within a given mass error tolerance. Unfortunately, a mass fingerprint does not include

all the theoretical peptide masses expected for the given digestion conditions. Furthermore, an important percentage of experimental peaks correspond to masses that do not match the peptide masses in the protein found. These two factors may increase the difficulty of an unambiguous assignment for a protein match. Trypsin autolysis, contaminant keratin derived tryptic peptides and matrix adduct peaks can be tabulated, but in addition to these peaks most of the digest spectra contain others not easily assignable, increasing the noise in database search results. Successful protein identification depends on several factors, being the most important: (i) MALDI peaks mass accuracy (ppm allowed), (ii) the relation between assigned and unassigned peaks in the spectrum and (iii) the size of the database used [89]. With this approach, a protein is considered to be successfully identified when five or more peptide masses are matched with a mass accuracy better than 30 ppm, 15% of the sequence is covered and the next best protein candidate has a significantly lower agreement with the measured data [54]. Several scoring functions have been developed to simplify the determination of confidence levels [90].

Since 1993, the combination of PMF and MALDI-TOF has been a very sensitive, fast and reliable method for protein identification. Most of the protein identification work in the last 12 years has been done by PMF. The methodology is very simple, the techniques for sample preparation are not complicated and search programs and protein databases can be acquired and installed in the lab computers to get fast results. However, most search programs (MASCOT: <http://www.matrixscience.co.uk>, PROFOUND: http://prowl.rockefeller.edu/profound_bin/WebProFound.exe, MSFIT: <http://prospector.ucsf.edu/ucsfhtml4.0/msfit.htm>, ALDENTE <http://www.expasy.org/tools/aldente>, etc.) may be used freely through the Internet. The whole process (sample preparation, data acquisition and database search) is now fully automated and several hundred proteins may be identified per day with only one instrument [91].

Nevertheless, as no sequence information is obtained by PMF, caution must be taken with results. A high degree of expertise is needed to interpret search program reports to avoid the possibility of giving false positives. Very restrictive conditions are required today by scientific journals to accept PMF data for publication [92].

Protein identification with peptide fragmentation data

Peptide ions in gas phase may decompose under various conditions: CID, metastable decay, etc. in a process known as MS/MS. The type of peptide fragment ions depends on the decomposition process. A low energy CID regimen is most frequently used for peptide fragment-ion production. Under this regimen, ions are accelerated acquiring energies in the order of a few electron volts (eV) and fragmentation occurs principally at the peptide amide bond. Two types of fragments are produced [8]: one of them preserves the *N*-terminus of the peptide while the other conserves the *C*-terminus (*b* and *y* ions, respectively, Figure 4) [9]. Neutral losses of these ions are also frequently produced under the low collision energy regimen. The number and intensity of fragment ions is mainly peptide dependent, but diverse fragmentation patterns can be produced using different instruments. Besides, successive stages of fragmentation, MS^n , may be performed when a trapping instrument is used [68].

The identification of proteins with peptide fragmentation data may be done using one of the following approaches:

- (1) Peptide fragmentation fingerprint (PFF): uninterpreted fragmentation spectra are compared, using *ad hoc* search programs, with theoretical spectra of peptides produced by *in silico* digestion of all the proteins in a database [20, 89].
- (2) Peptide sequence tag: peptide mass together with a short sequence produced by partial interpretation of the spectrum is used for database searching [22].

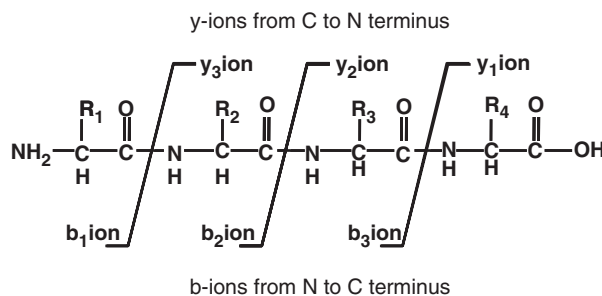


Figure 4: Principal ion series generated upon low energy CID of peptides. Cleavage at the peptide bonds produces two ions containing either the *C*-terminus (*y* ions) or the *N*-terminus (*b* ions).

The identification of peptides by MS/MS has substituted almost completely the more time consuming and relatively insensitive method of Edman degradation. Uninterpreted peptide CID fragmentation spectra are used to identify peptides and proteins with the help of programs like MASCOT [89] or SEQUEST [93], which use mass fragmentation data to explore either protein databases (SwissProt, NCBIInr, ...), or nucleotide data, as the incomplete nucleotide sequences contained in the diverse EST databases [94]. Each of these programs uses its own scoring system showing the probability of finding in a database a peptide identical to that producing a particular fragmentation spectrum. Depending on data quality and the instrumentation used, different search programs may have better performance and may be the obvious choice. Clear-cut results are produced in many cases with their own scoring systems but, frequently, an expert is needed to take an eye on the spectrum to be sure about the results. This is not compatible with high-throughput peptide and protein identification.

Peptide fragmentation spectra obtained by MS/MS in a mass spectrometer are automatically assigned using search programs, provided that proteins that originate these peptides are included, either in a protein or in a genomic database. Spectra from peptides, unmodified or with known modification, constitute the ideal conditions for peptide identification and facilitate retrieving results with high scores and no doubt about their authenticity. Nevertheless, depending on the peptide amount and on averaging time, poor quality spectra may be produced, precluding their automated identification. Peptides may fragment too in a non-homogeneous way giving poor sequence coverage. This effect is pronounced in spectra taken from non-tryptic peptides when internal arginine or lysine residues are present. Hence, with some frequency, the quality of the fragmentation spectrum makes the identification of a peptide very difficult and there is not enough discrimination for a positive assignment. This is particularly true when low sensitivity analysis is performed using LC/MS/MS or when peptide fragmentation is not homogeneous. In this case, MSⁿ experiments may help in the production of the expected results. Under these circumstances, the search program report consists of a list of peptides not containing an evident candidate and the right answer might be included or not in this list [95]. Manual interpretation of peptide fragmentation

spectra, besides requiring particular expertise, is very time consuming and is absolutely incompatible with high throughput peptide and protein identification. Hence, most of the labs involved in protein identification take the safest choice, discarding any search program result which is not clear-cut.

The ESI has been, generally, the choice to ionize peptides for fragmentation. It may be coupled easily to quadrupoles or ITs, which perform efficiently in ion selection and CID. Furthermore, peptide ions formed by ESI are mostly doubly or triply charged, making fragmentation easier. The miniaturized ESI sources as static nanospray or microspray coupled to HPLC separations has been the choice during the last 10 years to produce peptide ions for fragmentation. Electrospray coupled to different tandem mass spectrometers constitutes a very robust methodology and a great deal of proteins have been identified in such way, as it is demonstrated by hundreds of biological publications [25]. However, during the past few years, MALDI sources have been coupled to existing (IT [70], QTOF [83]) or new (TOF/TOF [85]) mass spectrometers. Efficient isolation and collision dissociation of ions may be performed using these instruments. As MALDI sample preparation is less demanding, there is an increasing trend towards using this source to ionize peptides for fragmentation experiments.

MALDI-TOF/TOF mass spectrometers (Figure 1(f)) enable protein identification by combining the high throughput of the PMF method with the specificity of the PFF method, since several MS/MS spectra can be acquired in a few seconds from a given sample and fragment ion data can be combined with precursor ion data for a highly reliable database search. In addition, the efficiency of MALDI-TOF/TOF instruments enable the study of proteomes from organisms with unknown genomes either by matching at least one conserved peptide sequence in the protein database or by *de novo* peptide sequencing from high quality MS/MS spectra. Insofar as many biological projects will continue to use 2D-PAGE, this new generation of MALDI instrumentation will strengthen further the role of MALDI MS in such research areas as protein identification and quantization, protein profiling, protein-protein interaction studies and PTM characterization, among others [25].

Unfortunately, the PFF approach is only useful when there is an exact coincidence between the

experimental data and a sequence included in the database. Any mass difference due to unexpected modifications precludes the identification. Conversely, the sequence tag approach, although requiring investing time for partial interpretation of the spectrum can be more tolerant with differences in peptide masses and sequences, allowing successful database searches when modifications occur.

***De novo* sequencing**

When protein identification remains elusive using the above described approaches, the interpretation of peptide fragmentation spectra is the last possibility to obtain information about the protein under study. With a deep knowledge of peptide decomposition mechanisms, frequently it is possible to reconstitute the sequence of a peptide from its fragmentation spectrum [96]. This process is known as *de novo* peptide sequencing, as it is independent of any information present in databases. Traditionally, *de novo* sequencing was the task for Edman sequencing, but as stated earlier this technique is very slow and insensitive by today's standards. However, the number of labs that have the expertise to attempt the interpretation of MS peptide spectra successfully has dramatically increased. The data produced are generally used for searching protein homologies using sequence analysis by BLAST, or directly to design primers in order to clone the specific gene.

Low-energy collision fragmentation MS/MS spectra of electrospray produced peptide ions have been extensively used for *de novo* peptide sequencing. Tandem instruments like TQ/IT or QTOF have been used successfully [97]. The MALDI is used with increasing frequency as a source of peptide ions for *de novo* sequencing. Apart from using ESI ions, good results have been presented with MALDI coupled to ITs, QqTOF and TOF/TOF [98] instruments, either using or not using diverse chemistries for peptide modification and simplification of the spectra. In this sense, the success in sequence interpretation using PSD in a MALDI-TOF mass spectrometer when peptides are N-terminal modified with a sulfonic acid moiety must be quoted [99].

Most of the commercialized instrumentation may be used for *de novo* sequencing, and the choice is dependent on availability and personal preferences. For example, some labs feel more comfortable with the resolution and exact masses provided by QqTOF instruments, while others prefer to sacrifice

resolution for the possibility of performing MS [3] and will choose ITs.

Interpretation of fragmentation spectra is largely dependent on the individual peptide sequence, causing the presence of internal peptide basic amino acids uneven fragmentation and major problems with sequence assignments. To offset this problem, tryptic digestion, which locates the two most basic amino acid residues, arginine and lysine at the C-terminus of the produced peptides, has been employed. With these basic residues fixed at the C-terminus, improved continuity of *b* and *y* ions (peptide bond fission fragment ions) is observed, making interpretation of spectra much easier. Differentiation of *b* from *y* ions is frequently obvious, but in some cases additional experiments would be needed, such as comparing unmodified spectra with a chemically modified spectra, i.e. esterified or N-acetylated. The tryptic digestion of the protein in a mixture of $^{18}\text{O}/^{16}\text{O}$ water, resulting in a duplicate set of *y* ions differing in mass by 2 Da [100], or N-terminal tagging of a peptide to define the *b* ion series, are different alternatives.

The possibility of completely interpreting peptide fragmentation spectra using computer programs is an old dream. Since 1997, Rich Johnson's *de novo* sequence program Lutefisk has been available on the Internet and recently most of the instrument manufacturers commercialize their own programs. Generally speaking, spectra taken at high resolution are preferable for computer interpretation. Although very helpful, neither of them can be considered as absolutely trustable and are completely useless without checking. However, they can save many hours of work when an expert faces the interpretation of a batch of spectra.

POST-TRANSLATIONAL MODIFICATIONS

Post-translational modifications (PTMs) are defined as the series of chemical reactions whereby a newly synthesized polypeptide chain is converted to a functional protein. A common characteristic of PTMs is that the accompanying change in amino acid structure produces the corresponding change in the formula weight of that amino acid relative to the original, unmodified residue. This mass change is usually the basis of the detection and characterization of PTMs by MS. The activity of most eukaryote and prokaryote proteins is modulated by PTMs [101].

The characterization of these modifications plays an important role in understanding most biological processes, such as activation/inactivation of enzyme activity [102–104], regulation of gene expression [105, 106], cellular localization [107], modulation in ligand–receptor interactions [108, 109], or destruction tag signalling [110]. Despite their pivotal role in biological function, the study of PTMs has been hampered by the lack of suitable analysis methods and many key modifications are yet to be discovered in a wide variety of biological processes.

The PTMs can be classified into three categories: proteolytic cleavage of part of the sequence (removal of an initiator methionine, a signal sequence, a transit peptide, etc.), adjunction of a chemical group (acetylation, glycosylation, phosphorylation, etc.) and formation of inter- or intra-peptidic linkages (disulfide bonds, thioether links, etc.). MS has a particular strength in its capability of characterizing PTMs and the use of new analysis technologies, e.g. the combination of different scan modes for ion filtering and precursor selection in hybrid systems (Figure 3), makes the study more amenable. The effectiveness of mass spectrometric analyses for characterizing PTMs lies on the ability to expand the list of amino acid sequences to include additional modified residue masses. The most significant issue that affects and limits the analysis of PTMs is the specific nature of modification in terms of the precise modification site, and the generally low stoichiometry of most modification reactions. As a result, the difficulty of characterizing a particular PTM of a protein is recognizing the modified peptide as an ion in the extensive set of mass spectrometric data yielded by the analysis of a protein digest. Electrospray is one of the softest ionization techniques for MS [11], which allows the detection and characterization of many PTMs, since direct sequence information is obtained. Highly complex protein mixtures can be directly identified and the possible post-translational modifications characterized by coupling 2D-nLC with tandem IT MS [26], using an electrospray source as external ion generator.

Phosphorylation

One of the most studied regulatory PTM is the phosphorylation in serine, threonine and tyrosine residues, which occurs through the action of protein kinases and can be reversed by the action of protein

phosphatases. Phosphorylation regulates many biochemical reactions involved, among other processes, in intracellular signalling [111–114], extracellular tissue development [115], response to stress [116], or in pathogenic processes [117], and is a common modification of proteins. Therefore, great effort has been directed towards developing methods for detecting and characterizing this modification. However, identification of phosphorylated residues is a difficult task. When phosphorylation occurs in tyrosine, collisionally induced dissociation of the modified peptide produces the corresponding fragment showing the increased mass of the phosphate group in the tyrosine residue. Then phosphotyrosine has a molecular weight of 243 Da. This behaviour is markedly different from the collisionally induced dissociation of phosphoserine and phosphothreonine peptides. Two aspects of the product ion spectra are noteworthy: first, the neutral loss of H_3PO_4 is very probable, which results in an ion 98 Da smaller from the parent ion. Second, both the *b* ions and *y* ions in the fragmentation spectrum that contain the phosphoserine or phosphothreonine are produced by consecutive fragmentation reactions breaking the amide bond and resulting in the loss of H_3PO_4 , or *vice versa*. Many laboratories used these fragmentation reactions to facilitate the recognition of serine and threonine phosphorylated peptide ions. In order to detect and characterize phosphorylated sites from a complex peptide mixture, usually it is necessary to enrich the mixture with phosphopeptides by immobilized metal-affinity chromatography [118] (IMAC) and using nano-HPLC coupled on line to nanospray ionization MS for the analysis. Using this approach, it is possible to characterize hundreds of phosphopeptides from a whole-cell lysate in a single experiment [119] and this methodology can easily be extended to display and quantify differential expression of phosphoproteins in two different systems. Some groups employed other approaches to identify phosphorylation sites in proteins from complex mixtures. One of these approaches is based on the replacement of phosphate groups of serines and threonines by ethanedithiol [120]. The resulting derivative peptides are combined with a biotin affinity tag in order to separate phosphopeptides from the non-phosphorylated tryptic peptides, and finally characterized by MS. Other approaches include a six-step derivatization/purification protocol that requires a much longer time to complete and attain the phosphorylated protein population [121].

In the recent years, new analytical methods based on MS have been developed for characterizing PTMs. The combination of different scan modes for ion filtering in hybrid systems (Figure 3) allows efficient selection of precursor ions, which release a specific fragment ion when induced to fragment in a LIT mass analyser. The examples are the selection of ions losing the phosphate group (-79 Da) when ionized in negative mode (proper ionization mode for ionization of phosphopeptides) and the selection of ions that have the enhanced neutral loss of phosphate in positive mode (-98 Da in singly charged parent ions) within a complex mixture of phosphorylated and non-phosphorylated peptides. The combination of classical TQ scan modes (precursor ion scanning or multiple reaction monitoring), with LIT scan modes allows the possibility of searching, isolating and characterizing phosphorylated peptides (Figure 3). A typical experiment aimed at phosphopeptide searching using hybrid instruments takes advantage of the rapid ionization polarity switching and combines precursor ion filtering in a TQ for parent ion selection with high-resolution scan for mass labelling and charge determination and enhanced product ion scan to induce fragmentation.

Ubiquitination

Ubiquitination is a PTM in which ubiquitin chains or single ubiquitin molecules are appended to target proteins, giving rise to poly- or monoubiquitination, respectively [122, 123]. Polyubiquitination targets proteins for destruction by the proteasome. The strategy for identifying the precise ubiquitination site by MS relies on searching the diglycine remnant of ubiquitin covalently bound to a lysine residue. Recently, a proteomic approach to enrich, recover and identify ubiquitin conjugates from yeast lysate [110] has been developed.

Xenobiotic modifications

Another most studied type of PTM is modification by xenobiotic species such as toxins or drugs. The general goal in the studies of xenobiotic protein modification is to help to elucidate the mechanisms of action of these agents by the analysis of the corresponding adducts. The careful design of chemical probes to be covalently attached to proteins of interest will strengthen our understanding of biology of drug development, and proteomic

approaches represent an essential analytical tool in this process. The chemical probes that have been developed to date have been shown to be useful for the identification of cysteine proteases involved in processes such as apoptosis [124], cataract formation [125] and malarial infection [126]. Other probes have been used to profile enzymes involved in clinically relevant conditions such as cancer progression or cancer invasiveness [127, 128]. In addition to the increasing number of reported chemical probes, many mechanism-based inhibitors or affinity labelling reagents have already been identified. Among the major classes of enzymes that have been the focus of drug discovery efforts, the two that are most amenable to be analysed by a proteomic approach are protein kinases [129] and protein phosphatases [130, 131]. The application of a proteomic approach may help elucidate the biological effects of the covalent binding between drug and protein, and may allow the precise characterization of the modification sites along the protein sequence. A recent research field that has been termed 'clinical proteomics' includes separating, identifying and characterizing proteins, defining and constructing clinical databases and sample collections, and testing clinically relevant hypotheses against complex data sets derived from high-throughput measurement techniques and well-phenotyped human populations. The resulting knowledge of the molecular processes and mechanisms underlying disease will lead to novel therapeutic strategies. Advances in MS have enabled the simultaneous quantification of large numbers of different proteins [25] at minute levels. Changes in the quantity and structure of proteins are associated with different disease states. Instead of searching for specific individual biomarkers to identify the presence of particular disorders [132, 133], proteomics enables massive identification of differential proteins by matching mass profiles of altered and control samples.

Acetylation

Protection against proteolysis mediated by aminopeptidases has been suggested as one of the common functions for acetylation ($+42$ Da), a modification present in most of the eukaryotic soluble proteins in mammalian cells [134]. *N*-terminal serine, alanine and methionine residues are normally the targets for acetylation. In recent years, several mass spectrometric approaches have been devised for selective identification of this PTM and sequencing of the

modified peptides [135]. These approaches have revealed an outstanding role for acetylation in nuclear receptors [136], human cancer [105] and plant histones [137], among many others.

Methylation

Several methylases have been described that specifically modify individual proteins. Mono-, di- and trimethylation (+14, +28 and +42 Da, respectively) take place at the amino groups of alanines and methionines and at side-chain lysines and are believed to play a role in the regulation of gene expression, protein localization, protein–protein interactions and protein–nucleic acid interactions [105, 134–143]. MS has been extensively used to demonstrate methylation of protein complexes [144] and of the *N*-terminal domains of histones [145–147], among many others.

Glycosylation

An enormous variety of glycosylations occur in proteins that influence their functionality in many ways. The formation of branching carbohydrate trees provides a shield for the tertiary fold of proteins or their functional domains. Some glycoproteins have more than 40% of carbohydrate content (e.g. erythropoietin). These covalent modifications are known to play a key role in many diseases by influencing regulatory and developmental or transport processes [134]. Normally, serine and threonine are the targets for *O*-glycosylation, while *N*-glycosylation takes place via asparagine residues. The carbohydrate chains are commonly composed of galactose, fucose, *N*-acetylglucosamine, acetylgalactosamine, mannose, sialic acid, *N*-acetylneuraminic acid and *s*-inositol. In some cases (e.g. yeast extracellular glycoproteins), oligosaccharides with hundreds of sugar units are attached. The disialylated biantennary and tetraantennary *N*-glycans may contribute notably to the 3D surface of glycoproteins. In these cases, the polypeptides may serve as a support for glycan presentation in glycoprotein/carbohydrate binding interactions.

The *N*- and *O*-glycans can be removed from glycoproteins upon treatment with *N*-glycosidase *F* and *O*-glycosidase, respectively. The comparative mass spectrometric analysis of treated and untreated glycosylated samples can reveal the type of modification (e.g. *N*-acetylhexoseamination entails a mass increase of +203, +406, +609 Da), and target

residues can be inferred by MS/MS analysis. A huge number of works [147–150] have demonstrated the ability of mass spectrometric approaches to reveal protein glycosylation sites, especially in combination with such techniques as isotope tagging [151, 152].

Nitrosylation

Nitric oxide (NO) is believed to regulate crucial processes within the cell through interaction with and modification of a variety of molecules, including proteins. Due to the lability of the S–NO bond, S-nitrosylated (+29 Da) protein species have been described mainly through indirect approaches such as the ‘biotin switch’, in which the cysteine-bound NO is replaced by a biotin tag [153]. This approach has enabled the identification of S-nitrosylated proteins in brain tissue [153], mesangial cells [154] and endothelial cells [155, 156].

SECOND GENERATION PROTEOMICS

Proteomic technologies have emerged as a new paradigm for biological process understanding, drug discovery and development of new biotechnology products. Due to this challenge, a great effort in different fields such as bio-analytical research, peptide/protein chemistry and bioinformatics have been made. The development of new analytical and data management tools to perform shotgun analysis will allow the scientific community to acquire a better understanding of biological processes, boosting the development of a new global approach to study life: Systems Biology.

In the last five years, proteomic procedures are changing dramatically. In 1999, in John Yates’ laboratory a new methodology to perform rapid simultaneous identification of hundreds of proteins was presented [157]. Two years later, this methodology was further refined and termed as multidimensional protein identification technology (MudPit) [26], and constitutes the basis of gel-free proteomics. Now, complex protein mixtures are digested in-solution, and the peptides produced are separated using chromatographic techniques based in orthogonal separation steps using strong cation exchange (SCX) and reverse-phase (RP) chromatography coupled to ESI–MS/MS. These analyses can be made in two different ways: (i) in the so called online method, both chromatographic separations are physically

interfaced so that the strong cation exchange separation is performed using salt steps. In each salt step, eluted peptides are trapped by the reverse phase column, washed and then analysed in a RP-LC-MS/MS run. The successive cycles of this process are performed; (ii) In the offline approach, peptides are separated in the first dimension using a salt gradient and eluted peptides are collected in fractions. The fractions are then dried down, reconstituted in an aqueous solvent and each one is analysed in a RP-LC-MS/MS run [158]. Both methodologies have advantages and drawbacks. The online approach is completely automatable but the offline method provides better coverage because of the better peptide separation in the first dimension. Peptide identification is carried out using software for database search with uninterpreted tandem mass spectra as mentioned before. Due to limitations in the scoring systems of current search engines, not all fragmented peptides can be confidently identified [159–161]. The development of several statistical methods for unambiguous peptide identification is highly necessary. Each database search engine has its own rules for matching the experimental tandem mass spectra of the peptides to the predicted mass spectra of the amino acid sequence contained in the database. While the scores assigned by these programs have been widely used for peptide identification, they do not readily provide an automated way to interpret the results, mainly because it is not possible to clearly distinguish between correct peptide assignments and false identifications (Figure 5). When dealing with small data sets, manual curation is a common routine. However, huge datasets generated by second generation proteomics experiments typically contain thousands of spectra, making manual inspection unfeasible. Several filtering criteria, based on the scores provided by database search engines and the properties of the identified peptides, have recently been developed to try to separate correct from incorrect peptide assignments. The statistical behaviour of the scores provided by algorithms such as SEQUEST or MASCOT has been studied using random or inverted databases in order to turn the raw scores into random match probabilities. The goal is to obtain the highest possible rate of successful identifications with the lowest possible rate of false positives [95, 159–161].

Although still under development, shotgun protein identification is increasingly used and now

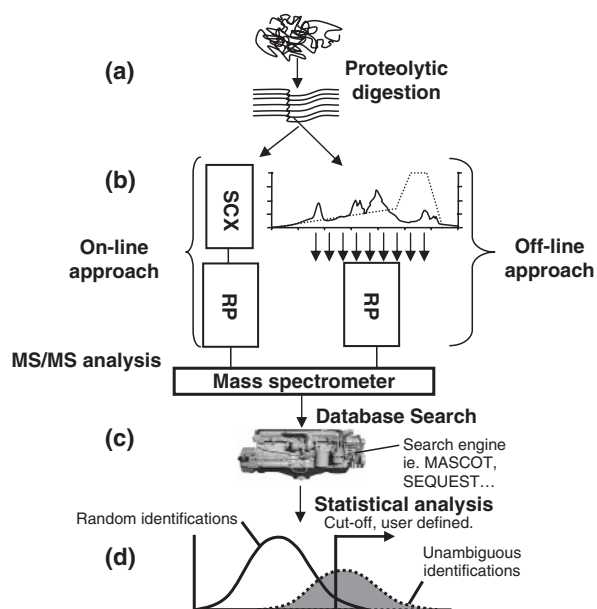


Figure 5: Large scale proteome characterization using second generation proteomics. (a) proteolytic digestion of whole extracts, (b) fractionation of the peptide pool using orthogonal chromatographic separations and MS analysis, (c) database search of the MS/MS spectra and (d) statistical analysis of the obtained results.

it is not only circumscribed to research laboratories involved in technical development. It is highly sensitive, can be completely automated, and allows a real high throughput in proteome analysis. This analytical strategy is being extensively used in an exploratory fashion for the 'blind' characterization of all kinds of biological samples, providing a comprehensive view on the studied system at the protein level, regardless of its complexity, thus laying the way to further hypothesis-based studies or experimental designs. The number of published articles is increasing each day and, as a representative example, it is worth mentioning the recent great success in the integration of these techniques with genomic and transcriptomic approaches in the study of the protozoan organisms belonging to the genus *Plasmodium*, which cause malaria and related diseases, and have a quite complex life cycle [162–166].

Recently, the efficient offline coupling of RP liquid chromatography with MALDI instruments has demonstrated its feasibility [167]. The outlet of the RP column is interfaced to a spotting device, which deposits a mixture of column eluent and matrix directly into the MALDI target. Despite the technique being not fully standardized

yet, a number of works have demonstrated its potential use for the detailed analysis of complex systems [168–170]. Its principal strength is the possibility of reanalysing any desired spots after an initial data processing.

QUANTIFICATION

Often one of the important aspects in proteomic studies is the need to measure the relative or absolute amounts of the proteins present in a biological sample. This is essential for studying the effect of a given agent on a biological system, for comparison of two different biological states, evaluation of yields in biotech process, etc. Quantitative or comparative proteome analysis was initially performed with 2D-PAGE [171, 172]. Although 2D-PAGE is still the most common approach, its limitations are well known: time consuming, labour intensive, low dynamic range and insufficient coverage of very large (>150 kDa), small (<15 kDa), hydrophobic, or otherwise insoluble proteins. In addition, usually more than 100 µg of total protein is necessary for each gel analysis.

For these reasons, more versatile MS-based approaches in conjunction with gel-free protein/peptide separations have been developed in recent years as an alternative to the classical gel-based technology. Today, stable isotopic labelling is the widest used methodology for quantization, either relative or absolute. The advantage of isotopic based methods relies on their efficiency when MS is applied. Several stable, non-radioactive isotopes like ^2H , ^{13}C , ^{15}N and ^{18}O have been used for this purpose. Three strategies have been applied to introduce stable isotopes into peptide and protein

chains to perform relative quantization (Figure 6): (i) the label can be introduced by chemical reaction of specific amino acid residues or functional groups with an isotopically labelled reagent, as with the isotope code affinity tags (ICAT) [173] or iTRAQ strategies [174]; (ii) enzyme catalysed reactions like $^{16}\text{O}/^{18}\text{O}$ exchange catalysed by trypsin or other proteases [175, 176] is a second and very promising approach; (iii) isotopically enriched media can be used to culture the cells to be studied [177–179]. All these procedures have been extensively evaluated and used during the last years, but none of them has proven to be superior, each one having advantages and disadvantages. In addition, absolute quantification strategies relying on the use of internal standards, such as synthetic peptides labelled with isotopically modified amino acids [180, 181] are being developed for very accurate quantitative measurements of specific *a priori* chosen targets.

This part of the review will focus on the most widely used, MS-based, non-gel approaches for quantitative proteomics, such as ICAT, metabolic labelling and $^{16}\text{O}/^{18}\text{O}$ exchange, together with recent applications of each one of these methodologies. Promising new developments, like iTRAQ and label-free LC-MS, will be commented upon as well.

Chemical labelling, ICAT and iTRAQ technologies

Although a wide variety of isotopically labelled chemicals have been reported, the ICAT procedure is the most widely used approach. It consists of two isotopically labelled molecules, which are differentiated only by their mass due to the substitution of several light isotopes (light reagent) by their heavy counterparts (heavy reagent). The reagent is

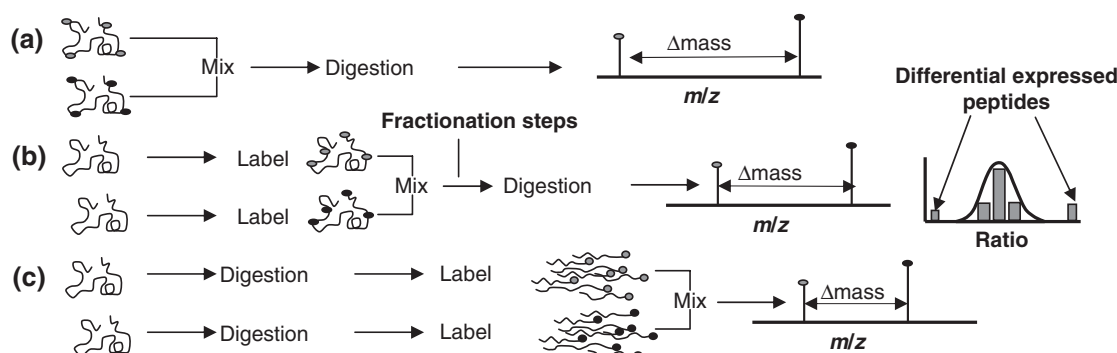


Figure 6: Strategy for quantitation differential expressed proteins in proteomics using MS. The three main strategies for introducing isotopic labels in proteins or peptides, (a) metabolic labelling, (b) chemical labelling and (c) enzymatic labelling.

composed of two different functionalities originally held together by a differentially deuterium-labelled linker [173]. A thiol specific reactive group was linked to a biotin moiety, allowing avidin purification of cysteine-containing peptides. In the linker structure, the substitution of eight hydrogen atoms by deuterium allowed the mass differentiation. The ICAT-labelled peptides could be separated from the rest of the peptides using affinity chromatography with avidin micro columns. In a further step, the elution of peptides from the avidin column produced labelled peptides which could be identified and quantified by MS. Nevertheless, some slight differences in chromatographic properties of the deuterated and the undeuterated forms of ICAT were observed. In addition, elution from the affinity column was incomplete in many cases. To improve its performance, the original ICAT reagent has suffered two modifications [182]. In the new reagent, deuterium is avoided and for differential labelling nine ^{12}C are substituted by nine ^{13}C in the linker. Substitution of ^{12}C by ^{13}C does not affect chromatographic behaviour. The addition of a cleavable bond to the molecule, allowed biotin to be removed, increasing recovery. Peptides purified with biotin moiety gave better fragmentation patterns, yielding more unambiguous identifications. Since only cysteine-containing peptides are labelled and purified, the sample is greatly simplified, enabling the quantitation of lower abundance proteins. Thus, this strategy is suitable for analysis of very complex mixtures, but not universally applicable to any tryptic digest. The use of this quantitative approach is not limited to any specific LC-MS platform, since a variety of sources and mass analysers have successfully been employed.

This technology has been the most used quantitative approach in the last few years. Many different works have been published, either 'blindly' applied to highly complex proteomes [183–190] or to specific macromolecular complexes [191, 192].

In parallel to the development of new ICAT reagents, several groups have used isotopically labelled forms of chemicals commonly used in protein chemistry to generate signal doublets. In contrast to these methods, it is worth mentioning a promising and completely different technique, called iTRAQ [174], in which the quantitative measurements are taken from tandem mass spectra rather than from survey scans/chromatographic peak areas. iTRAQ reagents consist of a set of

N-terminal-reacting isobaric compounds that, upon CID, release a mass-coded reporter ion. The combinations of heavy isotopes are introduced in the labelling group so that the precursor ion mass is the same for all the reagents, but each of them shows a reporter ion of different mass. The main advantage of this approach is the ability to simultaneously analyse as many samples as the number of iTRAQ reagents commercially available (multiplexed analysis, for now four reagents). The relative intensities of the reporter ions provide direct readouts of the relative contribution of each of the samples in the mixture.

Metabolic labelling

Metabolic labelling should be the most accurate methodology for quantification, because neither chemical nor fractionation steps are required to introduce the label and technical variation is thus minimized. Two main methodologies have been described. The first one consists of the introduction of ^{15}N -enriched ammonium sulphate or amino acids in the culture medium, where they are uptaken by cultured cells and incorporated to their endogenous proteins [177, 178, 193, 194]. A second one consists in the introduction of leucine or arginine, labelled with either deuterium or ^{13}C [179, 195]. In either method all the sample preparation steps are done after mixing the two differentially labelled samples, so any experimental treatment or artefact affects both samples equally. Peptide identification and quantitation can be done using MALDI or ESI coupled to any mass spectrometer. The separation between ^{15}N -labelled and unlabelled mass peaks is unpredictable because the incorporation yield of ^{15}N is not known. In the second strategy, the mass difference between the peptide pairs is not fixed, either, but it is predictable from the peptide sequence.

It is obvious that these strategies can only be applied to cultured cells, but they have been reported to be very powerful in elucidating elemental biological processes [196–202]. Recently, these strategies have also been successfully used for labelling rather complex multicellular organisms such as *Caenorhabditis elegans*, *Drosophila melanogaster* [203], plants [204] or mammals [205]. Furthermore, pairwise comparisons between cultured cell lines and dissected tissues are also highly promising [206].

Quantitative proteomics using enzymatic $^{16}\text{O}/^{18}\text{O}$ labelling

This methodology is based on the trypsin catalysed oxygen exchange at the free C-terminus from a peptide [175, 176]. When a peptide is incubated with H_2^{18}O in the presence of trypsin, there is a water molecule exchange resulting in the addition of one or two ^{18}O atoms to the carboxy-terminus of the peptide. Oxygen exchange is sequence-independent for tryptic peptides and there are almost no experimental constraints in the sample preparation protocol. The proteins are digested with trypsin using ^{16}O water. Further incubation of the digest with trypsin either with H_2^{16}O or H_2^{18}O water completes the differential labelling. Besides trypsin, other proteases have also been reported to perform well for the oxygen exchange [176]. The great advantage of this method is its great versatility: virtually any protease-generated peptide can be labelled and it can also be successfully applied to very low amounts of sample [207]. For the mass spectrometric analysis, high resolution scans are preferred because of the narrow spacing and partial overlap of the labelled and unlabelled species. The number of published works is, for now, scarce [207–211].

Label-free method

Quantitative proteomic analysis using label-free methods consists of looking for statistical differences in ion signal intensities or chromatographic peak areas among between several HPLC–MS runs of the samples to be compared [212–215]. These methods are based on the linear relationship between the signal and the concentration of the analyte in the chromatographic eluent, but they have been poorly characterized. It is expected that, in many cases, due to problems derived intrinsically from mass spectrometric analysis such as ion suppression effects or the influence of differential data-dependent scanning conditions on peak intensity, the accuracy of the measurements can be reduced if the sample is very complex. Nevertheless, the continuous and fast development of MS techniques coupled with ultra-high efficiency separation methods promises reliable results in a short time. In this sense, high resolution FTICR MS combined with accurately controlled chromatographic conditions, the so-called ‘accurate mass and time’ approach [216], holds great promise in quantitation power, resorting to

MS/MS identification only for those peptides of interest which show significant change in signal intensity.

TOP-DOWN PROTEOMICS

All the paragraphs above review ‘bottom up’ proteomics, which refers to those experiments analysing protein digests, and consequently involve working at peptide level. However, these strategies do not allow the study of combinations of site-specific mutations and PTMs in full, intact proteins. The so-called ‘top-down’ proteomics is an exciting field that promises to offer new information taking advantage of the exact molecular weight and direct fragmentation of the protein ions in the gas phase. Due to the importance of taking very accurate mass measurements, FTICR analysers are preferred, but others can also be used [217]. Fragmentation of intact isolated proteins in high-resolution mass spectrometers enables extensive structural characterization, leading to accurate determination of modifications, errors or derivatized active sites, or sequencing of the whole protein in a single experiment [218, 219]. The main disadvantages of this new methodology are its lower efficiency in protein identification, sensitivity and throughput compared with bottom-up strategies. Hence, its application in solving biological problems is not currently widespread because of the still insufficient technical development (both in hardware and software). Perhaps, its main limitations are the lack of an efficient protein separation technology that can be easily interfaced to the mass spectrometer, the signal suppression effects observed when electrospray ionization is used and the still not well-controlled fragmentation mechanisms inside the ICR cell [220]. Nevertheless, whenever these strategies have been applied as an integrated technology together with the bottom-up approaches, the assignment of PTM sites is simplified; confidence of identifications increased, and protein isoforms differentiated [221].

CONCLUSIONS

As it can be seen from the reviewed literature, any mass spectrometer having tandem mass analysis capabilities, regardless the ion source, mass analyser or separation technology to which it is interfaced,

can be used for both protein identification and quantitative studies. Specialized instruments with increased resolution and scan speed are providing and will continue to provide enhanced analytical capabilities, showing better performance in certain types of analyses. Nevertheless, most proteomics applications reviewed are amenable to being implemented in a wide variety of platforms. The fast developments of mass spectrometers in parallel to the advances in protein chemistry provide us with new ways for asking biological questions. It is likely that gel-free MS-based proteomics will occupy a central role in biological research, providing accurate measurements of changes in biological functions, which will be elegantly coded by researchers as changes in molecular mass. In order to accomplish these goals, new interesting configurations of ion sources and mass analysers will be explored, new scan modes will be designed, and physical improvements of the current spectrometers will provide enhanced 'firepower'.

Key Points

- Development of soft ionization techniques allowed the study of biomolecules by mass spectrometry.
- Mass spectrometry provides the most accurate mass measurements of large biomolecules.
- PMF allows rapid protein identification with high sensitivity.
- Protein sequence databases and database search engines play a central role in peptide and protein identification.
- Mass spectrometry is highly suitable for the study of post-translational modifications.

References

1. Borman S, Russell H, Siuzdak G. A mass spectrometry timeline. *Today's Chemist* 2003;**13**:47–9.
2. Matsudaira PT. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. *J Biol Chem* 1987;**262**:10035–8.
3. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 1970;**227**:680–5.
4. Barber M, Bordoli RS, Sedgewick RD, et al. Fast Atom Bombardment of solids (FAB). A new ion source for mass spectrometry. *J Chemistry Soc – Chemical Comm* 1981;**11**:325–7.
5. Mcfarlane RD, Torgerson DF. Californium-252 plasma desorption mass spectroscopy. *Science* 1976;**191**:920.
6. Blakley CR, Vestal ML. Thermospray interface for liquid chromatography/mass spectrometry. *Anal Chem* 1983;**55**:750.
7. Jennings KR. Collision-induced decomposition of aromatic molecular ions. *Int J Mass Spectrom Ion Phys* 1968;**1**:227–35.
8. Biemann K. Mass spectrometry of peptides and proteins. *Annu Rev Biochem* 1992;**61**:977–1010.
9. Roepstorff P, Fohlman J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed Mass Spectrom* 1984;**11**:601.
10. Hunt DF, Yates JR 3rd, Shabanowitz J, et al. Protein sequencing by tandem mass spectrometry. *Proc Natl Acad Sci USA* 1986;**83**:6233–7.
11. Fenn JB, Mann M, Meng CK, et al. Electrospray ionisation for mass spectrometry of large biomolecules. *Science* 1989;**246**:64–71.
12. Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular mass exceeding 10 000 Daltons. *Anal Chem* 1988;**60**:2299–301.
13. Henzel WJ, Stults JT, Watanabe C. *Proceedings of the Third Symposium of the Protein Society*. Seattle, WA, 1989.
14. Henzel WJ, Watanabe C, Stults JT. Protein identification: the origins of peptide mass fingerprinting. *J Am Soc Mass Spectrom* 2003;**14**:931–42.
15. Pappin DJ, Hojrup P, Bleasby AJ. Rapid identification of proteins by peptide-mass fingerprinting. *Curr Biol* 1993;**3**:327–32.
16. James P, Quadroni M, Carafoli E, et al. Protein identification by mass profile fingerprinting. *Biochem Biophys Res Commun* 1993;**195**:58–64.
17. Henzel WJ, Billeci TM, Stults JT, et al. Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. *Proc Natl Acad Sci USA* 1993;**90**:5011–5.
18. Mann M, Hojrup P, Roepstorff P. Use of mass spectrometric molecular weight information to identify proteins in sequence databases. *Biol Mass Spectrom* 1993;**22**:338–45.
19. Yates JR 3rd, Speicher S, Griffin PR, et al. Peptide mass maps: a highly informative approach to protein identification. *Anal Biochem* 1993;**214**:397–408.
20. Eng J, McCormack AL, Yates JR III. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Amer Soc Mass Spectrom* 1994;**5**:976–89.
21. Yates JR 3rd, Eng JK, McCormack AL, et al. Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. *Anal Chem* 1995;**67**:1426–36.
22. Mann M, Wilm M. Error-tolerant identification of peptides in sequence databases by peptide sequence tags. *Anal Chem* 1994;**66**:4390–9.
23. Boguski MS, Lowe TM, Tolstoshev CM. bEST-database for "expressed sequence tags". *Nat Genet* 1993;**4**:332–3.
24. Adams MD, Kelley JM, Gocayne JD, et al. Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* 1991;**252**:1651–6.
25. Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature* 2003;**422**:198–207.
26. Washburn MP, Wolters D, Yates JR 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 2001;**19**:242–7.

27. Ong SE, Foster LJ, Mann M. Mass spectrometric-based approaches in quantitative proteomics. *Methods* 2003;**29**:24–30.
28. Laiko VV, Baldwin MA, Burlingame AL. Atmospheric pressure matrix-assisted laser desorption/ionization mass spectrometry. *Anal Chem* 2000;**72**:652–7.
29. Hillenkamp F, Karas M, Beavis RC, *et al.* Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. *Anal Chem* 1991;**63**:1193A–1203A.
30. Karas M, Glückmann M, Schäfer J. Ionization in matrix-assisted laser desorption/ionization: singly charged molecular ions are the lucky survivors. *J Mass Spectrom* 2000;**35**:1–12.
31. Hillenkamp F, Karas M, Beavis RC, *et al.* Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. *Anal Chem* 1991;**63**(24):1193A–1203A.
32. Kratzer R, Eckerskorn C, Karas M, *et al.* Suppression effects in enzymatic peptide ladder sequencing using ultraviolet – matrix assisted laser desorption/ionization – mass spectrometry. *Electrophoresis* 1998;**19**:1910–9.
33. Krause E, Wenschuh H, Jungblut PR. The dominance of arginine-containing peptides in MALDI-derived tryptic mass fingerprints of proteins. *Anal Chem* 1999;**71**:4160–5.
34. Beardsley RL, Reilly JP. Optimization of guanidination procedures for MALDI mass mapping. *Anal Chem* 2002;**74**:1884–90.
35. Annan RS, Carr SA. Phosphopeptide analysis by matrix-assisted laser desorption time-of-flight mass spectrometry. *Anal Chem* 1996;**68**:3413–21.
36. Billeci TM, Stults JT. Tryptic mapping of recombinant proteins by matrix-assisted laser desorption/ionization mass spectrometry. *Anal Chem* 1993;**65**:1709–16.
37. Asara JM, Allison J. Enhanced detection of phosphopeptides in matrix-assisted laser desorption/ionization mass spectrometry using ammonium salts. *J Am Soc Mass Spectrom* 1999;**10**:35–44.
38. Gusev AI, Wilkinson WR, Proctor A, *et al.* Direct quantitative analysis of peptides using matrix assisted laser desorption ionisation. *Anal Bioanal Chem* 1996;**354**:455–63.
39. van Montfort BA, Canas B, Duurkens R, *et al.* Improved in-gel approaches to generate peptide maps of integral membrane proteins with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *J Mass Spectrom* 2002;**37**:322–30.
40. Gharahdaghi F, Kirchner M, Fernandez J, *et al.* Peptide-mass profiles of polyvinylidene difluoride-bound proteins by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry in the presence of nonionic detergents. *Anal Biochem* 1996;**233**:94–9.
41. Yu YQ, Gilar M, Lee PJ, *et al.* Enzyme-friendly, mass spectrometry-compatible surfactant for in-solution enzymatic digestion of proteins. *Anal Chem* 2003;**75**:6023–8.
42. Schleuder D, Hillenkamp F, Strupat K. IR-MALDI-mass analysis of electrophoretically separated proteins directly from the membrane: comparison of different membranes, application to on-membrane digestion, and protein identification by database searching. *Anal Chem* 1999;**71**:3238–47.
43. Eckerskorn C, Strupat K, Schleuder D, *et al.* Analysis of proteins by direct-scanning infrared-MALDI mass spectrometry after 2D-PAGE separation and electrophoretic blotting. *Anal Chem* 1997;**69**:2888–92.
44. Berkenkamp S, Karas M, Hillenkamp F. Ice as a matrix for IR-matrix-assisted laser desorption/ionization: mass spectra from a protein single crystal. *Proc Natl Acad Sci USA* 1996;**93**:7003–7.
45. Whitehouse CM, Dreyer RN, Yamashita M, *et al.* Electrospray interface for liquid chromatographs and mass spectrometers. *Anal Chem* 1985;**57**:675–9.
46. Wilm M, Mann M. Analytical properties of the nanoelectrospray ion source. *Anal Chem* 1996;**68**:1–8.
47. Davis MT, Stahl DC, Hefta SA, *et al.* A microscale electrospray interface for on-line, capillary liquid chromatography/tandem mass spectrometry of complex peptide mixtures. *Anal Chem* 1995;**67**:4549–56.
48. Robinson CV. Protein secondary structure investigated by electrospray ionization. *Methods Mol Biol* 1996;**61**:129–39.
49. Robinson CV. Watching and weighting-chaperone complexes in action. *Nat Methods* 2005;**2**:331–2.
50. Chen R, Cheng X, Mitchell DW, *et al.* Trapping, detection, and mass determination of coliphage T4 DNA ions by electrospray ionization fourier transform ion cyclotron resonance mass spectrometry. *Anal Chem* 1995;**67**:1159–63.
51. McLuckey SA, Wells JM. Mass analysis at the advent of the 21st century. *Chem Rev* 2001;**101**:571–606.
52. Yates JR 3rd. Mass spectrometry and the age of proteome. *J Mass Spectrom* 1998;**33**:1–19.
53. Weickhardt C, Moritz F, Grotemeyer J. Time-of-flight mass spectrometry: state-of-the-art in chemical analysis and molecular science. *Mass Spectrom Rev* 1996;**15**:139–62.
54. Jensen ON, Podtelejnikov A, Mann M. Delayed extraction improves specificity in database searches by matrix-assisted laser desorption/ionization peptide maps. *Rapid Commun in Mass Spectrom* 1996;**10**:1371–8.
55. Cornish TJ, Cotter RJ. Tandem time-of-flight mass spectrometer. *Anal Chem* 1993;**65**:1043–7.
56. Burlingame AL, Boyd RK, Gaskell DM. Mass spectrometry. *Anal Chem* 1994;**66**:634R–683R.
57. Kaufmann R. Matrix assisted laser desorption ionization mass spectroscopy for profiling organic constituents of water samples. *J Biotechnol* 1995;**41**:155–75.
58. Stults JT. Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS). *Curr Opin Struct Biol* 1995;**5**:691–8.
59. Mann M, Talbo G. Developments in matrix-assisted laser desorption/ionization peptide mass spectrometry. *Curr Opinion Biotechnol* 1996;**7**:11–9.
60. Karas M. Matrix-assisted laser desorption ionization MS: a progress report. *Biochem Soc Trans* 1996;**24**:897–900.
61. Murray KK. DNA sequencing by mass spectrometry. *J Mass Spectrom* 1996;**31**:1203–15.
62. Harvey DJ. Matrix-assisted laser desorption/ionisation mass spectrometry of oligosaccharides and glycoconjugates. *J Chromatogr A* 1996;**720**:429–46.
63. Leary JL, Schmidt RL. Quadrupole mass spectrometers: an intuitive look at the math. *Journal of Chemical Education* 1996;**73**:1142–4.
64. Steel C, Henchman M. Understanding the quadrupole mass filter through computer stimulation. *J Chemical Edu* 1998;**75**:1049–54.

65. Chowdhury SK, Katta V, Chait BT. An electrospray-ionization mass spectrometer with new features. *Rapid Commun Mass Spectrom* 1990;**4**:81–7.
66. Guan Z, Drader JJ, Campbell VL, et al. Real-time monitoring of the gas phase reactions of a single ion population using the remeasurement experiment in Fourier transform ion cyclotron resonance mass spectrometry. *Anal Chem* 1995;**67**:1453–8.
67. Schwartz JC, Jardine I. Quadrupole ion trap mass spectrometry. *Methods Enzymol* 1996;**270**:552–86.
68. Marina A, Garcia MA, Albar JP, et al. High-sensitivity analysis and sequencing of peptides and proteins by quadrupole ion trap mass spectrometry. *J Mass Spectrom* 1999;**34**:17–27.
69. Yates JR 3rd. Mass spectral analysis in proteomics. *Annu Rev Biophys Biomol Struct* 2004;**33**:297–316.
70. Levander F, James P. Automated protein identification by the combination of MALDI MS and MS/MS spectra from different instruments. *J Proteome Res* 2005;**4**:71–4.
71. Douglas DJ, Frank AJ, Mao D. Linear ion traps in mass spectrometry. *Mass Spectrom Rev* 2005;**24**:1–29.
72. Marshall AG, Hendrickson CL, Jackson GS. Fourier transform ion cyclotron resonance mass spectrometry: a primer. *Mass Spectrom Rev* 1998;**17**:1–35.
73. Gorshkov MV, Pasa Tolic L, Udseth HR, et al. Electrospray ionization-Fourier transform ion cyclotron resonance mass spectrometry at 11.5 tesla: instrument design and initial results. *J Am Soc Mass Spectrom* 1998;**9**:692–700.
74. Lee SW, Berger SJ, Martinovic S, et al. Direct mass spectrometric analysis of intact proteins of the yeast large ribosomal subunit using capillary LC/FTICR. *Proc Natl Acad Sci USA* 2002;**99**:5942–7.
75. Bogdanov B, Smith RD. Proteomics by FTICR mass spectrometry: top down and bottom up. *Mass Spectrom Rev* 2005;**24**:168–200.
76. Hunt DF, Buko AM, Ballard JM, et al. Sequence analysis of polypeptides by collision activated dissociation on a triple quadrupole mass spectrometer. *Biomed Mass Spectrom* 1982;**8**:397–408.
77. Morris HR, Paxton T, Dell A, et al. High sensitivity collisionally-activated decomposition tandem mass spectrometry on a novel quadrupole/orthogonal-acceleration time-of-flight mass spectrometer. *Rapid Commun Mass Spectrom* 1996;**10**:889–96.
78. Le Blanc JC, Hager JW, Ilisiu AM, et al. Unique scanning capabilities of a new hybrid linear ion trap mass spectrometer (Q TRAP) used for high sensitivity proteomics applications. *Proteomics* 2003;**3**:859–69.
79. Wilm M, Shevchenko A, Houthaeve T, et al. Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature* 1996;**379**:466–9.
80. Wilm M, Neubauer G, Mann M. Parent ion scans of unseparated peptide mixtures. *Anal Chem* 1996;**68**:527–33.
81. Steen H, Kuster B, Mann M. Quadrupole time-of-flight versus triple-quadrupole mass spectrometry for the determination of phosphopeptides by precursor ion scanning. *J Mass Spectrom* 2001;**36**:782–90.
82. Bateman RH, Carruthers R, Hoyes JB, et al. A novel precursor ion discovery method on a hybrid quadrupole orthogonal acceleration time-of-flight (Q-TOF) mass spectrometer for studying protein phosphorylation. *J Am Soc Mass Spectrom* 2002;**13**:792–803.
83. Shevchenko A, Loboda A, Shevchenko A, et al. MALDI quadrupole time-of-flight mass spectrometry: a powerful tool for proteomic research. *Anal Chem* 2000;**72**:2132–41.
84. Medzihradszky KF, Campbell JM, Baldwin MA, et al. The characteristics of peptide collision-induced dissociation using a high-performance MALDI-TOF/TOF tandem mass spectrometer. *Anal Chem* 2000;**72**:552–8.
85. Baldwin MA, Medzihradszky KF, Lock CM, et al. Matrix-assisted laser desorption/ionization coupled with quadrupole/orthogonal acceleration time-of-flight mass spectrometry for protein discovery, identification, and structural analysis. *Anal Chem* 2001;**73**:1707–20.
86. Belov ME, Gorshkov MV, Udseth HR, et al. Controlled ion fragmentation in a 2-D quadrupole ion trap for external ion accumulation in ESI FTICR mass spectrometry. *J Am Soc Mass Spectrom* 2001;**12**:1312–9.
87. Pappin DJ. Peptide mass fingerprinting using MALDI-TOF mass spectrometry. *Methods Mol Biol* 2003;**211**:211–9.
88. Schuerenberg M, Luebbert C, Eickhoff H, et al. Prestructured MALDI-MS sample supports. *Anal Chem* 2000;**72**:3436–42.
89. Perkins DN, Pappin DJ, Creasy DM, et al. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 1999;**20**:3551–67.
90. Chamrad DC, Korting G, Stuhler K, et al. Evaluation of algorithms for protein identification from sequence databases using mass spectrometry data. *Proteomics* 2004;**4**:619–28.
91. Pan S, Zhang H, Rush J, et al. High throughput proteome screening for biomarker detection. *Mol Cell Proteomics* 2005;**4**:182–90.
92. Carr S, Aebersold R, Baldwin M, et al. The need for guidelines in publication of peptide and protein identification data: Working Group on Publication Guidelines for Peptide and Protein Identification Data. *Mol Cell Proteomics* 2004;**3**:531–3.
93. Sadygov RG, Cociorva D, Yates JR 3rd. Large-scale database searching using tandem mass spectra: looking up the answer in the back of the book. *Nat Methods* 2004;**1**:195–202.
94. Choudhary JS, Blackstock WP, Creasy DM, et al. Matching peptide mass spectra to EST and genomic DNA databases. *Trends Biotechnol* 2001;**19**(Suppl 10):S17–S22.
95. Lopez-Ferrer D, Martinez-Bartolome S, Villar M, et al. Statistical model for large-scale peptide identification in databases from tandem mass spectra using SEQUEST. *Anal Chem* 2004;**76**:6853–60.
96. Papayannopoulos I. The interpretation of collision-induced dissociation tandem mass spectra of peptides. *Mass Spectrom Rev* 1995;**14**:49–73.
97. Shevchenko A, Chernushevich I, Wilm M, et al. De novo peptide sequencing by nanoelectrospray tandem mass spectrometry using triple quadrupole and quadrupole/time-of-flight instruments. *Methods Mol Biol* 2000;**146**:1–16.
98. Yergey AL, Coorsen JR, Backlund PS Jr, et al. De novo sequencing of peptides using MALDI/TOF-TOF. *J Am Soc Mass Spectrom* 2002;**13**:784–91.
99. Keough T, Youngquist RS, Lacey MP. A method for high-sensitivity peptide sequencing using postsource decay matrix-assisted laser desorption/ionization mass spectrometry. *Proc Natl Acad Sci USA* 1999;**96**(13):7131–6.

100. Uttenweiler-Joseph S, Neubauer G, Christoforidis S, *et al.* Automated de novo sequencing of proteins using the differential scanning technique. *Proteomics* 2001;**1**:668–82.
101. Mann M, Jensen ON. Proteomic analysis of post-translational modifications. *Nat Biotechnol* 2003;**21**:255–61.
102. Warden SM, Richardson C, O'Donnell J, *et al.* Post-translational modifications of the beta-1 subunit of AMP-activated protein kinase affect enzyme activity and cellular localization. *Biochem J* 2001;**354**:275–83.
103. Estrada-Mondaca S, Lougarre A, Fournier D. Drosophila acetylcholinesterase: effect of post-translational [correction of post-traductional] modifications on the production in the baculovirus system and substrate metabolism. *Arch Insect Biochem Physiol* 1998;**38**:84–90.
104. Ozols J. Luminal orientation and post-translational modification of the liver microsomal 11 beta-hydroxysteroid dehydrogenase. *J Biol Chem* 1995;**270**:2305–12.
105. Fraga MF, Ballestar E, Villar-Garea A, *et al.* Loss of acetylated Lysine 16 and trimethylated Lysine 20 of Histone H4 is a common hallmark of human cancer. *Nat Genet* 2005;**37**:391–400.
106. Deterding JL, Banks GC, Tomer KB, *et al.* Understanding global changes in histone H1 phosphorylation using mass spectrometry. *Methods* 2004;**33**:53–8.
107. Teasdale RD, Matheson F, Gleeson PA. Post-translational modifications distinguish cell surface from Golgi-retained beta 1,4 galactosyltransferase molecules. Golgi localization involves active retention. *Glycobiology* 1994;**4**:917–28.
108. Nepomuceno RR, Ruiz S, Park M, *et al.* C1qR_P Is a heavily O-Glycosylated cell surface protein involved in the regulation of phagocytic activity. *J Immunol* 1999;**162**:3583–9.
109. Carter AM, Hills F, O'Gorman DB, *et al.* The insulin-like growth factor system in mammalian pregnancy. *Placenta* 2004;**18**:53–6.
110. Peng J, Schwartz D, Elias JE, *et al.* A proteomic approach to understanding protein ubiquitination. *Nat Biotechnol* 2003;**21**:921–6.
111. Monilola A Olayioye. Intracellular signaling pathways of ErbB2/HER-2 and family members. *Breast Cancer Res* 2001;**3**:385–9.
112. Sherryline Jogie-Brahim, Hae-Ki Min, Youngman Oh. Potential of proteomics towards the investigation of the IGF-independent actions of IGFBP-3. *Expert Review of Proteomics* 2005;**2**:71–86.
113. Steen H, Kuster B, Fernandez M, *et al.* Tyrosine phosphorylation mapping of the epidermal growth factor receptor signaling pathway. *J Biol Chem* 2002;**277**:1031–9.
114. Pelech S. Tracking cell signaling protein expression and phosphorylation by innovative proteomic solutions. *Curr Pharm Biotechnol* 2004;**5**:69–77.
115. Sørensen ES, Højrup P, Petersen TE. Post-translational modifications of bovine osteopontin: identification of twenty-eight phosphorylation and three O-glycosylation sites. *Protein Science* 1995;**4**:2040–9.
116. Saito S, Yamaguchi H, Higashimoto Y, *et al.* Phosphorylation site interdependence of human p53 post-translational modifications in response to stress. *J Biol Chem* 2003;**278**:37536–44.
117. Kim JK, Mastronardi FG, Wood DD, *et al.* Multiple sclerosis. An important role for post-translational modifications of myelin basic protein in pathogenesis. *Mol Cell Proteomics* 2003;**2**:453–62.
118. Andersson L, Porath J. Isolation of phosphoproteins by immobilized metal (Fe³⁺) affinity chromatography. *Anal Biochem* 1996;**154**:250–4.
119. Ficarro SB, McClelland ML, Stukenberg PT, *et al.* Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat Biotechnol* 2002;**20**:301–5.
120. Nagasu OY, Chait BT. Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. *Nat Biotechnol* 2001;**19**:379–82.
121. Zhou H, Watts J, Aebersold R. A systematic approach to the analysis of protein phosphorylation. *Nat Biotechnol* 2001;**19**:375–8.
122. Hicke L. A new ticket for entry into budding vesicles ubiquitin. *Cell* 2001;**106**:527–30.
123. Weissman AM. Themes and variations on ubiquitination. *Mol Cell Biol* 2001;**2**:169–78.
124. Faleiro L, Kobayashi R, Fearnhead H, *et al.* Multiple species of CPP32 and Mch2 are the major active caspases present in apoptotic cells. *EMBO J* 1997;**16**: 2271–81.
125. Baruch A, Greenbaum D, Levy ET, *et al.* Defining a link between gap junction communication, proteolysis, and cataract formation. *J Biol Chem* 2001;**276**:28999–9006.
126. Greenbaum D, Baruch A, Grainger M, *et al.* Identification of a protease-dependent invasion pathway in the human malarial parasite, *Plasmodium falciparum*. *Science* 2002;**298**:2002–6.
127. Adam GC, Sørensen EJ, Cravatt BF. Proteomic profiling of mechanistically distinct enzyme classes using a common chemotype. *Nat Biotechnol* 2002;**20**:805–9.
128. Jessani N, Liu Y, Humphrey M, Cravatt BF. Enzyme activity profiles of the secreted and membrane proteome that depict cancer cell invasiveness. *Proc Nat Acad Sci USA* 2002;**99**:10335–40.
129. Garcia-Echeverria C, Traxier P, Evans DB. ATP site-directed competitive and irreversible inhibitors of protein kinases. *Med Res Rev* 2000;**20**:28–57.
130. Denu JM, Dixon JE. Protein tyrosine phosphatases: mechanisms of catalysis and regulation. *Curr Opin Chem Biol* 1998;**2**:633–41.
131. Arabcaï G, Guo X, Beebe KD, *et al.* Alpha-Haloacetophenone derivatives as photoreversible covalent inhibitors of protein tyrosine phosphatases. *J Am Chem Soc* 1999;**121**:5085–6.
132. Sorace JM, Zhan M. A data review and re-assessment of ovarian cancer serum proteomic profiling. *BCM Bioinformatics* 2003;**4**:24.
133. Hoom EJ, Pisitkun T, Zietse R, *et al.* Prospects for urinary proteomics: Exosomes as a source of urinary biomarkers. *Nephrology* 2005;**10**:283–90.
134. Han KK, Martinage A. Post-translational chemical modification(s) of proteins. *Int J Biochem* 1992;**24**:19–28.
135. Annan RS, Carr SA. The essential role of mass spectrometry in characterizing protein structure: mapping posttranslational modifications. *J Prot Chem* 1997;**16**:391–402.
136. Huq MD, Wei LN. Post-translational modification of nuclear co-repressor receptor interacting protein 140 by acetylation. *Mol Cell Proteomics* 2005;**4**:975–83.

137. Johnson L, Mollah S, Garcia BA, *et al.* Mass spectrometry analysis of Arabidopsis histone H3 reveals distinct combinations of post-translational modifications. *Nucleic Acids Res* 2004;**32**:6511–8.
138. Shen EC, Henry MF, Weiss VH, *et al.* Arginine methylation facilitates the nuclear export of hnRNP proteins. *Genes Dev* 1998;**12**:679–91.
139. Bedford MT, Frankel A, Yaffe MB, *et al.* Arginine methylation inhibits the binding of proline-rich ligands to Src Homology 3, but not WW, domains. *J Biol Chem* 2000;**275**:16030–6.
140. Brahms H, Meheus L, de Brabandere V, *et al.* Symmetrical dimethylation of arginine residues in spliceosomal Sm protein B/B' and the Sm-like protein LSm4, and their interaction with the SMN protein. *RNA* 2001;**7**:1531–42.
141. Gary JD, Clarke S. RNA and protein interactions modulated by protein arginine methylation. *Prog Nucl Acid Res Mol Biol* 1998;**61**:65–130.
142. Denman RB. Methylation of the Arginine-Glycine-rich region in the fragile X mental retardation protein FMRP differentially affects RNA Binding. *Cell Mol Biol Lett* 2002;**7**:877–83.
143. Boisvert FM, Côté J, Boulanger MC, *et al.* A proteomic analysis of arginine-methylated protein complexes. *Mol Cell Proteom* 2003;**2**:1319–30.
144. Zhang L, Freitas MA. Comparison of peptide mass mapping and electron capture dissociation as assays for histone posttranslational modifications. *Int J Mass Spectrom* 2004;**234**:213–25.
145. Zhang K, Yau PM, Chandrasekhar B, *et al.* Differentiation between peptides containing acetylated or tri-methylated lysines by mass spectrometry: An application for determining lysine acetylation and methylation of histone H3. *Proteomics* 2004;**4**:1–10.
146. Bonaldi T, Imhof A, Regula JT. A combination of different mass spectroscopic techniques for the analysis of dynamic changes of histone modifications. *Proteomics* 2004;**4**:1382–96.
147. Rudd PM, Colomina C, Royle L, *et al.* A high-performance liquid chromatography based strategy for rapid, sensitive sequencing of N-linked oligosaccharide modifications to proteins in sodium dodecyl sulphate polyacrylamide electrophoresis gel bands. *Proteomics* 2001;**1**:285–94.
148. Dell A, Morris HR. Glycoprotein structure determination by mass spectrometry. *Science* 2001;**291**:2351–6.
149. Mechref Y, Novotny MV. Structural investigations of glycoconjugates at high sensitivity. *Chem Rev* 2002;**102**:321–69.
150. Zaia J. Mass spectrometry of oligosaccharides. *Mass Spectrom Rev* 2004;**23**:161–227.
151. Zhang H, Li XJ, Martin DB, *et al.* Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat Biotechnol* 2003;**21**:660–6.
152. Kaji H, Saito Y, Yamauchi Y, *et al.* Lectin affinity capture, isotope-coded tagging and mass spectrometry to identify N-linked glycoproteins. *Nat Biotechnol* 2003;**21**:667–72.
153. Jaffrey SR, Erdjument-Bromage H, Ferris CD, *et al.* Protein S-nitrosylation: a physiological signal for neuronal nitric oxide. *Nat Cell Biol* 2001;**3**:193–7.
154. Kuncewicz T, Sheta EA, Goldknopf IL, *et al.* Proteomic analysis of S-nitrosylated proteins in mesangial cells. *Mol Cell Proteomics* 2003;**2**:156–63.
155. Martínez-Ruiz A, Lamas S. Detection and proteomic identification of S-nitrosylated proteins in endothelial cells. *Arch Biochem Biophys* 2004;**423**:193–200.
156. Martínez-Ruiz A, Villanueva L, González de Orduña C, *et al.* S-nitrosylation of Hsp90 promotes the inhibition of its ATPase and endothelial nitric oxide synthase regulatory activities. *Proc Natl Acad Sci USA* 2005;**102**:8525–30.
157. Link AJ, Eng J, Schieltz DM, *et al.* Direct analysis of protein complexes using mass spectrometry. *Nat Biotechnol* 1999;**17**:676–82.
158. Peng J, Elias JE, Thoreen CC, *et al.* Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J Proteome Res* 2003;**2**:43–50.
159. Keller A, Nesvizhskii AI, Kolker E, *et al.* Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* 2002;**74**:5383–92.
160. Tabb DL, MacCoss MJ, Wu C, *et al.* Similarity among tandem mass spectra from proteomic experiments: detection, significance, and utility. *Anal Chem* 2003;**75**:2470–7.
161. Sadygov RG, Yates JR 3rd. A hypergeometric probability model for protein identification and validation using tandem mass spectral data and protein sequence databases. *Anal Chem* 2003;**75**:3792–8.
162. Florens L, Washburn MP, Raine JD, *et al.* A proteomic view of the Plasmodium falciparum life cycle. *Nature* 2002;**419**:520–6.
163. Carlton JM, Angiuoli SV, Suh BB, *et al.* Genome sequence and comparative analysis of the model rodent malaria parasite *Plasmodium yoelii yoelii*. *Nature* 2002;**419**:512–9.
164. Doolan DL, Southwood S, Freilich DA, *et al.* Identification of Plasmodium falciparum antigens by antigenic analysis of genomic and proteomic data. *Proc Natl Acad Sci USA* 2003;**100**:9952–7.
165. Florens L, Liu X, Wang Y, *et al.* Proteomics approach reveals novel proteins on the surface of malaria-infected erythrocytes. *Mol Biochem Parasitol* 2004;**135**:1–11.
166. Hall N, Karras M, Raine JD, *et al.* A comprehensive survey of the Plasmodium life cycle by genomic, transcriptomic, and proteomic analyses. *Science* 2005;**307**:82–6.
167. Scherl A, Francois P, Converset V, *et al.* Nonredundant mass spectrometry: a strategy to integrate mass spectrometry acquisition and analysis. *Proteomics* 2004;**4**:917–27.
168. Gelpi E. Interfaces for coupled liquid-phase separation/mass spectrometry techniques. An update on recent developments. *J Mass Spectrom* 2002;**37**:241–53.
169. Bodnar WM, Blackburn RK, Krise JM, *et al.* Exploiting the complementary nature of LC/MALDI/MS/MS and LC/ESI/MS/MS for increased proteome coverage. *J Am Soc Mass Spectrom* 2003;**14**:971–9.
170. Krokhn O, Li Y, Andonov A, *et al.* Mass spectrometric characterization of proteins from the SARS virus: A preliminary report. *Mol Cell Proteomics* 2003;**2**:346–56.
171. O'Farrell PH. High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* 1975;**250**:4007–21.

172. Klose J. Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik* 1975;**26**:231–43.
173. Gygi SP, Rist B, Gerber SA, *et al.* Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* 1999;**17**:994–9.
174. Ross P, Huang Y, Marchese JN, *et al.* Multiplexed protein quantitation in *saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* 2004;**3**:1154–69.
175. Yao X, Freas A, Ramirez J, *et al.* Proteolytic ^{18}O labeling for comparative proteomics: model studies with two serotypes of adenovirus. *Anal Chem* 2001;**73**:2836–42.
176. Yao X, Afonso C, Fenselau C. Dissection of proteolytic ^{18}O labeling: endoprotease-catalyzed ^{16}O to ^{18}O exchange of truncated peptide substrates. *J Proteom Res* 2003;**2**:147–52.
177. Oda Y, Huang K, Cross FR, *et al.* Accurate quantitation of protein expression and site-specific phosphorylation. *Proc Natl Acad Sci USA* 1999;**96**:6591–6.
178. Pasa-Tolic L, Jensen PK, Anderson GA, *et al.* High throughput proteome-wide precision measurements of protein expression using mass spectrometry. *J Am Chem Soc* 1999;**121**:7949–50.
179. Ong SE, Blagoev B, Kramatchova I, *et al.* Stable isotope labelling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 2002;**5**:376–86.
180. Gerber SA, Rush J, Stemman O, *et al.* Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc Natl Acad Sci USA* 2003;**100**:6940–45.
181. Wright ME, Han DK, Aebersold R. Mass spectrometry-based expression profiling of clinical prostate cancer. *Mol Cell Proteomics* 2005;**4**:545–54.
182. Yi EC, Li X, Cooke K, *et al.* Increased quantitative proteome coverage with $^{13}\text{C}/^{12}\text{C}$ -based, acid-cleavable isotope-coded affinity tag reagent and modified data acquisition scheme. *Proteomics* 2005;**5**:380–87.
183. Han DK, Eng J, Zhou H, *et al.* Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. *Nat Biotechnol* 2001;**19**:946–51.
184. Guina T, Purvine SO, Yi EC, *et al.* Quantitative proteomic analysis indicates increased synthesis of a quinolone by *Pseudomonas aeruginosa* isolates from cystic fibrosis airways. *Proc Natl Acad Sci USA* 2003;**100**:2771–6.
185. Schmidt F, Donahoe S, Hagens K, *et al.* Complementary analysis of the *Mycobacterium tuberculosis* proteome by two-dimensional electrophoresis and isotope-coded affinity tag technology. *Mol Cell Proteomics* 2004;**3**:24–42.
186. Li C, Hong Y, Tan YX, *et al.* Accurate qualitative and quantitative proteomic analysis of clinical hepatocellular carcinoma using laser capture microdissection coupled with isotope-coded affinity tag and two-dimensional liquid chromatography mass spectrometry. *Mol Cell Proteomics* 2004;**3**:399–409.
187. Yan W, Lee H, Yi EC, *et al.* System-based proteomic analysis of the interferon response in human liver cells. *Genome Biol* 2004;**5**:R54.
188. Yan W, Lee H, Deutsch EW, *et al.* A dataset of human liver proteins identified by protein profiling via isotope-coded affinity tag (ICAT) and tandem mass spectrometry. *Mol Cell Proteomics* 2004;**3**:1039–41.
189. Zhou H, Ai JH, Li SJ, *et al.* Accurate qualitative and quantitative proteomic analysis of clinical hepatocellular carcinoma using laser capture microdissection coupled with isotope-coded affinity tag and two-dimensional liquid chromatography mass spectrometry. *Mol Cell Proteomics* 2004;**3**:399–409.
190. Tam EM, Morrison CJ, Wu YI, *et al.* Membrane protease proteomics: Isotope-coded affinity tag MS identification of undescribed MT1-matrix metalloproteinase substrates. *Proc Natl Acad Sci USA* 2004;**101**:6917–22.
191. Shiio Y, Donohoe S, Yi EC, *et al.* Quantitative proteomic analysis of Myc oncoprotein function. *EMBO J* 2002;**21**:5088–96.
192. Igarashi K, Francastel C, Chi TH, *et al.* Dynamic changes in transcription factor complexes during erythroid differentiation revealed by quantitative proteomics. *Nat Struct Mol Biol* 2004;**11**:73–80.
193. Conrads TP, Alving K, Veenstra TD, *et al.* Quantitative analysis of bacterial and mammalian proteomes using a combination of cysteine affinity tags and ^{15}N -metabolic labeling. *Anal Chem* 2001;**73**:2132–9.
194. Washburn MP, Ulaszek R, Deciu C, *et al.* Analysis of quantitative proteomic data generated via multidimensional protein identification technology. *Anal Chem* 2002;**74**:1650–7.
195. Jiang H, English AM. Quantitative analysis of the yeast proteome by incorporation of isotopically labeled leucine. *J Proteome Res* 2002;**1**:345–50.
196. Foster LJ, De Hoog CL, Mann M. Unbiased quantitative proteomics of lipid rafts reveals high specificity for signaling factors. *Proc Natl Acad Sci USA* 2003;**100**:5813–8.
197. Blagoev B, Kratchmarova I, Ong SE, *et al.* A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nat Biotechnol* 2003;**21**:315–8.
198. Ong SE, Mittler G, Mann M. Identifying and quantifying in vivo methylation sites by heavy methyl SILAC. *Nature Methods* 2004;**1**:119–26.
199. Hinsby AM, Olsen JV, Mann M. Tyrosine phosphoproteomics of fibroblast growth factor signaling – a role for insulin receptor substrate-4. *J Biol Chem* 2004;**279**:46438–47.
200. Gu S, Liu Z, Pan S, *et al.* Global investigation of p53-induced apoptosis through quantitative proteomic profiling using comparative amino acid-coded tagging. *Mol Cell Proteomics* 2004;**3**:998–1008.
201. Andersen JS, Lam YW, Leung AK, *et al.* Nucleolar proteome dynamics. *Nature* 2005;**433**:77–83.
202. Gruhler A, Olsen JV, Mohammed S, *et al.* Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. *Mol Cell Proteomics* 2005;**4**:310–27.
203. Krijgsveld J, Ketting RF, Mahmoudi T, *et al.* Metabolic labeling of *C. elegans* and *D. melanogaster* for quantitative proteomics. *Nat Biotechnol* 2003;**21**:927–31.
204. Ippel JH, Pouvreau L, Kroef T, *et al.* In vivo uniform (^{15}N)-isotope labelling of plants: using the greenhouse for structural proteomics. *Proteomics* 2004;**4**:226–34.
205. Wu CC, MacCoss MJ, Howell KE, *et al.* Metabolic labeling of mammalian organisms with stable isotopes for

- quantitative proteomic analysis. *Anal Chem* 2004;**76**: 4951–9.
206. Ishihama Y, Sato T, Tabata T, *et al.* Quantitative mouse brain proteomics using culture-derived isotope tags as internal standards. *Nat Biotechnol* 2005;**23**:617–21.
207. Zang L, Toy DP, Hancock WS, *et al.* Proteomic analysis of ductal carcinoma of the breast using laser capture microdissection, LC-MS, and $^{16}\text{O}/^{18}\text{O}$ isotopic labeling. *J Proteom Res* 2004;**3**:604–12.
208. Brown C, Fenselau CJ. Investigation of doxorubicin resistance in MCF-7 breast cancer cells using shot-gun comparative proteomics with proteolytic ^{18}O labeling. *Proteom Res* 2004;**3**:455–62.
209. Qian WJ, Monroe ME, Liu T, *et al.* Quantitative proteome analysis of human plasma following in vivo lipopolysaccharide administration using $^{16}\text{O}/^{18}\text{O}$ labeling and the accurate mass and time tag approach. *Mol Cell Proteomics* 2005;**4**:700–9.
210. Chen X, Cushman SW, Pannell LK, *et al.* Quantitative proteomic analysis of the secretory proteins from rat adipose cells using a 2D liquid chromatography-MS/MS approach. *J Proteome Res* 2005;**4**:570–7.
211. Blonder J, Hale ML, Chan KC, *et al.* Quantitative profiling of the detergent-resistant membrane proteome of iota-b toxin induced vero cells. *J Proteome Res* 2005;**4**: 523–31.
212. Chelius D, Bondarenko PV. Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. *J Proteome Res* 2002;**1**:317–23.
213. Wang W, Zhou H, Lin H, *et al.* Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. *Anal Chem* 2003;**75**: 4818–26.
214. Wiener MC, Sachs JR, Deyanova EG, Yates NA. Differential mass spectrometry: a label-free LC-MS method for finding significant differences in complex peptide and protein mixtures. *Anal Chem* 2004;**76**: 6085–96.
215. Higgs RE, Knierman MD, Gelfanova V, *et al.* Comprehensive label-free method for the relative quantification of proteins from biological samples. *J Proteome Res* 2005;**4**:1442–50.
216. Conrads TP, Anderson GA, Veenstra TD, *et al.* Utility of accurate mass tags for proteome-wide protein identification. *Anal Chem* 2000;**72**:3349–54.
217. Kelleher NL. Top-down proteomics. *Anal Chem* 2004;**1**: 197A–203A.
218. Pesavento JJ, Kim YB, Taylor GK, *et al.* Shotgun annotation of histone modifications: a new approach for streamlined characterization of proteins by top down mass spectrometry. *J Am Chem Soc* 2004;**24**:3386–7.
219. Horn DM, Zubarev RA, McLafferty FW. Automated de novo sequencing of proteins by tandem high-resolution mass spectrometry. *Proc Natl Acad Sci USA* 2000;**97**: 10313–7.
220. McLafferty FW, Fridriksson EK, Horn DM, *et al.* Biomolecule mass spectrometry. *Science* 1999;**284**: 1289–90.
221. Nemeth-Cawley JF, Tangarone BS, Rouse JC. “Top Down” characterization is a complementary technique to peptide sequencing for identifying protein species in complex mixtures. *J Proteome Res* 2003;**2**:495–505.