Tandem Mass Spectrometry for Peptide and Protein Sequence Analysis

Joshua J. Coon[1], John E.P. Syka[1,2], Jeffrey Shabanowitz[1], and Donald F. Hunt[1]
[1]University of Virginia, Charlottesville, VA and [2]Thermo Electron, San Jose, CA, USA

Proteins are involved in nearly every aspect of cellular function. In fact, the characterization of proteins has become such a significant part of modern biology, it has inspired a new discipline: proteomics—the classification of the protein complement expressed by the genome of an organism. Technology development has driven, and continues to drive, rapid evolution in this field. Seventeen years ago innovations in biomolecule ionization [electrospray (1) and matrix-assisted laser desorption/ionization (2), ESI and MALDI, respectively] placed mass spectrometry (MS) at the forefront of this emerging discipline. Although biological MS was pursued prior to this work, these techniques are generally credited for expanding the field substantially. With these tools, mass spectrometrists could for the first time, easily and robustly, convert condensed phase peptides or even whole proteins to intact gas phase ions.

Protein MS Methodology

Over the past few decades many MS-based protein identification strategies have emerged. Peptide mass fingerprinting was among the first to gain widespread popularity. In this method, an isolated protein is digested enzymatically, and the resulting peptide molecular weights are measured [typically with a MALDI time-of-flight (MALDI-TOF) mass spectrometer]. The measured peptide masses are then searched against a database of proteins that have been digested in silico. Technical developments in chromatography and MS instrumentation have made a more recent strategy, entitled shotgun proteomics, quite popular (3). Here an entire proteome is enzymatically digested, chromatographically separated, and interrogated with tandem mass spectrometry (MS/MS) (see below). Finally today, a small but growing number of mass spectrometrists are advancing yet another strategy that seeks to directly analyze intact proteins—top-down proteomics (4). No matter which of these methods are used, technical developments in mass spectrometer technology continue to advance existing and generate new proteomic strategies.

In a typical shotgun proteomics experiment, a cell lysate, containing as many as several thousand proteins, is analyzed (Figure 1A). The sample is digested with a proteolytic enzyme resulting in a complex mixture of peptides (approximately 40 peptides/protein; Figure 1B). Next, the digested sample is chromatographically separated (in one or multiple dimensions) and introduced to the mass spectrometer by means of a nanoflow high-performance liquid chromatography (nHPLC; approximately 50 nL/min) column integrated directly to an ESI source on the mass spectrometer (Figure 1, C and D). The ESI source converts condensed phase ions, eluting from the HPLC column, to multiply protonated molecules (cations; the number of attached protons is proportional to the number of basic residues contained in the peptide) in the gas phase—a requirement for MS analysis. The mass spectrometer first records the mass/charge (m/z) of each peptide ion (Figure 1E) and then selects the peptide ions individually to obtain sequence information via MS/MS (Figure 1F).

Tandem MS Collision-Activated Dissociation

MS/MS, the process of peptide ion fragmentation with subsequent m/z measurement, is typically induced by isolating the protonated peptide m/z of interest and subjecting it to several hundred collisions with rare gas atoms (5). This process, termed collision-activated dissociation (CAD), supplies sufficient internal energy to induce covalent bond breakage. In the gas phase, the amide bonds of the peptide backbone are typically the preferred sites of protonation. These protonated amide linkages are weakened and, upon collisional-activation, are favored for cleavage to create a series of homologous product ions. For example, the MS/MS spectrum in Figure 1F contains the consecutive peptide fragment ions PE, PEP, etc. Thus, the process aims to produce a collection of peptide fragment ions that differ in mass by a single amino acid, allowing one to read the amino acid sequence of the precursor peptide. Note, however, that production of this series relies on a random distribution of protonated amide bonds within the starting precursor peptide ion population (e.g.,

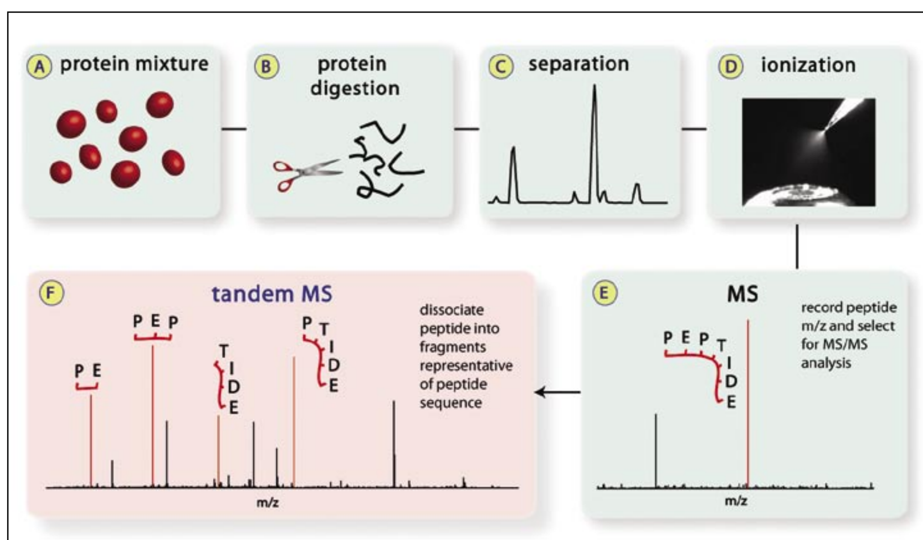

Figure 1. Sequence of events defining a contemporary mass spectrometry (MS)-based proteomics experiment. (A) The process begins with a mixture of proteins from any number of sources. (B) Enzymes are used to digest the proteins into peptides in the condensed phase. (C) Digested peptides are separated with some form of chromatography. (D) Peptides are individually ionized via electrospray ionization (ESI) and sampled into the mass spectrometer inlet. (E) The charged peptides are separated based on their mass/charge (m/z) within the mass spectrometer. (F) Abundant peptide cations are selected for fragmentation via tandem mass spectrometry (MS/MS). Following peptide ion dissociation the newly generated fragments are m/z analyzed to produce the MS/MS spectrum. *nESI photograph displayed in panel D courtesy of New Objective, Woburn, MA.

Sarepta Exhibit 1018, page 1

in an ion trap mass spectrometer the MS/MS experiment starts with approximately 10,000 peptide precursor ions).

During the collisional-activation of a selected peptide precursor ion, the energy is deposited on a relatively slow time scale (picoseconds). This imparted energy is redistributed throughout the peptide precursor ion and ultimately induces cleavage of the weakest bond(s). When peptides contain important posttranslational modifications (PTMs) (e.g., phosphorylation, glycosylation, sulfonation, etc.), the preferred dissociation pathways described above can change. In phosphorylated peptides, for example, the phosphate group competes with the amide bonds of the backbone as the preferred site of protonation (in the gas phase). Consequently, collisional-activation of these peptides readily displaces phosphoric acid from the peptide following nucleophilic attack of the protonated sidechain. Regardless of the mechanisms defining these preferential cleavages, the net result is often the same: CAD MS/MS spectra of PTM-containing peptides are frequently devoid of the peptide ion fragment m/z (those corresponding to the consecutive backbone cleavages), which are necessary for sequence identification. This limitation has, for a long time, precluded the widespread global characterization of protein PTMs by MS.

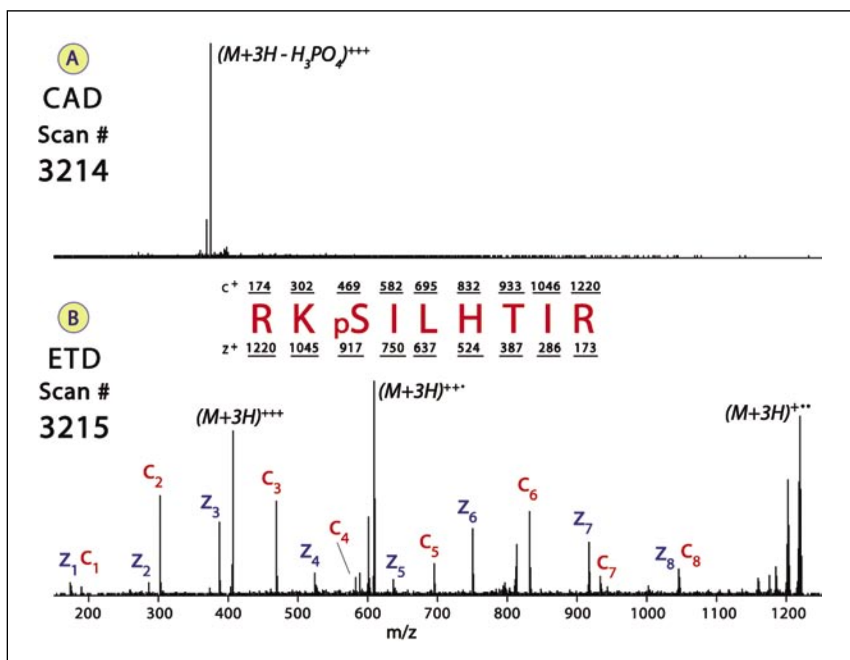

Figure 2. Tandem mass spectrometry (MS/MS) spectra obtained from a phosphopeptide eluted during a nanoflow high-performance liquid chromatography MS/MS (nHPLC-MS/MS) experiment. (A) The MS/MS spectrum produced following conventional collisional-activation. Note this spectrum is dominated by a single mass/charge (m/z) corresponding to loss of a phosphoric acid moiety. No peptide backbone cleavage is observed. Sequence identification is, therefore, impossible. (B) The MS/MS spectrum that is produced following electron transfer dissociation (ETD) fragmentation. Here, every single backbone cleavage product is observed. The sequence is easily assigned as RKpSILHTIR. Both panels display single-scan mass spectra. CAD, collision-activated dissociation.

Tandem MS Electron Capture Dissociation

An alternative method for peptide ion dissociation, electron capture dissociation (ECD), was introduced by McLafferty and coworkers 7 years ago (6). In that work, low energy electrons were reacted with peptide cations in the magnetic field of a Fourier transform ion cyclotron resonance MS (FT-ICR-MS). The reaction resulted in the attachment of electrons to the protonated peptides producing peptide cations containing an additional electron. The odd electron peptide then undergoes very rapid (femtoseconds) rearrangement with subsequent dissociation. Unlike the collision-activated process, ECD does not cleave chemical modifications from the peptide, but rather induces random breakage of the peptide backbone—cleavage that is indifferent to either peptide sequence or length. For several years this remarkable development was exclusively available to those in possession of the most expensive mass spectrometer technology, FT-ICR-MS. This is mainly because the more widely accessible instruments, those that use radio frequency (RF) fields to confine peptide ions (e.g., ion trap or quadrupole TOF), are not capable of simultaneously storing peptide cations and electrons.

Obviously, beyond its proven efficacy, the ECD fragmentation process held an enormous amount of latent analytical utility, if only it were accessible to the entire proteomics community. With this in mind, our laboratory recently began developing a new ion dissociation method to translate the unique features of ECD to the common benchtop ion trap mass spectrometers routinely used in proteomics research. Our plan, however, did not involve reacting electrons with peptide cations as in ECD. Instead, we chose to use negatively charged ions (anions) as vehicles for electron delivery (see Equation 1).

$$(\text{Peptide} + 3\text{H})^{3+} + \text{Anion}^{-\bullet} \rightarrow (\text{Peptide} + 3\text{H})^{2+\bullet} + \text{Anion} \quad \text{[Eq. 1]}$$

Though the RF fields used by the ion trap cannot confine both cations and electrons, the device can simultaneously contain cations and anions. Given the appropriate anion, the reaction should proceed as described in Equation 1 to donate an electron to the peptide. Subsequently, the peptide would contain an extra electron, and that should induce peptide backbone fragmentation, just as in ECD. Further, gas phase peptide cations and small organic anions react rapidly (ion-ion reactions, milliseconds time scale)—reactions whose duration and timing are easily controlled.

Ion-ion reactions had been studied for nearly a decade in 3-D ion traps, but the electron transfer reaction described in Equation 1 remained elusive (7). Due to a number of limitations with conventional 3-D ion traps, our efforts centered around construction of a novel ion-ion device: the quadrupole linear ion trap (QLT) (8,9). This mass spectrometer consists of a set of four hyperbolic rods (electrodes) in which RF fields are used to confine peptide cations along the central axis. The commercial version of this instrument (Finnigan LTQ™; Thermo Electron, San Jose, CA, USA) places an ESI ion source at one end of the device (for peptide ion generation) and leaves the other end unoccupied. For our experiments we placed an additional ion source on the vacant end of the QLT. This way peptide cations were generated from the front, as normal, and anions from various compounds could be created and injected into the QLT from the back.

Tandem MS Electron Transfer Dissociation

Reactions of numerous candidate anions with peptide cations revealed polyaromatic hydrocarbons as one class of compounds that were particularly effective at promoting the electron trans-

Sarepta Exhibit 1018, page 2

fer reaction described in Equation 1. And just as in ECD, labile PTMs remain intact, while peptide backbone bonds are cleaved with relatively little concern to peptide sequence, charge, or size. Unlike ECD, electron transfer dissociation (ETD) can be performed with lower-cost benchtop mass spectrometers on a time scale that permits coupling with online chromatographic separations. Figure 2 displays MS/MS spectra obtained from a phosphopeptide eluted during an nHPLC-MS/MS experiment. To provide a comparison, both CAD and ETD fragmentation methods were sequentially applied to the eluting phosphopeptide. Figure 2A displays the first MS/MS scan acquired using the conventional collisional-activation; this spectrum is characterized by a single m/z that corresponds to the loss of H_3PO_4. Sequence assignment with this spectrum is therefore impossible. The very next scan utilized an ETD reaction involving a polyaromatic hydrocarbon anion and generated the spectrum shown in Figure 2B. Upon ETD fragmentation, every possible backbone fragment is observed, defining the unknown peptide sequence as RKpSILHTIR. So far, we have applied ETD to globally characterize protein phosphorylation, but the method also preserves other CAD labile modifications such as glycosylation and sulfonation (unpublished data).

With each passing day the demand for comprehensive protein characterization technologies grows stronger. Mass spectrometry has been, and remains, at the core of these technologies. Successful mass spectrometric sequence analysis, however, hinges on the ability to systematically dissect peptide backbone bonds in a deliberate and predictable manner (MS/MS). The conventional peptide fragmentation tool, CAD, fails in this regard if the peptide is too long (approximately >20 residues) or contains either labile PTMs or multiple basic residues. ETD, on the other hand, is indifferent to peptide sequence, the presence of PTMs, or even peptide length.

Because of these unique features and its widespread accessibility to common benchtop mass spectrometers, we believe ETD will play an important role in the developing field of protein MS. Obviously ETD's initial distinction will be founded in its ability for global protein PTM mapping (e.g., identifying phosphorylation, glycosylation, sulfonation, etc.) sites of chemical modification. Alone this ability is of considerable value; however, ETD opens other analytical doors that may restructure the proteomics landscape. A unique characteristic ETD shares with ECD is the capacity to dissociate nearly every backbone bond of large peptides and even intact proteins. Note that while intact proteins can be dissociated with CAD, this process routinely produces only a few backbone cleavages making sequence identification challenging. Combined with other ion-ion techniques, ETD is likely to deliver top-down proteomics (analyzing intact proteins rather than digested protein peptides) to low-cost, benchtop mass spectrometers (unpublished data). Further, the distinct and predictable features of ETD fragmentation (i.e., near complete backbone fragmentation that is indifferent to peptide sequence) are likely to drive the development of new peptide sequencing computer algorithms—algorithms that may relieve the need for predicted protein databases for spectral correlation.

Acknowledgments

The authors acknowledge the National Institutes of Health grant nos. GM37537 and AI33993 (D.F.H.), RR018688 (J.J.C.), the National Science Foundation grant no. MCB-0209793 (D.F.H.), and Thermo Electron for their support, specifically Jae Schwartz, George Stafford, and Scott Quarmby. We also thank Gary Valaskovic (New Objective) for providing the nESI photograph displayed in Figure 1 and An Chi for preparing the sample from which Figure 2 was derived.

References

1. **Fenn, J.B., M. Mann, C.K. Meng, S.F. Wong, and C.M. Whitehouse.** 1989. Electrospray ionization for mass-spectrometry of large biomolecules. Science *246*:64-71.
2. **Karas, M. and F. Hillenkamp.** 1988. Laser desorption ionization of proteins with molecular masses exceeding 10000 daltons. Anal. Chem. *60*:2299-2301.
3. **Wolters, D.A., M.P. Washburn, and J.R. Yates.** 2001. An automated multidimensional protein identification technology for shotgun proteomics. Anal. Chem. *73*:5683-5690.
4. **Kelleher, N.L.** 2004. Top-down proteomics. Anal. Chem. *76*:196a-203a.
5. **Hunt, D.F., J.R. Yates, J. Shabanowitz, S. Winston, and C.R. Hauer.** 1986. Protein sequencing by tandem mass-spectrometry. Proc. Natl. Acad. Sci. USA *83*:6233-6237.
6. **Zubarev, R.A., N.L. Kelleher, and F.W. McLafferty.** 1998. Electron capture dissociation of multiply charged protein cations. A nonergodic process. J. Am. Chem. Soc. *120*:3265-3266.
7. **McLuckey, S.A. and J.L. Stephenson.** 1998. Ion ion chemistry of high-mass multiply charged ions. Mass Spectrom. Rev. *17*:369-407.
8. **Syka, J.E.P., J.J. Coon, M.J. Schroeder, J. Shabanowitz, and D.F. Hunt.** 2004. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. Proc. Natl. Acad. Sci. USA *101*:9528-9533.
9. **Coon, J.J., J.E.P. Syka, J.C. Schwartz, J. Shabanowitz, and D.F. Hunt.** 2004. Anion dependence in the partitioning between proton and electron transfer in ion/ion reactions. Int. J. Mass Spectrom. *236*:33-42.

Sarepta Exhibit 1018, page 3