Developments in mass spectrometry for the analysis of complex protein mixtures
Gurusahai Khalsa-Moyers and William Hayes McDonald
Advance Access publication date 18 May 2006

Abstract
State-of-the-art proteomics workflows involve multiple interdependent steps: sample preparation, protein—peptide separation, mass spectrometry and data analysis. While improvements in any of these steps can increase the depth and breadth of analysis, advances in mass spectrometry have catalysed many of the most important developments. We discuss common classes of mass analysers and how these analysers are put together to produce some of the most popular mass spectrometry platforms. The capabilities of these platforms determine how they can be used in a variety of common proteomic strategies and, in turn, what types of biological questions can be addressed. Moving forward, powerful new hybrid mass spectrometers and application of emerging types of tandem mass spectrometry promise that our ability to analyse complex mixtures of proteins will continue to advance.

Keywords: proteomics; technologies; mass spectrometers; tandem mass spectrometry

INTRODUCTION
Over the past several years, our ability to characterize complex mixtures has advanced dramatically. This has been especially true for complex biological samples and is a major enabler of the ‘omics (genomics, transcriptomics, metabolomics, proteomics, etc.) revolution. Whether this charge towards systems biology is a true paradigm shift or merely a change in the scale and scope of data being collected is an open debate and certainly outside the framework of this review. Regardless, new technologies are changing the scope of the questions that can be addressed in biology and are extending our measurements out of the classic model organisms into broader studies of disease and even into environmental microbial communities.

The linguistic and experimental antecedent of proteomics, protein chemistry, has been practised for decades. Both one-dimensional sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (SDS-PAGE) and, providing even greater separative capacity, two-dimensional gel electrophoresis (2D-PAGE) are powerful tools for the separation and analysis of proteins. One common workflow in proteomics involves 2D-PAGE followed by mass-spectrometry-based identification. Because of differences in molecular weight and isoelectric points, different post-translationally modified forms can often be resolved as separate spots on the gel. While these patterns can be observed, it is often quite problematic to determine the specific combination of modifications that constitute a particular spot. Top-down proteomics, a term coined by McLafferty and associates [1], refers to the accurate measurement of intact protein mass followed by direct fragmentation of whole proteins [1–3]. Analysis of fragment ions leads to identification of N- and C-termini, large sections of protein primary sequence, and positions of post-translational modifications (PTMs). The major advantages of top-down proteomics are the availability of the entire protein sequence for characterization via tandem mass spectrometry (MS/MS) and the ability to measure multiple forms of PTMs [2].
The term shotgun proteomics (also known as bottom-up) was named in analogy to shotgun DNA sequencing strategies. It refers to strategies that involve the proteolytic digestion of a mixture of proteins and the subsequent analysis of the resulting even more complex mixture of peptides, usually by MS/MS [4]. The fundamental reason that shotgun proteomics has demonstrated a deeper analytical capacity than many protein-separation first techniques is that the analytical challenges associated with separating and analysing peptides are much more straightforward than those for intact proteins. Several of these strategies will be highlighted through the remainder of this review.

Of the 'omics (genomics, transcriptomics, metabonomics, proteomics, etc.), proteomics has been historically most closely tied to equipment advances in mass spectrometry. The remainder of this review is focused on these advances, emerging technologies and most importantly how they are being deployed. However, it is essential to note that these analyses do not simply rely on the magic of the mass spectrometer. Sample preparation, separation, ionization, spectral analysis and data filtering and contextualization all play essential roles in the process. While we will not be discussing these in depth, it is important that they be introduced in order to provide context for both the current and future advances in MS.

SAMPLE PREPARATION
Sample preparation is an often overlooked aspect of a proteomics workflow, at least for those who have not had to put together a proteomics workflow themselves. Proper tuning of protocols not only for focusing on the necessary set of proteins but also for compatibility with downstream analyses results in series of compromises. For instance, sets of ionic detergents that could be easily employed in SDS-PAGE might not be compatible with a 2D-PAGE protocol. Considerations would further change if a shotgun procedure is planned. For instance, several protocols have been developed for targeting the important subclass of integral membrane and associated proteins (reviewed in [5]). Major advances in the effectiveness of proteomics workflows have arisen from advances in separating and analysing the peptides from these proteolytic digests.

CHROMATOGRAPHIC SEPARATIONS
While 2D-PAGE and chromatographic separations of proteins have been a mainstay of biochemistry and continue to make important contributions to proteomic analyses, peptides provide a much more straightforward separation challenge. This is especially true if a subset of the peptides provide sufficient information for unequivocal identification of the constituent proteins. For interface with electrospray ionization (ESI) (see subsequently), some of the most popular and effective techniques use high-performance liquid chromatography (HPLC). A liquid interface allows for coupling the separation step directly to the ionization source of the mass spectrometer. Reverse-phase (RP) chromatography is the most commonly used technique for separating peptides, although many other techniques and variations have been employed with success (reviewed in [6, 7]). For the most complex mixtures of proteins or peptides, it is necessary to add another dimension of separation. This is most commonly accomplished using a strong-cation-exchange (SCX) column for the first dimension, followed by multiple RP separations. A popular version of this multidimensional setup developed in the Yates lab uses an integrated SCX-RP column and is known as multidimensional protein identification technology (MudPIT) [8–10]. Many other possibilities exist for interfacing peptide separation with mass spectrometry. For instance, fractions from an HPLC separation can be collected on a matrix-assisted laser desorption ionization (MALDI) (see the following text) target for analysis. This has the added advantage of allowing the HPLC separation to be ‘frozen’, and thus the downstream analysis is not limited to the time frame of the HPLC [11]. Other separation strategies that have been utilized in proteomics workflows include other types of HPLC, isoelectric focusing, capillary electrophoresis and chip-based separations [12, 13].

IONIZATION
Part of the revolution in biological MS was facilitated by the development of new methods for soft ionization, specifically ESI [14] and MALDI [15]. It would be difficult to overstate the importance of these techniques and even more of a challenge to describe all of the many variations on how they have been employed. As mentioned previously, the ability to interface these ionization methods with a variety of upstream processing and separation strategies have made them even more powerful. Low flow micro-capillary HPLC separations have been particularly
well suited to coupling with nano-ESI (e.g. flow rate of 300 nl/min) and provide the maximum sensitivity for analysis. Interesting variations on MALDI include the Ciphergen protein chip assays [16, 17] and tissue-section mass spectrometry based ‘imaging’. Recently, a new ionization technique, desorption electrospray ionization (DESI) has the potential to lead to new workflows for the direct analysis of a variety of molecular species [18–20]. Also of note for the future is the potential to utilize another ESI variant, electrosonic spray ionization (ESSI), to further broaden the classes of molecules which are able to be ionized [21].

DATA ANALYSIS
With the emergence of extended separations and automated sample processing and data acquisition, the volume of data quickly outpaces any possibility of manual evaluation. This has proven true for every high-throughput workflow. For instance, protein spots from 2D-PAGE can be identified by mapping the masses of tryptic peptide fragments back to those predicted from database sequencing. Automated programs exist for performing such comparisons (e.g. FRAGFIT [22]). MS/MS ‘sequencing’ of peptides presents an even greater challenge. This is most often addressed by comparing the fragmentation pattern of a given MS/MS spectrum with modelled spectra generated from sequences within the appropriate protein database. Sequest [23, 24] and subsequent algorithms including the popular Mascot search package [25] have automated this process and continue to be important areas of development. However, even with automation of the search itself, there still remains a substantial challenge in organizing and contextualizing the search results. Not only must these data be filtered and collated (e.g. via DTASelect [26], Interact [27], etc.), but also some statistical evaluation of the quality of those data at the protein and peptide levels must be performed (e.g. Peptide/Protein Prophet [28, 29], Q-score [30] and others [31, 32]). Finally, since many peptides for which we have spectra may not be found in the database, another focus of development is in de novo algorithms (e.g. Lutefisk [33, 34], GutenTag [35] and PEAKS [36]) that generate peptide sequence (or at least a short sequence tag) directly from the MS/MS spectrum. Algorithmic development is also pushing forward ‘top-down’ analysis of intact proteins by mass spectrometry (e.g. ProSight [37]). Again, an in-depth discussion of these algorithms and their applications is beyond the scope of this review, but developments in this area will continue to allow researchers to make better sense of the data being collected in high-throughput proteomics workflows.

Types of mass analysers
Even with all of the developments in proteomics workflows, new mass spectrometry techniques will continue to advance the types of experiments that can be performed. A portion of these are hardware-based. New analysers, hybridizations and the addition of more mechanisms for performing MS/MS promise capabilities not only for the analysis of peptides but also for intact proteins. Often, the magic of these instruments is in how the different mass analysers are arranged and interfaced with a particular ionization and ion detection strategy. The different geometries of these hybridizations will affect not only the performance of the individual mass analysers but also the types of MS/MS that can be performed. We will begin by describing the most commonly used analysers, along with some new implementations, and then describe the more popular ways of combining these analysers.

The multipole mass filter is an incredibly important tool and is used in many mass spectrometers, often in series with other mass analysers. Usually implemented as four round or hyperbolic rods (quadrupole, often affectionately shortened to quad), it uses differential alternating voltages (AC) to pass through ions of a particular mass-to-charge \( m/z \) ratio. Depending on how these voltages are applied, quadrupole mass filters can be used to selectively pass ions of a particular \( m/z \) (Q) or to pass through a much broader ion packet (also known as RF-only quadrupole, q). Often the latter form is filled with an inert gas and serves as collision cells for collision-induced dissociation (CID). Variations on RF-only multipoles include those with higher numbers of rods, hexapoles and octopoles; these are primarily used to transmit and/or accumulate ions. For a mass-selectable quadrupole, a complete spectrum is obtained by ‘scanning’ across the \( m/z \) range and measuring the relative signal at a particular \( m/z \). Resolution is affected by the speed with which this scan takes place, the voltages which can be applied, and tolerances to which the rods are milled.
Another popular type of mass analyser is the time-of-flight (TOF). It measures the \( m/z \) of a particular ion based on the time required, after an initial acceleration, for that ion to traverse a field-free region. The larger the ion, the longer the time needed for its flight. One way to increase the resolution of this analyser is to increase the length of the ion flight path. Reflectron instruments accomplish this by reflecting the ions back down the flight tube using a series of plates to decelerate and eventually reflect the ions. An additional pass (also known as W optics) effectively quadruples the flight path for the ions. Reflectron instruments further act to improve mass-to-charge resolution by correcting for differences in initial ion formation time and energies.

The quadrupole ion trap (QIT) has long been a workhorse instrument in many proteomics workflows. It uses a combination of RF and DC voltages to select ions of a particular \( m/z \); in this case, instead of passing on those ions, they are trapped in three dimensions. Generation of a mass spectrum involves trapping of a broader range of \( m/z \) values and then scanning them out of the trap to a detector using a ramping of RF voltages. These instruments are particularly well-suited to MS/MS experiments since ions of specific \( m/z \) can be trapped and then, using specific waveforms, excited for collision-induced dissociation CID (see below). The resulting fragment ions can then be scanned out to generate an MS/MS spectrum. These instruments allow even higher-order tandem mass spectrometry (MS\(^n\)) experiments in which a fragment ion from the initial MS/MS can be isolated, activated and fragmented further. This process can continue for as long as there are sufficient ions.

A relative of the QIT and the quadrupole mass filter is the linear, or 2D, ion trap (LIT). These analysers resemble quadrupoles except that additional DC potentials allow for trapping of ions along the long axis. Scanning is performed through a ramping protocol that either ejects ions radially (e.g. Thermo LTQ) or axially (e.g. ABI/Sciex Q-Trap) (Figure 1). A major advantage of the 2D traps is that they possess greater trapping volumes and can thus analyse more ions per cycle, with concomitant improvements in sensitivity and dynamic range. Depending on the specific geometries of the LIT, they can perform similar manipulations to those performed in the quadrupole trap.

**Figure 1:** Common formats for analysis of peptide and proteins via mass spectrometry. MALDI uses a laser pulse to ionize peptide or protein sample embedded in a matrix material. Often coupled to MALDI ionization sources are TOF mass analysers. TOF analysers work on the principle of ion flight time down a flight tube, where ions of lower \( m/z \) reach the detector before those of larger \( m/z \). Two TOF analysers can be connected in a sequential manner and fragmentation induced in a collision cell between the two TOF analysers, a ‘tandem-in-space’ MS/MS. ESI is often coupled to ion trap mass analysers. Voltages, both rf and DC, are selectively applied to the ion trap in order to focus and select ions of a particular \( m/z \). MS/MS (as indicated by activated ion) can be either tandem in time or tandem in space, depending on the instrument. The ESI ion trap and the ESI–LTQ both perform tandem-in-time mass spectrometry in which the precursor ions are first selected by application of voltage to the trap, selected ions fragmented and subsequent fragment ions detected. Tandem MS in the ESI-QqTOF and ESI-Q-Trap both operate as tandem in space, with precursor ion activation occurring in the second quadrupole.
The ‘grand-daddy’ analysers used in proteomics applications are the Fourier transform ion cyclotron resonance mass spectrometers (FT-ICR or FTMS) (Figure 2). They measure mass indirectly by oscillating ions in a strong magnetic field. Because these ions will oscillate as a function of their $m/z$, measuring the frequency of these oscillations allows inference of $m/z$ using a Fourier transform (FT). These instruments provide the highest combination of mass accuracy, resolution and intrascan dynamic range. While they have long been limited to expert-only laboratories, new instrumentation promises to bring this technology to the masses. The first truly novel mass analyser introduced in the last 25 years, the orbitrap, also uses an FT-based strategy to measure $m/z$. However, the trapping is performed electrostatically (as opposed to magnetically) and the frequency oscillations are measured along the long axis of the trapping cell (Figure 2). Even though magnetic FT-ICR instruments should continue to provide the highest levels of performance, the orbitrap provides a compelling

![Figure 2](https://academic.oup.com/bfg/article-abstract/5/2/98/210883)

**Figure 2:** High-power hybrid instruments for the analysis of complex mixtures. Shown are diagrams of hybrid mass spectrometers that link robust front-end MS/MS with the high mass accuracy, resolution and ion capacities of Fourier transform (FT) mass analysers. The first hybrid Qq-FTMS couples a resolving quadrupole and a collision cell (q) to allow for both mass selection and CID fragmentation prior to the FT-ICR cell for high-resolution detection. CID, Electron Capture Dissociation (ECD) and infrared multiphoton dissociation (IRMPD) MS/MS capabilities (see Figure 3) are still possible within the FT-ICR cell itself. The second hybrid links the LTQ with an FT-ICR cell. This hybrid allows the additional flexibility of a fully intact detection system within the LTQ itself and can increase duty cycle by performing partially parallel experiments in the two analysers (e.g. multiple tandem mass spectrometry scans in the LTQ during a single high-resolution MS scan in the FT-ICR cell). As with the Qq-FTMS, MS/MS experiments are possible within the FT-ICR cell via ECD and IRMPD. The final hybrid also uses the LTQ as its front-end but links it to the Orbitrap electrostatic trapping (rather than magnetic) FT-based mass analyser. This analyser obviates the need for high-field-strength magnets and their cryogenic requirements. Unlike the FT-ICR analysers, it is not possible to perform MS/MS in the Orbitrap itself.
alternative for many applications due to the lack of need for a large superconducting magnet and concomitant requirements for liquid helium and nitrogen.

While not mass analysers per se, ion mobility separations are another potential unit of hybridization that provides some tantalizing possibilities. Ion mobility allows for separation as a function of how long ions take to traverse a gas-filled area based primarily on their shape and charge. Even ions of identical $m/z$ may traverse this section at different rates and thus effect an additional dimension of separation. Inclusion of ion mobility drift tubes as a dimension of separation before ions enter the mass spectrometer provides the possibility of increasing the dynamic range of analysis while at the same time decreasing the total time required for analysis [38], especially for the most complex mixtures. Clemmer and colleagues have developed an ion mobility technique in which a split-field drift tube is used. The first portion of the drift tube is maintained at low-field such that ions are allowed to pass through, while the second portion of the tube is alternated between high field and low field [39]. Application of high field to the second portion of the drift tube causes increased collision of the ions, and thus yields fragmentation ions which can then be analysed simultaneously with the parent ion. The alternate low field in the second portion of the drift tube allows for passage of ions into the mass analyser (TOF) for analysis. This technique, described extensively elsewhere [38, 39], has been successfully used to profile the proteomes of *Drosophila melanogaster* embryos and adult heads [40, 41]. Ion mobility principles can also be employed to focus on a specific mobility in a way that is analogous to a quadrupole mass filter. One such method, high-field asymmetric waveform ion mobility spectrometry, or FAIMS, acts as a filter to separate gas-phase ions independent of their $m/z$ ratio by the application of a compensation voltage (CV) that will allow only certain ions to pass through the FAIMS device. FAIMS coupled with ESI and MS (ESI–FAIMS–MS) has been applied to the separation of ions from a tryptic digest of pig haemoglobin [42], and from a tryptic digest of enolase 1 [43]. ESI–FAIMS–MS analysis in both instances demonstrated an increased signal-to-noise ratio such that low-abundance peptides could be detected. Further, an ability to distinguish between isobaric peptides was also shown with use of ESI–FAIMS–MS.

### Putting together the pieces—common mass spectrometry hybrids

It is a daunting challenge to discuss all the possible combinations of ionization sources and mass analysers that could be useful for proteomics. We will try to highlight some of the ‘grizzled veterans’ and ‘up-and-coming superstars’ that are the primary tools of the proteomicist. For ease of discussion, they will be divided into general classes based on their ionization source, MALDI or ESI. However, with the modularity of new instruments and the ability to rapidly ‘bolt on’ new ionization sources, these distinctions become a bit blurred in practice.

For 2D–PAGE-based proteomics workflows, one of the most used mass spectrometers is the MALDI-TOF (see Figure 1). For these experiments a particular protein ‘spot’ is excised, digested with trypsin and the resulting peptides applied to a MALDI target plate with the appropriate matrix to facilitate ionization. As mentioned previously, masses of the resulting peptides allow for comparison of predicted peptide digestion patterns of proteins in the database. While post-source decay (PSD) allows for limited MS/MS capabilities, other instruments are required for robust MS/MS. One of these, the TOF–TOF, places a gas-filled collisional chamber between two TOF analysers (Figure 1). Dedicated MALDI versions of QqTOFs (see the following text) are also available to improve capabilities to perform MS/MS. With the availability of sources that can easily be switched between ESI and MALDI most of the popular platforms also have the potential to use MALDI including: MALDI-QIT, MALDI-LIT and even MALDI-FTMS. Because MALDI primarily produces singly charged ions, there can be some limitations in the molecular mass of the ions (depending on the mass analyser) and in the predictability of MS/MS fragmentation (depending on the form of MS/MS being performed).

ESI is another stalwart of proteomics workflows. It has been especially useful for the aforementioned shotgun analysis (in which a mixture of proteins is digested directly with limited or no prior separation at the protein level). A powerful leveraging of this technique couples the peptide separation step directly to the ESI source followed by the subsequent collection of MS/MS spectra. While these types of experiments were originally performed on triple quadrupole instruments (QqQ), the real workhorse instruments for shotgun analysis are the ESI–QITs and LITs (Figure 1). However, the additional
resolution and mass accuracy of the QqTOFs have made them popular choices as well for these workflows (Figure 1). ESI is also the method of choice for top-down proteomics (i.e. measurements on undigested proteins) allowing for the ionization of extremely large proteins and even non-covalent protein complexes.

**Tandem mass spectrometry (MS/MS)**

One of the primary technological advances that has made possible, and subsequently practical, many of the powerful proteomics workflows are developments in the ability to perform MS/MS. As previously alluded, during MS/MS, a precursor ion of interest is isolated, activated, fragmented and the resulting product ions analysed. Fragmentation of precursor ions in MS/MS is accomplished either through application of ‘slow heating’ methods or of more recently developed gas-phase ion–ion reactions or ion–electron reactions.

‘Slow heating’ ion activation methods such as low-energy CID [44, 45] and infrared multiphoton dissociation (IRMPD) (reviewed more extensively in [44, 45]) involve slow additions of energy to the precursor ions such that the internal energy of the ions exceeds the energy needed for fragmentation [46] (see Figure 3A). The most commonly used slow heating fragmentation method is low-energy CID (also known as collisionally activated or aided dissociation, or CAD). This method is employed in QqQ, QqTOF, QIT, LIT and in the ICR cell of an FT-ICR [45].

In a CID experiment, precursor ions are isolated and subjected to collision with a neutral, inert gas. The energy imparted to the precursor ion upon collision is converted from kinetic energy to internal energy [45]. Multiple collisions result in an increase in internal energy of the precursor ion such that the peptide or protein backbone fragments, primarily at amide bonds, resulting primarily in b- and y-type fragment ions (Figure 3). Within ion traps and ICR cells, the collision is considered to be tandem in time because the initial isolation and subsequent activation and fragmentation steps occur in the same location. Tandem-in-space experiments are performed in QqQ, QqTOF, Qq-FTMS instruments. Precursor ions enter the first quadrupole, and selected ions are sent to the second quadrupole, where they are focused by the application of an RF-only voltage. The collision cell is filled with neutral gas, and ion activation results from multiple collisions of precursor ions with the gas. Resulting fragments are then measured in the downstream detector (e.g. TOF in the QqTOF instrument). Another instrument that performs a tandem-in-space MS/MS is the TOF–TOF. However, because the ions have higher energies during the fragmentation process, it produces somewhat different MS/MS fragmentation patterns and is considered a high-energy CID process (Figure 3b) [45].

Infrared multiphoton dissociation (IRMPD) is an alternative slow heating dissociation method involving non-resonant ion activation and subsequent dissociation via photon absorption. Historically used in analysis of small molecules [45], IRMPD is now becoming more popular in analysis of larger peptides and proteins [45, 46]. In an IRMPD experiment, ions are typically activated with a low-power CO2 laser of wavelength 10.6 μm [45]. This wavelength makes IRMPD particularly well-suited for analysis of phosphopeptides because the phosphate group preferentially absorbs at this wavelength and thus phosphopeptides can be selectively dissociated using IRMPD [47]. Since it is also a slow heating method, IRMPD gives nearly identical fragment ions to those seen with dissociation by CID [48] (Figure 3).

Complementary to the slow heating ion activation methods in MS/MS are the ion–electron and ion–ion reactions (Figure 3c). Electron capture dissociation (ECD) [49, 50] is based on ion–electron reactions in which the capture of electrons by a gaseous positive ion leads to fragmentation and neutralization of the positive ion [45]. When applied to multiply charged peptide or protein cations, ECD results in extensive cleavage of the backbone at N-Cβ bonds to yield c and z ions (Figure 3) [49]. ECD also preferentially cleaves disulfide bonds but leaves other PTMs such as sulfonation and gamma-carboxylation intact [51]. Ion residence time in ICR cells is typically greater than that in ion traps, making FT-ICR better suited to ECD. Additionally, ECD efficiency is greatest for electron energies less than 1 eV, energies difficult to achieve in an ion trap instrument [51] due to the strong RF potentials needed to trap ions. Because ECD provides complementary information to slow heating activation, ECD is an available option on all commercial FT-ICR instruments and has been applied in the structural characterization of numerous PTMs: O-linked glycosylation [48, 52], methionine oxidation, acylation [52] and phosphorylation.
A combination of ECD with vibrational slow heating methods, termed activation induced-ECD, or AI-ECD, was developed by McLafferty and co-workers [53] and gives a more complete sequence coverage and more thorough characterization of sites of PTMs. Machine design was evaluated by Marshall and colleagues for optimization of both ECD and IRMPD within an FT-ICR. It was determined that an on-axis electron source inside the magnet bore and an off-axis infrared laser resulted in optimum performance of both ECD and IRMPD [54].

Since ECD is not amenable to ion trap instruments, an analogous ion–ion technique was developed by Hunt and colleagues [55] and is an extension of earlier electron transfer reaction studies in nucleic acids done by McLuckey and colleagues [56–58], and in peptides done by Zubarev and colleagues [59]. Electron transfer dissociation (ETD) is an ion–ion reaction between singly charged anions and multiply charged peptides. The anions are allowed to interact with the multiply charged peptides and result in proton transfer with or without dissociation. Proton transfer results in charge reduction while dissociation leads to c- and z-fragment ions. ECD: used in FT-ICR instruments, ECD relies on interaction of precursor ions with low energy electrons introduced into the ICR cell. The capture of low energy electrons by the multiply charged protein or peptide precursor leads to neutralization and dissociation of the ion. Backbone cleavage occurs rapidly at or near the site of electron capture, resulting in c- and z-fragment ions as in ETD.
proceed. This serves to allow only the reaction of interest to occur. This is termed ‘parallel ion parking’ and was achieved by McLuckey and coworkers [61] upon application of a tailored waveform that ‘parks’ ions of all other than the reagent anion and the peptide cation of interest. Through the application of a single-frequency dipolar resonance voltage to the end caps, ions of a selected \( m/z \) were accelerated resulting in a reduced reaction rate. Further reduction of sequential ion–ion reactions was achieved by McLuckey and coworkers [60] upon application of a tailored waveform that ‘parks’ ions of all \( m/z \) other than the reagent anion and the peptide cation of interest. This is termed ‘parallel ion parking’ and serves to allow only the reaction of interest to proceed [60].

**Using the tools—global measurements of what, when, with whom and how much**

As mentioned previously, the ultimate goal of proteomics is to completely characterize the what, when, with whom and how much every protein is doing within a cell. We will now discuss some of the important work that is beginning to solve subsets of this grand challenge.

The classic proteomics experiment attempts to address the issue of which proteins and how much of those proteins are present in a sample and then how those levels change over time. 2D-PAGE still provides excellent ability to visualize these protein spots and to quantitate changes (reviewed in [62]). Protein identification can either be performed comprehensively or only on those subsets of proteins that are observed to change, and is typically based on excising spots, performing in-gel digestions and then analysing those peptides by mass spectrometry. Large format gels and/or narrow pH ranges in the first dimension increase the separative capacity of 2D-PAGE, but require a high degree of gel-to-gel reproducibility for robust results. A dye-based technique, fluorescence 2D differential in-gel electrophoresis (DIGE) provides a platform to compare two states on a single gel. DIGE-based workflows are sometimes being employed in concert with shotgun techniques, such as used in Yu et al. [63] for the discovery of pancreatic cancer biomarkers within serum, to give a more comprehensive proteomic readout.

Multidimensional peptide separations, such as the aforementioned MudPIT, have made possible the direct shotgun analysis of cellular proteomes. Initial experiments in yeast allowed for the identification of 1484 proteins including those that would be difficult, if not impossible, to detect by gel-based strategies [9, 10]. Added separative steps pushed the numbers of identifiable proteins even higher [31]. Another variation uses 1D-PAGE, in-gel digestion of a ladder of gel slices and subsequent LC–MS/MS analysis of the peptides [64]. More recently, shotgun–proteomics-based techniques are moving out of model organisms into even more complex systems such as host–pathogen interactions for the malaria parasite [65, 66] and simple communities found in the wild such as the acid mine drainage microbial communities [67]. A variant shotgun–bottom-up approach uses accurate peptide masses (via FTMS) and HPLC retention times that have been cross-validated against a peptide MS/MS to rapidly survey a proteome without the need for the time-consuming acquisition of MS/MS spectra every experiment. These validated accurate mass and time (vAMT) tags are being used by Smith and colleagues [68–70] to rapidly screen through proteomic samples.

While not able to access the breadth of the proteome, top-down techniques are becoming increasingly routine. For instance, Kelleher et al. [71] were able to demonstrate the confident identification of 72 gene products from the thermophilic bacterium, *Methanococcus jannaschii*, using 2D protein fractionation followed by high-resolution MS/MS on a custom quadrupole-FTMS instrument. Several PTMs (methylation, acetylation and disulfide bonds) were also identified. Combination of top-down and bottom-up techniques such as those used by Verberkmoes et al. [72] identified a total of 868 proteins from *Shewanella oneidensis*, with accurate mass measurements made via top-down techniques on 70 intact proteins. Use of both techniques took advantage of the strengths of each and allowed not only for identification of proteins but also elucidation of PTMs [72].

One of the challenges associated with the shotgun proteomic experiments is that the results are at best only limitedly quantifiable [73]. However, much more precise quantitation is possible through two general mechanisms of incorporating stable isotope labels onto the peptides. These differential labelling
strategies provide an internally controlled measurement that does not require the run-to-run precision that non-labelling strategies require. One class of techniques is via chemical labelling where proteins or peptides from two (or more) states are labelled with chemically identical but isotopically distinct tags. One such approach uses the isotope coded affinity tag (ICAT) reagents, to differentially label cysteine residues in the samples [74]. Because the technique can only be used to quantitate cysteine containing peptides, the scope of quantitation is somewhat limited. Another popular labelling strategy involves performing the trypsin digestion in the presence of limited. Another popular labelling strategy involves peptides, the scope of quantitation is somewhat can only be used to quantitate cysteine containing residues in the samples [74]. Because the technique can only be used to quantitate cysteine containing peptides, the scope of quantitation is somewhat limited. Another popular labelling strategy involves performing the trypsin digestion in the presence of either $^{16}\text{O}$ or $^{18}\text{O}$ water [75, 76]. Thus, every labelled peptide is shifted by 4 Da. Another class of technique for differentially labelling samples is to introduce a differential isotope into the growth media or food for the organism. While demonstrations of this strategy have been performed in unicellular organisms [77–79] and tissue culture [80–82], such techniques are even applicable to metazoan systems including fruit flies and round worms [83]. However, probably the most impressive demonstration of this technique showed that it could be adapted for mice and rats [84]. More technically challenging (and impossible for some samples) in vivo labelling strategies have the potential for much greater precision of quantitation since the mixing of samples can take place much further upstream in the workflow and thus minimizing variations during sample preparation. Finally, absolute quantitation of peptides can be obtained by spiking in known amounts of isotopically labelled peptide standards into the protein–peptide mixture. This common practice in small molecule quantitation was originally described by Barr et al. [85] for proteins and has more recently taken on the moniker of AQUA (Absolute QUAntitation) [86].

A key component to understanding protein function is to determine with which other proteins within a cell that it interacts to perform its function. For years, biochemical isolation based on activity and subsequent identification of protein complexes has been a key experimental strategy to characterize cellular processes. The advent of molecular genetic techniques allowing for the rapid engineering and expression of proteins has opened up new possibilities. One of these is engineering protein constructs that express ‘tags’ that allow for their selective isolation. These are regularly used for in depth studies of particular protein complexes and when coupled with either gel-based mass spectrometry analysis or direct shotgun analysis have proven to be incredibly sensitive and powerful. Such strategies have also been employed for systematic studies of protein complexes in yeast [87, 88] and bacteria [89]. Coupling bottom-up and top-down analysis as described for the analysis of R. palustris ribosomes [90] will become increasingly more common. Further investigation of structure and composition of large complexes will be facilitated by improvement of methods that allow mass spectrometry of intact non-covalent complexes. For example, Robinson and Rostom were able to detect an intact 800 kDa GroEL chaperonin complex consisting of 14 subunits in a QqTOF instrument [91].

The subcellular localization of a protein also has profound impact on its function. While microscopic localization based on fluorescence detection of tagged proteins or through the use of antibodies will continue to provide the highest spatial resolution, mass spectrometry-based techniques can also be used for crude levels of characterization. These techniques are based on cell biology fractionations of particular subcellular organelles or broader classes of proteins. By characterizing the complete protein complement of these fractionated samples, one can get a ‘grocery list’ of proteins present within a particular cellular location. Several excellent examples include the characterization of the Golgi apparatus from rat [92, 93], mitochondria from rice [94] and mouse [95] and the nucleolus from human [96] and Arabadopsis [97]. It is in some of these characterizations that one begins to get overlap between supermolecular protein complexes and non-membrane-bound organelles.

Another aspect that bears on protein function is their regulation by PTMs. There are over 200 described protein PTMs that vary from incredibly well-studied (e.g. phosphorylation) to those whose role is poorly understood [98]. These modifications can impact the stability, interactions, biochemical activity and localization of a protein. Many techniques have been described to determine the presence and location of PTMs on purified proteins. A generalized method for characterizing PTMs involves an extension to the basic shotgun proteomics protocol in which additional parallel enzymatic digestions are performed to provide peptides that more confidently establish the presence and precise location of the PTM [99]. Other shotgun-based techniques specifically enrich only for those modified...
peptides. A good example of this is the use of immobilized metal affinity chromatography (IMAC) to specifically enrich for phosphopeptides [100, 101]. This enrichment step can be essential since many functionally important modifications are made to low-abundance proteins and, often, to relatively low stoichiometry. One problem with a shotgun approach is that, during the digestion process, peptides from different isoforms of the protein are pooled together. This can lead to a blurring of which modifications were derived from a specific molecular form of the protein. A top-down protocol has the ability to define the array of modifications present on the different forms of the protein. Coupling bottom-up and top-down protocols has the potential to leverage the strengths of both.

CONCLUSIONS
The analysis of the entire cellular complement of proteins is a truly daunting challenge. This is especially difficult when one also attempts to measure other aspects such as interacting partners and PTM states. Finally, these measurements must be made across a dynamic range of 10^6 or greater with a sensitivity that requires only a small amount of starting material. Over the past several years, significant progress has been made towards reaching these long-term goals. During this time advances in sample preparation, front-end separations, mass spectrometers and algorithms have played key roles in this progress. Moreover, because of their interdependence, the proper optimization and integration of the various steps in a proteomic workflow has proven the difference between an interesting technical strategy and something that is truly useful for biological analysis. Previous work has also demonstrated that there is no ‘one size fits all’ method, and a complete picture will require the integration of multiple analytical strategies. It is also clear that advances in mass spectrometry will continue to catalyse developments in the field. New powerful mass spectrometers, in terms of their dynamic range and mass accuracy, are being equipped with an even greater repertoire of MS/MS capabilities. As companies begin to provide these new platforms as robust easy-to-use instruments, we should continue to see exciting developments in the field that provide ever more useful and specific information to the biologist.

SUMMARY STATEMENTS
Proteomic workflows involve an integrated series of analytical steps including: sample preparation, protein–peptide separations, mass spectrometry and data analysis.

Technology developments in mass spectrometry have been key to many advances in the proteomics field.

These include improved mass analysers, how those mass analysers are combined and ultimately how easy the systems are to operate.

Combinations of mass analysers also allow the use of new forms of MS/MS and provide more powerful tools to probe the structure of proteins and peptides.

Acknowledgments
The authors acknowledge funding from the US Department of Energy’s Office of Biological and Environmental Sciences through the Genomics: GTL Center for Molecular and Cellular Systems and support by the ORNL Laboratory Director’s Research and Development Program, US Department of Energy, under contract No. DE-AC05-00OR22725 with Oak Ridge National Laboratory, managed and operated by UT-Battelle, LLC. K.-M. acknowledges support from the University of Tennessee (Knoxville)-ORNL Graduate School of Genome Science and Technology. We would like to thank Doug Goeringer for critical reading of this manuscript.

References


