PROTEOMICS BY FTICR MASS SPECTROMETRY: TOP DOWN AND BOTTOM UP

Bogdan Bogdanov and Richard D. Smith*

Biological Sciences Division and Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA 99352

This review provides a broad overview of recent Fourier transform ion cyclotron resonance (FTICR) applications and technological developments relevant to the field of proteomics. Both the "bottom up" (peptide level) and "top down" (intact protein level) approaches are discussed and illustrated with examples. "Bottom up" topics include peptide fragmentation, the accurate mass and time (AMT) tag approach and dynamic range extension technology, aspects of quantitative proteomics measurements, post-translational modifications, and developments in FTICR operation software focused on peptide and protein identification. Topics related to the "top down" approach include various aspects of high mass measurements, protein tandem mass spectrometry, methods for the study of protein conformations, and protein complexes as well as advanced technologies that may become of practical utility in the coming years. Finally, early examples of the integration of both FTICR approaches to bio-medical proteomics applications are presented, along with an outlook for future directions. © 2004 Wiley Periodicals, Inc., Mass Spec Rev 24:168–200, 2005

Keywords: *FTICR; peptides; proteins; fragmentation; chromatography*

I. INTRODUCTION

Fourier transform ion cyclotron resonance (FTICR) mass spectrometry (MS) (Marshall, Hendrickson, & Jackson, 1998) provides the highest mass resolution, mass resolving power, mass accuracy, and sensitivity of present MS technologies and is an increasingly useful technique in biological and biochemical research, despite its relatively high purchase, operational, and maintenance costs (Marshall, 2000). Many technological advances over the past 10 years have made, and anticipated developments will continue to make, FTICR a powerful technology for addressing important challenges currently faced in biology, biochemistry, and medicine. FTICR manufacturers are increasingly and more rapidly incorporating improvements into their products based upon developments from the academic/research community, thus enabling a broader scientific community to have access to complex instrumentation with tremendous potential. This trend is expected to accelerate as costs decrease in response to increases in demand (i.e., sales) and competition.

One field in particular that has benefited from, and also driven, FTICR developments is proteomics. Proteomics may be defined as the direct qualitative and quantitative analysis of the full complement or subset of the proteins present in an organism, tissue, or cell under a given set of physiological or environmental conditions. The rapidly increasing number of genomes sequenced cannot provide information as to which proteins may be present or how they are modified in specific situations. MS, combined with separation techniques, such as liquid chromatography (LC) or capillary electrophoresis (CE) to better address the complexity of most biological samples, has become pre-eminent for proteomics analyses.

This review provides an extensive summary of advances in FTICR during the period 1998–2002 as applied to the field of proteomics. These advances are discussed in relation to approaches used for applications, i.e., peptide (bottom up) and protein (top down) analyses, so researchers with experience in MS and interested in moving into proteomics, as well as biologists or biochemists without any prior experience with FTICR will have a broad overview of the status of the field and what may be possible in the future. Many excellent reviews are currently available on FTICR, and interested readers are referred to them for additional information (Holliman, Rempel, & Gross, 1994; McLafferty, 1994; Dienes et al., 1996; Marshall, 1996; Marshall & Guan, 1996; Green & Lebrilla, 1997; Marshall, Hendrickson, & Jackson, 1998).

II. PEPTIDE IDENTIFICATION: THE "BOTTOM UP" APPROACH

"Bottom up" strategies to proteome analysis involve cleaving the protein into peptide fragments that are smaller but still sufficiently distinctive to allow protein identification. Two generally applicable "bottom up" approaches for protein identification include one based upon mass measurements for a set of peptide digestion products from the parent protein and the other based on MS fragmentation (MS/MS) of one or more of these peptides. The first approach is referred to as peptide mass fingerprinting. A set of peptide fragments unique to each protein is created and their masses used as a "fingerprint" to identify the original protein. The peptide mass fingerprinting approach has been broadly

applied with conventional MS instrumentation, but its throughput is limited because it requires prior isolation of a single protein (or a simple mixture). The second approach to protein identification is based on the information from dissociation (e.g., using collisional activation or some other energy deposition process) of one or more polypeptides that have typically been isolated by the first stage of the MS analysis (Yates, McCormack, & Eng, 1996; McCormack et al., 1997). This approach, which forms the basis for further approaches discussed below, greatly reduces the number of polypeptides needed for confident protein identification with conventional instrumentation. The bottom up approaches generally do not provide information on the entire protein sequence, and similarly suffer from the fact that peptides from many possible forms of the parent gene products (arising from partial degradation and various covalent modifications) are generally indistinguishable in the absence of additional information (e.g., the MW of the parent protein).

A. Peptide Fragmentation

Peptide fragmentation using various dissociation techniques has been studied extensively for the purposes of peptide and subsequent protein identification. The most common peptide fragmentation methods used with FTICR instrumentation are sustained off-resonance irradiation (SORI) collision-induced dissociation (CID) (Gauthier, Trautman, & Jaconson, 1991), infrared multi-photon dissociation (IRMPD), both with a CO_2 laser (Little et al., 1994) and blackbody radiation (Price, Schnier, & Williams, 1996), and (much less frequently) surface-induced dissociation (SID) (Ijames & Wilkins, 1990; Chorush et al., 1995; Laskin, Denisov, & Futrell, 2000, 2001). Peptides primarily fragment to form b- and y-type ions for the slower multi-step activation processes (all of the above except for CID), and peptide fragmentation efficiency and relative fragment ion abundances are highly dependent on the peptide amino acid sequence. Compared to, e.g., peptide CID using quadrupole-time-of-flight (Q-TOF) or quadrupole ion trap (QIT) instruments, CID conducted in the FTICR cell is generally slower only because of the time required to re-establish high vacuum before excitation and detection can be performed. A major trend is to perform CID external to the FTICR cell, generally in an rf-multipole device (Sannes-Lowry & Hofstadler, 2000; Belov et al., 2001a), and where the results will be comparable to the fragmentation observed with more conventional instrumentation.

For the purpose of developing more controllable and predictable fragmentation processes, Lee et al. investigated fragmentation of singly and doubly sodiated peptides generated by matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI) using SORI (Lee, Kim, & Beauchamp, 1998). They observed that sodiated peptides exhibited highly selective cleavage at aspartic acid residues. Peptides that lacked acidic residues underwent sequential cleavage from the C-terminus. They proposed a salt-bridge chemistry mechanism, in which the sodium ion stabilizes the salt-bridge intermediate formed by proton transfer from the aspartic acid residue to the adjacent amide nitrogen. They suggested that this new fragmentation might also allow the possibility of mapping aspartic acid residues in proteins.

Most of the peptide fragmentation studies to date with FTICR have used positively charged ions, yet many biologically significant protein and peptides contain several acidic residues (e.g., glutamic and aspartic acid) or post-translational modifications (e.g., phosphates or sulfonates). Whereas these types of samples may be difficult to analyze by MS in the positive ion mode (e.g., low solution pH), analysis in the negative ion mode can yield stronger fragment ion signal intensities. Ewing and Cassady investigated the dissociation of multiply charged negative peptide ions that contained multiple acidic residues (Ewing & Cassady, 2000). Using SORI-CID, the dominant fragments observed were c, y, and internal fragments, although some a, b, x, and z ions were also formed. Many fragment ions were formed from cleavage adjacent to acidic residues (aspartic, glutamic, and cysteic acids), both N- and C-terminal, and at the C-terminus. In addition, loss of small neutral molecules (e.g., NH_3, H_2O, CO_2) and radicals (e.g., $CH_3^{\bullet}$) from both the parent and fragment ions was observed. Partial primary sequence information was also obtained for some of the systems investigated.

In general, acquiring a SORI-CID spectrum takes >2 s and even then, depending on the peptide amino acid sequence, problems can arise because of low fragmentation efficiency (relatively low signal-to-noise ratios (S/Ns) of the fragment ions) and selective fragmentation (i.e., of only the lowest energy pathways). Laskin et al. designed an FTICR equipped for in-cell SID at a high acquisition rate (0.3 s/spectrum), high trapping efficiency of fragment ions, well controlled kinetic energy of the parent ions, and high fragmentation efficiency (Laskin et al., 2002). The time scale makes SID potentially useful for high-throughput MS/MS experiments.

Costello and co-workers used a modified SORI-CID technique to fragment singly protonated peptide molecules generated by MALDI. Instead of applying a constant frequency offset (e.g., $\Delta f = 1,000$ Hz) for the whole mass-to-charge ratio (m/z)-range, a constant percentage offset (e.g., $\Delta f = 0.015 \times f$ Hz) was used; only the SORI amplitude (V_{pp}) was an adjustable parameter, and a value of 20 V produced efficient fragmentation for ions covering an m/z-range of 300–3,500 Thompson (Th) (Mirgorodskaya, O'Connor, & Costello, 2002).

Hunt and co-workers obtained a sub-femtomole detection limit using IRMPD MS/MS, although the practical "working" level achieved is ∼5 femtomole loaded on-column (Martin et al., 2000).

Electron capture dissociation (ECD), originally pioneered by McLafferty and co-workers and so far unique to FTICR, has been the focus of a large number of publications since its introduction in 1998 (Zubarev, Kelleher, & McLafferty, 1998). Initially applied to larger multiply charged protein molecules, ECD has also been applied to induce fragmentation in peptide ions that carry at least two charges. Capture of low-energy (<1 eV) electrons by $[M + nH]^{n+}$ ions yields $[M + nH]^{(n-1)+\bullet}$ ions. These ions will fragment rapidly and more randomly (because of the fast recombination of thermal electrons and various protonated sites in peptide or protein ions) than by, e.g., CID, via cleavage of the N-αC bonds, thereby producing N-terminal c-type ions and C-terminal $z^{\bullet}$-type ions. A unique feature of ECD in peptide and protein identification is that fragments can retain post-translational modifications that are labile under conventional activation processes, often allowing

unambiguous location of the modifications (see "Post-Translational Modifications"). Similarly, ECD has the ability to cleave disulfide bonds. Thus, ECD and CID often provide complementary information. High quality spectra have been obtained from 10 nM solutions of substance P and neurotensin (Håkansson et al., 2001), but the time scale needed to provide good S/N is generally still too long for coupling ECD with high performance LC (HPLC) or for use in a high-throughput environment. However, within these limitations, ECD has been used successfully with both HPLC and CE (see "Chromatographic Coupling").

B. The Accurate Mass and Time (AMT) Tag Approach to Peptide Identification
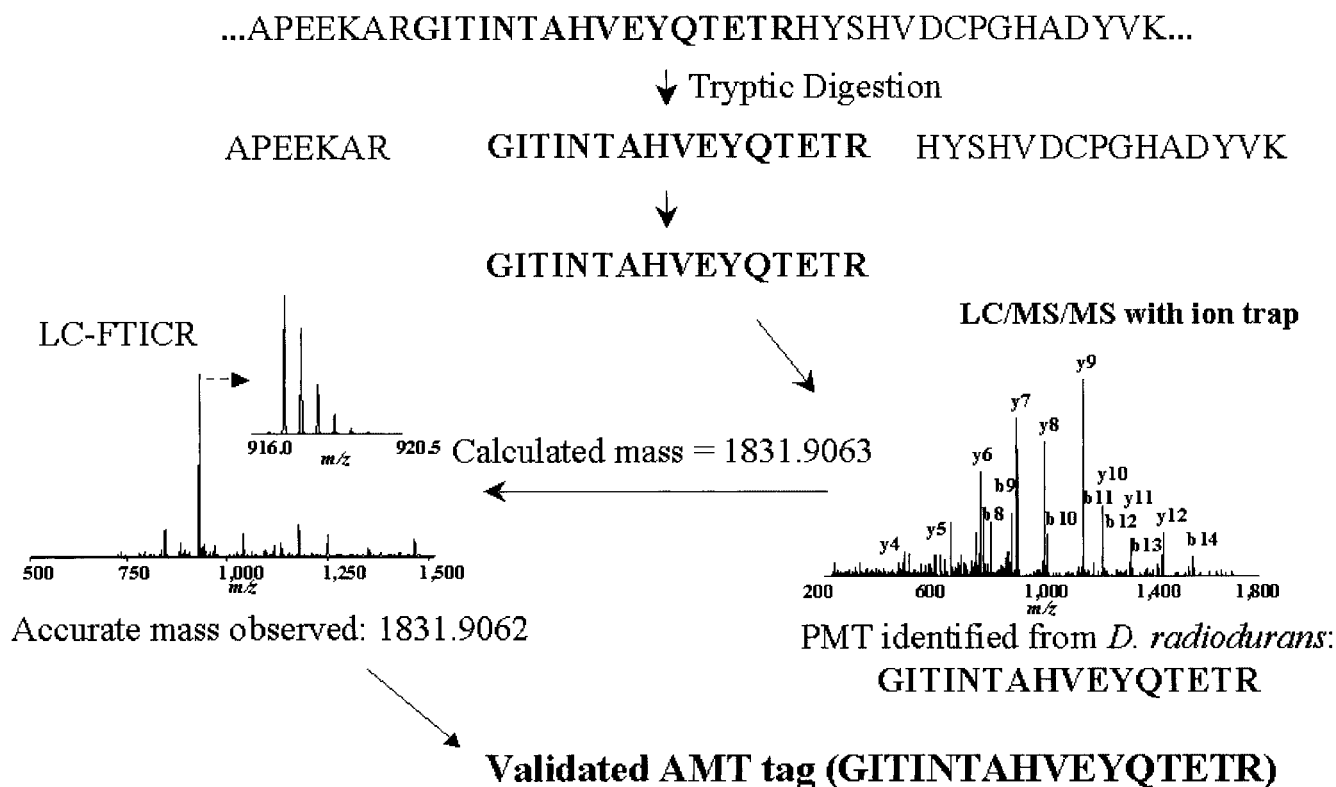
Over the past 3 years, our laboratory has focused on developing a bottom up proteomics approach that allows rapid characterization of large numbers of proteins directly from cell lystates, e.g., without the need for prior separation by 1- or 2-D gel electrophoresis. This global proteomics strategy involves the use of polypeptide "AMT" tags and is based on the assumption that the molecular masses of many peptides (or modified peptides) from an enzymatic digestion of proteins are effectively unique among all the possible peptides predicted from an annotated genome, provided they can be measured with high enough mass measurement accuracy (MMA). Thus, the approach shares similarities with peptide fingerprinting, when applied to complex mixtures of proteins. In practice, the specificity of the approach is enhanced by the use of separation times (e.g., elution times from LC separations). Because FTICR can provide sub parts-per-million (ppm) MMA, it is the most powerful tool for this approach, e.g., providing greater confidence in protein identifications. FTICR also provides enhanced sensitivity, allowing identification of lower abundance proteins and resulting in more comprehensive proteome coverage. Another key advantage of this approach is its much higher throughput compared to other MS methods because of the avoidance of routine tandem MS (MS/MS) measurements.

The AMT tag strategy first uses conventional MS/MS techniques to make an initial set of peptide identifications (and to provide corresponding LC elution times), referred to as peptide potential mass and time (PMT) tags. These PMT tags represent peptides from certain proteins identified in data-dependent MS/MS experiments performed on any MS platform, e.g., QIT mass spectrometers, and generally without the benefit of accurate mass information. Tryptic peptide mixtures are analyzed by HPLC-MS/MS and the resulting spectra are subsequently analyzed by a program such as SEQUEST (Eng, McCormack, & Yates, 1994) or MASCOT (Perkins et al., 1999) to identify peptides. Once identified, these peptides constitute a PMT tag database. A number of methods can be used to maximize the proteome coverage. In the case of microorganisms, we have demonstrated that growing cell cultures under many different conditions increased the numbers of detected proteins and performing many DC-MS/MS experiments allows a large percentage of the proteome to be identified and facilitates PMT tag creation. If samples from the same proteome are then analyzed using LC-FTICR MS and the PMT tag peptides are identified with some reasonable set of constraints (e.g., an MMA of ≤ 1 ppm and 5% normalized elution time (NET) tolerance), then the PMT tag is "validated" and promoted to an AMT tag (Fig. 1). Thus, these AMT tags serve as confident markers for their corresponding proteins in future experiments, providing an effective "look-up" table of masses and separation times. This approach eliminates the subsequent need for time-consuming MS/MS measurements, and consequently greatly increases analysis throughput. A critical aspect of this approach though is the initial time investment required to create the PMT tag database by exhaustive shotgun MS/MS measurements. The process can involve various pre-fractionation steps, etc., that are generally similar to the steps that would be employed to provide more extensive proteome coverage with conventional approaches using MS/MS. These steps are also needed to characterize the low level peptides that can be detected by LC-FTICR, but not by LC-MS/MS (in the absence of initial stages of sample fractionation). However, the end result of these additional efforts is that subsequent studies are enabled that do not require MS/MS.

Using the AMT tag approach >61% of the predicted proteome of *Deinococcus radiodurans* was identified with high confidence (Lipton et al., 2002). This percentage represented the broadest proteome coverage for any organism at that time and included 715 proteins previously annotated as either hypothetical or conserved hypothetical. In a single LC-FTICR MS analysis, the masses of ~1,500 AMT tags were typically detected, corresponding to ~700 open reading frames (ORFs) (depending on the cell culture conditions) and 15–20% of the *D. radiodurans* proteome (Fig. 2). Since the initial report, additional data processing/analysis improvements and additional analyses have increased overall proteome coverage to 83% (Paša-Tolić et al., 2002b). The AMT tag approach potentially requires more extensive separations (e.g., another dimension of chromatography or fractionation of some sort in applications with large genomes). Alternatively, one can set more stringent criteria for protein identification such as the detection of two or more species (AMT tags) that correspond to the same protein. However, in many cases, it would clearly be beneficial to increase the MMA further, as well as to use more precise separation time information (Palmblad et al., 2002a; Petritis et al., 2003). Since many genome sequences are incomplete or contain many undiscovered errors, many interesting challenges arise. For example, as the size of a proteome increases and when modifications exist or are possible, even higher MMA than routinely provided by FTICR to date may become desirable to further increase the throughput of MS based proteomics. Our LC-FTICR studies typically use MS resolutions of 50,000–100,000; a level almost always sufficient to resolve the set of peptides eluting at any single time point. Thus, MMA is generally more important in these studies than raw resolving power, and significant efforts are now being extended to better control the experimental factors that limit MMA (e.g., excessive space charge in the FTICR cell) and to apply improved calibration methods.

Some of these challenges can be met by improving the quality of separations. In fact, a large advantage of combining high-resolution separations with FTICR is that the need for MS resolution (as well as other FTICR attributes) generally become more modest. The advantage of combining information from separations along with MMA provides potential for enormously increased selectivity and sensitivity. A MMA of 1 ppm or better

...APEEKARGITINTAHVEYQTETRHYSHVDCPGHADYVK...

↓ Tryptic Digestion

APEEKAR   GITINTAHVEYQTETR   HYSHVDCPGHADYVK

↓

GITINTAHVEYQTETR

LC-FTICR

LC/MS/MS with ion trap

916.0   *m/z*   920.5 Calculated mass = 1831.9063

500   750   1,000   1,250   1,500
m/z

200   600   1,000   1,400   1,800
m/z

Accurate mass observed: 1831.9062

PMT identified from *D. radiodurans*:
GITINTAHVEYQTETR

Validated AMT tag (GITINTAHVEYQTETR)

FIGURE 1. Identification of an accurate mass and time (AMT) tag for elongation factor Tu (EF-Tu). In this example a tryptic peptide from EF-Tu was identified by tandem mass spectrometry (MS/MS) using an LCQ ion-trap mass spectrometer. The accurate mass of this PMT tag was calculated based on its sequence (i.e., 1,831.9063 Da) and its elution time recorded. In the liquid chromatography (LC)-Fourier transform ion cyclotron resonance (FTICR)-MS analysis of the same sample, a doubly charged peptide was observed at the same elution time, having a mass within 1 ppm (i.e., 1,831.9062 Da) of the calculated mass of this peptide. This peptide is then considered an AMT tag and uniquely identifies EF-Tu within the *Deinococcus radiodurans* proteome. (Adapted from Paša-Tolić et al., 2002b, with permission from John Wiley Ltd., copyright 2002.)

may also not be sufficient for some applications. For example two peptide isobars, RVMRGMR and RSHRGHR, were baseline resolved at a resolving power of ∼3,300,000 (Bossio & Marshall, 2002). The elemental compositions of these molecules differed by N_4O vs. S_2H_8 (0.00045 Da), which is less than the mass of an electron (0.00055 Da). The ultra high-resolution provided here by FTICR may be attributed to various factors: high magnetic field, large Penning trap diameter, digital quadrature hetrodyne detection, extended ion cooling by stepwise reduction of the cell trapping potential, low cell pressure, selected waveform inverse Fourier transform (SWIFT) isolation of a narrow *m/z* range, and multiple spectra summation. In another experiment, a mass resolving power of 552,000 and high MMAs of 0.032 and 0.053 ppm were obtained to unambiguously identify the $[DY(S-O_3H)MGWMDF-NH_2-2H]^{2-}$ and $[DY(PO_3H_2)MGWMDF-NH_2-2H]^{2-}$ peptides in a broadband mode (*m/z* = 400–1,400 Th) (He, Hendrickson, & Marshall, 2001). Using a 32 megabyte digitized time domain signal and co adding 64 spectra made the time scale of the experiment (3 min/spectrum) too slow for on-line chromatographic coupling, as is relevant to analysis of, e.g., proteolytic digests or other biological samples.

C. Dynamic Range Enhancement Applied to MS

The large variation of protein relative abundances generally of interest in proteomics (often >6 orders of magnitude) presents a major hurdle. The maximum dynamic range for a single FTICR mass spectrum is typically limited to approximately 10^3. Both FTICR sensitivity and dynamic range are typically limited by the charge capacity of either the external accumulation multipole (e.g., quadrupole, hexapole, octapole) or the FTICR cell. "Overfilling" of the external multipole may lead to biased accumulation or extensive ion activation and dissociation, and ions in excess of the capacity of the FTICR will lead to space charge effects (e.g., frequency shifts, peak coalescence). However, when combined with high-resolution LC separations, the overall dynamic range of measurements can exceed 10^6 (Shen et al., 2004) because of a combination of factors that include "overloading" of the separations (and wider peaks for more abundant species) and non-linear ionization efficiencies for more abundant species (Smith, Shen, & Tang, 2004). Because the dynamic range can be limited in any single spectrum, the dynamic range is often limited by the co-elution of more abundant
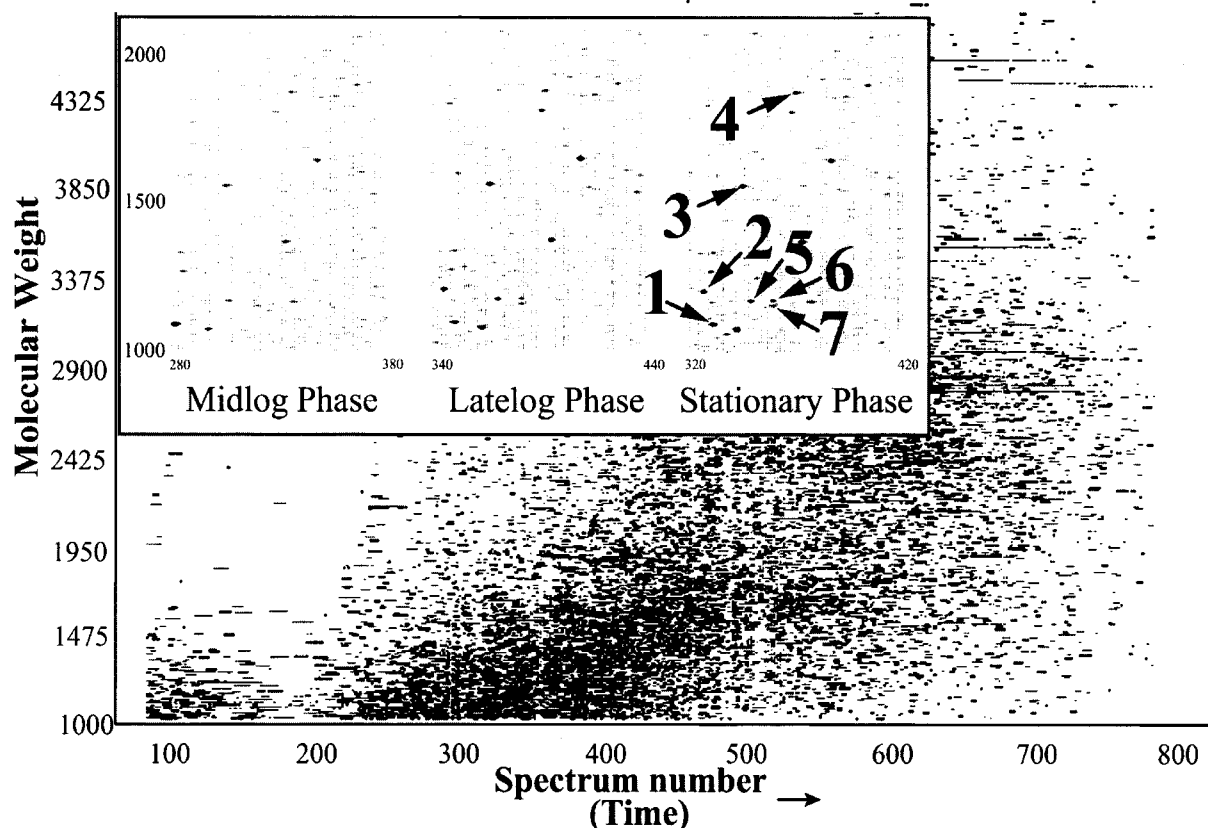
FIGURE 2. Two-dimensional display of a capillary LC-FTICR analysis in which ~50,000 putative polypeptides were detected from a tryptic digest of *D. radiodurans* proteins harvested in mid-log phase (OD600 0.3–0.4; 30°C). Inset shows portions of displays for peptides from *D. radiodurans* harvested in mid-log (**Left**), late-log (**Center**), and poststationary (**Right**) phases (spot size reflects relative abundance). Each individual spot corresponds to a peptide that can be identified along with the parent protein using AMT tags. Spot identifications: 1, DR1314 (hypothetical protein); 2, DR1790 (homolog of yellow royal jelly protein of insects); 3, DR1314 (hypothetical protein); 4, DR0309 (EF-Tu); 5, DR2577 (S-layer protein); 6, DR0989 (membrane protein); and 7, DR1124 (S-layer protein). (Adapted from Lipton et al., 2002, with permission from National Academy of Sciences, USA, copyright 2002.)

species. To increase the dynamic range in the presence of abundant components, the technique of "dynamic range enhancement applied to mass spectrometry" (DREAMS) was developed (Belov and Smith, 2001b). This technique was actually a variation on the dynamic range expansion method reported earlier (Bruce, Anderson, & Smith, 1996). Bruce et al. identified species at 1×10^{-8} M from a solution that also contained another component at 9×10^{-5} M. They also observed a new low abundance putative variant of bovine insulin, which was identified by MS/MS. The concept behind DREAMS is data-dependent ejection of high abundance ions in a selection quadrupole so that ions present at low abundance can be accumulated for longer times in an accumulation quadrupole. Ejection is accomplished by applying a secular radio frequency (RF) on two opposite rods of the selection quadrupole so that ions experience an oscillating motion perpendicular to the quadrupole axis.

A typical DREAMS FTICR experiment consists of two sequences. First a normal mass spectrum is recorded with a relatively short ion accumulation time (e.g., 100 ms). Based on this spectrum the data control software generates the DREAMS RF waveform (for up to 10 ions) and a DREAMS mass spectrum is recorded using a longer ion accumulation time (e.g., 1 s). Therefore, in a typical DREAMS experiment two ion chromatograms can be reconstructed. The total time required to acquire a "normal" and DREAMS mass spectra pair is 8–20 s, and is presently limited by data processing and transfer times, and DREAMS RF waveform generation. Initial experiments by Belov et al. used a segmented selection quadrupole operating at a Mathieu parameter (q) of ~0.45 (mass resolution ~100), and a 3.5 T FTICR to analyze a 0.10 mg/mL BSA tryptic digest solution by LC-ESI-DREAMS-FTICR MS (Belov, Anderson, & Smith, 2002). Compared to the normal experiment (60 unique tryptic fragments identified within 5 ppm MMA), the DREAMS experiment identified 22 additional tryptic fragments (~35% increase). Recently this approach was used to analyze a global tryptic digest of [14]N- and ^{15}N-labeled peptides from soluble
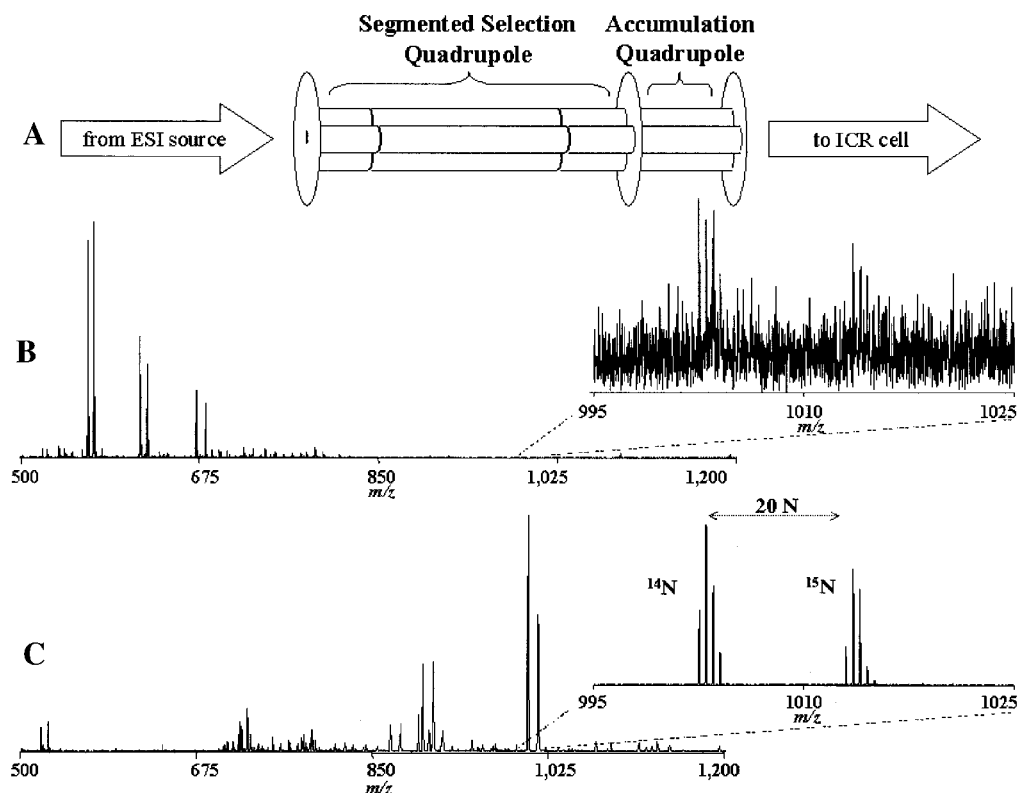
FIGURE 3. **A**: Illustration of DREAMS set-up—utilizing the information from the preceding spectrum, the most abundant species are ejected in the selection quadrupole, prior to extended accumulation in an external 2-D trap (i.e., accumulation quadrupole). **B**: Representative FTICR spectrum acquired during RPLC separation of the mixture of identical aliquots of a natural isotopic abundance and ^{15}N-labeled version of a mouse B16 cells obtained using broadband mode acquisition and a 100 ms accumulation time. **C**: Corresponding representative spectrum obtained using RF—only selective acquisition and a 300 ms accumulation time. Resonant frequencies for RF—only dipolar excitation were identified during a broadband ion acquisition (B), and up to five species having relative abundance >10% were then data-dependently ejected during a selective acquisition that immediately followed (C). (Adapted from Paša-Tolić et al., 2002b, with permission from John Wiley Ltd., copyright 2002.)

protein extracts from mouse B16 cells, mixed in approximately 1:1 ratio (Paša-Tolić et al., 2002a). A 120 cm long column was used and data were acquired for 10 hr. Many additional species were detected with much greater signal intensities in the DREAMS spectra, especially at longer retention times (Fig. 3). Analysis of the normal spectra resulted in detection of 9,896 ^{14}N/^{15}N-labeled peptide pairs. In the DREAMS spectra, 8,856 ^{14}N/^{15}N-labeled peptide pairs were detected, of which 7,917 were new ones not detected in the normal spectra (Fig. 4). The average abundance ratios (AR) and standard deviation for the normal and DREAMS ^{14}N/^{15}N-labeled peptide pairs were 1.0 ± 0.28 and 1.015 ± 0.31, respectively. Unfortunately, an analysis of the increase in detected ORFs could not be performed because of the incompleteness and unreliability of the mouse genome and its predicted proteins. However, a similar experiment using ^{14}N- and ^{15}N-labeled peptides from *D. radiodurans* resulted in the detection of 9,559 peptide pairs in the normal set of spectra and 7,926 in the DREAMS sets of spectra. Using the *D. radiodurans* AMT tag database, 2,244 unique peptides, corre-

sponding to 965 ORFs, and 2,259 unique peptides, corresponding to 1,000 ORFs, were identified in the normal and DREAM sets of spectra, respectively. Of these, a total of 279 ORFs (30%) identified in the DREAMS dataset were "new" low abundance proteins.

Improvements in resolution of the ion ejection step, variable ion accumulation time, and shorter ion cooling time are expected to further improve the duty cycle, sensitivity, and dynamic range of the DREAMS approach. Peptides presently not included in the AMT tag database, since they were not identified as PMT tags because of low abundance of the parent protein and the relatively low sensitivity of QIT instrumentation, cannot be identified with the AMT tag approach described earlier. Thus, the PMT tag generation stage generally requires the use of additional stages of separation or fractionation prior to the "shotgun" LC-MS/MS analyses with conventional MS instrumentation. However, LC-FTICR with multiplexed MS/MS may become increasing competitive for this purpose, particularly because of the higher data quality achieved. Automated gain control (AGC) in
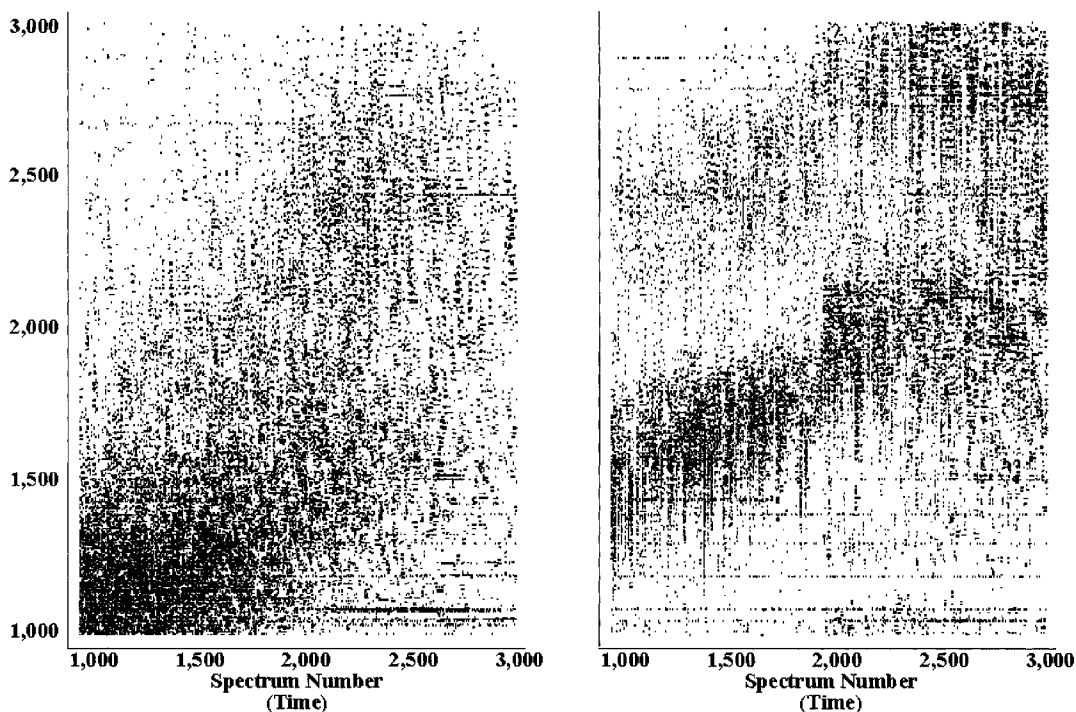
FIGURE 4. 2-D display showing the molecular masses and spectrum number for peptide pairs detected from a 1:1 mixture of ^{14}N- and ^{15}N-labeled peptides from mouse B16 cells. The left panel shows peptide pairs detected in the normal spectrum set and the right panel shows peptide pairs detected in the DREAMS spectrum set. A total of 17,813 unique peptide pairs were detected using the conservative criteria described in the text, an 80% increase over the number observed from the normal spectrum set alone. (Adapted from Paša-Tolić et al., 2002b, with permission from John Wiley Ltd., copyright 2002.)

conjunction with DREAMS serves to improve both the S/N and MMA measurements for low abundance peptides. Finally, we note that the principle of DREAMS is not limited solely to proteomics research and FTICR instrumentation; implementation for other complex analysis problems, e.g., fuel or environmental mixtures, or with ion trap instrumentation, is realistic.

D. Chromatographic Coupling

Obtaining global proteome coverage involves dealing with extremely complex mixtures of many peptides. Despite the high-resolution and MMA, even FTICR will not be able to deal with such complexity in most cases, unless the mixture has been separated by some kind of chromatography. The most commonly used separation technique is LC, mainly coupled with ESI. Using high-efficiency multiple-capillary LC operated at 10,000 psi [85 cm column length, 50 μm i.d., 3 μm C18 particle packing material (300 Å pore size), 3 hr gradient] combined with an 11.4 T FTICR, a yeast cytosolic tryptic digest was analyzed (40 μg total sample) (Shen et al., 2001). The capillary LC provided a peak capacity of ~650, and the 2-D-capillary-LC-FTICR MS analysis provided a combined resolving power of $>6 \times 10^7$. In total $>100,000$ polypeptides and $>1,000$ proteins could be characterized from a single LC-FTICR analysis. The dynamic range for peak intensities in a single mass spectrum was ~1,000. However,

the overall dynamic range for peptide and protein abundances can be even greater based on the quality of the separation, ESI related phenomena, and instrumental capabilities such as AGC and DREAMS (Smith, Shen, & Tang, 2004). Continuing advances in separation capabilities (higher pressures, smaller i.d. columns, multi-dimensional separations) are expected to further improve the dynamic range and sensitivity. Recently, attomole sensitivity has been obtained for a simple peptide mixture by modifying a 7 T nano-LC micro-ESI/FTICR mass spectrometer, including a shorter octopole, ~100% duty cycle, improved nano-LC micro-ESI emitter tips, and reverse-phase HPLC resins that require no ion-pairing agent (Quenzer et al., 2001). The reported attomole sensitivity represented the total amount of sample loaded onto the nano-LC column, eluted, and subsequently detected.

CE was first coupled to FTICR by Hofstadler et al. (1993) and has recently attracted considerable interest (Wetterhall et al., 2002). Benefits of CE vs. HPLC are higher separation efficiencies, shorter analysis time, and lower sample consumption. Disadvantages can include more limited sample loading compared with LC, limited separation of neutral species, and more variable results. Wetterhall et al. used CE-ESI/FTICR to analyze a tryptic digest of human cerebrospinal fluid (CSF). More than 10,000 peaks were generated, corresponding to 1,500 unique mass classes. Fourteen proteins were identified with 99% confidence (from a 150 protein database containing proteins known to be

present in human CSF and plasma), and 30 with 95% confidence, based on MMA error of ≤5 ppm. ECD also has been incorporated in LC- and CE-FTICR MS experiments to analyze peptide mixtures and protein digests. Improvements in the electron injection system design made it possible to acquire fragment ion spectra using 10–100 ms electron pulses (Tsybin et al., 2001). However, the ECD sensitivity was low; so multiple spectra (minimum of four) had to be added together to improve S/N to enhance identification probability. The time scales of ECD and typical chromatographic peak widths are of the same order (2–4 s). A typical experiment consisted of alternately acquiring LC(CE)-FTICR MS and LC(CE)-FTICR ECD MS/MS spectra by switching the electron gun off and on. No ion or charge state selection was performed. For a mixture of substance P, melittin, and oxidized insulin B-chain, enough sequence tags were generated to unambiguously identify the three peptides; whereas for a BSA tryptic digest, ECD fragments from 15 tryptic peptides could be identified by database searching (Palmblad et al., 2002b). Two analyses on human serum albumin protein tryptic digest were performed using LC(CE)-FTICR ECD MS/MS (Tsybin et al., 2002): a CE-FTICR MS analysis followed by a CE-ECD-FTICR MS analysis. The two data sets were matched with regard to parent and fragment ions. Based on the fragment abundance and charge state, only a few of the tryptic fragments were considered as potential candidates for ECD MS/MS. Even though advances have been made, the relatively low sensitivity with ECD and slow acquisition rate of the FTICR still limits the applicability of ECD for large-scale whole proteome analyses.

E. High-Throughput Aspects

One of the major challenges in proteomics, as well as in genomics and in combinatorial chemistry for drug discovery and materials developments, is high-throughput analysis. High-throughput efforts generally involve the incorporation of FTICR, separations, and data transfer, analysis, and data representation methods into one integrated, automated unit that ideally runs 24/7. Almost all efforts to date have focused on LC-ESI/FTICR MS, but recently LC-MALDI/FTICR MS has been presented as the basis for a high-throughput platform. When high quality mass spectra are required, enormous amounts of data can be generated even in a relatively short (e.g., 3 hr) LC separation with 256 kB to 1 MB data points per spectrum every 1–5 s. For extremely complicated samples, extending the separation time is not a practical option. Higher magnetic fields can allow faster spectrum acquisition; however, this leads to more data that have to be stored and processed. Whereas FTICR does not directly add great benefits compared to high-throughput platforms that incorporate QIT or TOF instruments with conventional approaches, the higher MMA afforded by FTICR can enable the AMT tag approach and thus provide the basis for greatly improved throughput. Another advantage of FTICR is higher sensitivity and dynamic range of measurements, which allows analysis of smaller sample amounts compared to other MS techniques and provides significantly greater proteome coverage for a given sample size and analysis time using the AMT tag approach.

FTICR can also offer advantages when MS/MS approaches are applied. Unlike QIT or TOF instruments, FTICR has the unique ability to simultaneously isolate various ions of different

m/z values by using SWIFT or correlated harmonic excitation field techniques. These isolated ions can then undergo sequential MS/MS by various dissociation techniques. Masselon et al. (2000) implemented data-dependent multiplexed tandem FTICR MS with reversed phase LC (RPLC) to enhance peptide identification throughput. "On-the-fly" data-dependent experiments that included a dynamic exclusion option to maximize the number of peptides identified in a single LC run by reducing the redundant fragmentation and thus further enhancing the throughput were controlled by in-house developed ICR-2LS software. Two sequences were continuously repeated: MS followed by MS/MS. After acquisition of the MS spectrum, a SWIFT waveform was generated to isolate multiple parent ions (up to 4), followed by application of multiple frequency SORI waveforms, generated as a superposition of individual SORI waveforms, to simultaneously dissociate the ions. IRMPD or ECD methods do not require the injection of a collision gas into the ICR cell and thus shorten the MS/MS sequence time. Parameters were chosen such that extensive fragmentation occurred for most peptides. In addition, a relatively wide dynamic range was realized, demonstrated by high-quality dissociation spectra for both high- and low-abundance species. Analysis by LC-FTICR multiplexed MS/MS (IRMPD and SORI-CID) of a BSA tryptic digest resulted in identification of 44 peptides (18 peptides using both methods) (Li et al., 2001). In general, smaller peptides were identified by IRMPD, whereas larger peptides were more effectively identified by SORI-CID. Overall, SORI-CID fragmentation provided more useful fragmentation. Note that low abundance species may be omitted for MS/MS analysis, and the duty cycle is relatively slow compared to MS/MS experiments using QITs. The addition of data-dependent automation has recently been reported for the modular ion cyclotron resonance data acquisition system (MIDAS) (Freitas, King, & Shi, 2003). Real-time generation of single notch SWIFT excitation events will allow automation of FTICR MS/MS experiments for those using the MIDAS system.

A major issue in high-throughput proteome analysis is the large range of relative abundances for eluting peptides and cyclotron frequency shifts that result from large variations of ion populations in the FTICR cell. Ideally, one would like to have a similar ion population as used for the external calibration; however, in practice this is not easy. Recently, AGC was incorporated into LC-FTICR experiments to overcome this problem by dynamically varying the accumulation time of eluting peptides in the external multipole based on the total signal chromatogram of low resolution pre-scan data (64 kB) at constant time (200 ms) to avoid overfilling the cell (Belov et al., 2003a). The latter event may result in m/z discrimination or induce dissociation in the accumulation multipole. Based on the total signal chromatogram and the base peak intensity, the accumulation time may be varied between 50 ms and 1 s. For a tryptic digest of 4 casein proteins (0.05 mg/mL) using a 0.25 s accumulation time, 2,165 putative peptides were identified (based upon MMA less than 50 ppm with a 3.5 T instrument) from a single LC-FTICR MS experiment. In a similar experiment using AGC, 6,513 putative peptides were identified. The large increase in putative peptide identifications can be attributed to the improved MMA because of the more similar frequency shifts resulting from the more uniform size of the ion populations.

The use of LC-MALDI/FTICR MS instrumentation for proteomics (Peters et al., 2002) mentioned earlier, has features that make this arrangement attractive. Like most LC/ESI combinations, RP-μHPLC separates complex peptide mixtures, and the eluting peptides are automatically deposited in real-time with matrix onto surface-patterned MALDI targets to minimize degradation of the chromatographic separation quality. The resulting plates are quickly analyzed in the MS mode using an FTICR with a high-speed MALDI interface. Based upon the high-resolution and MMA, peaks of interest are identified for further MS/MS analysis. This approach potentially eliminates the future need to rerun the separation as required for on-line LC-MS separations. Peptide mapping was greatly enhanced by labeling lysine residues with in-house developed 2-methoxy-4,5-dihydro-1X-midazole, where $X = H$ or D. Labeling enabled differential quantification by the 4 Da mass difference per lysine and increased the sequence coverage by a factor of two. In addition, the label favored ion fragmentation towards C-terminal fragment ions, thus simplifying interpretation of the tandem mass spectra. Internal calibration to ensure high ($\sim$1 ppm) MMA was achieved by irradiation of a spot containing calibration peptides and MALDI matrix. After irradiation and subsequent accumulation of sample ions, the two ion populations were mixed and simultaneously transferred into the FTICR cell. The total time to acquire a mass spectrum was 7–10 s, so around 384 internally calibrated spectra could be acquired in 1 hr. The high MMA allowed identification based solely on mass, as with the AMT tag approach. For unidentified peptides, subsequent MS/MS spectra could be obtained after automated data processing. An advantage of the MALDI approach is that one is not limited to a maximum of four parent ions to be investigated per MALDI spot. A disadvantage of this approach is the much broader m/z range compared to ESI because of the nature of the MALDI ionization process and the extended analysis times if large numbers of peptides are selected for MS/MS. In general, the resolution decreases as m/z increases. For example, a 5 μg tryptic digest sample of soluble yeast proteins was separated and deposited every 10 s with matrix on a MALDI target. Over 11,000 unique monoisotopic masses were found, with many spectra containing over 200 monoisotopic masses.

Expectations are that FTICR instruments will become more widely integrated in high-throughput platforms and that manufacturers are already developing ready-to-use systems, including automated digestion, faster separations, data-dependent experiments (various MS/MS techniques, DREAMS, targeted MS/MS), and data processing plus visualization capabilities. In our lab we have developed and successfully operated a prototype automated high-throughput LC-FTICR for proteomics for more than a year. We expect that LC will remain the most widely used on-line separation technique, either combined with ESI or MALDI, although efforts are underway to automate and make other separation techniques more reproducible for use in a high-throughput environment.

F. Quantitative Proteomics

Obtaining information, in terms of both relative and absolute abundances as to which proteins in a proteome change under certain conditions is important for elucidating biological cycles and the functions of proteins involved in these cycles.

1. Isotope-Coded Affinity Tags (ICAT)

Because of the enormous complexity of protein digests, it is often advantageous to first reduce the complexity of these digests by introducing some type of specificity enabling isolation of a subproteome. A frequently used approach is to label peptides at amino acids that are relatively less common. By labeling a control sample with a light isotope and a perturbed sample, for instance grown under different conditions for bacteria cultures, with a heavy isotope, and then mixing the two samples, one can obtain qualitative information on the relative ARs of both proteomes. The widely used ICAT include labeling cysteine amino acid residues.

Using an AMT tag approach, ICAT labeling was used to identify and quantify changes in protein expression between control and camptothecin-treated mouse cortical neurons (Yu et al., 2002). FTICR enabled quantitative measurements to be made with a high degree of precision. More importantly, the high sensitivity and dynamic range allowed the relative ARs of low-abundant species to be detected. The ARs for more than 125 proteins were identified and quantified; an AR of 1.7 or more was considered as a significant change. A particularly large percentage of the enzymes involved in the glycolysis and tricarboxylic acid cycles were observed. Unfortunately, because the ICAT-D_0 and ICAT-D_8 labeled peptides elute at close times, quantification becomes more difficult since the AR may be different from spectrum to spectrum. Thus the integrated AR for all spectra containing the peptides of interest provides the most accurate relative quantification.

2. $^{12}C/^{13}C$, $^{14}N/^{15}N$, and $^{16}O/^{18}O$ Labeled Peptides

The elution time spread and the fairly large modification introduced by ICAT labeling have stimulated investigations of alternative approaches to circumvent these limitations. Recently, Berger et al. (2002) investigated a complex peptide mixture produced from global Lys-C digestion of soluble yeast proteins, which were combined in a 1:1 ratio after growth in synthetic media containing either normal lysine or ^{13}C-labeled lysine (six carbon atoms). Data-dependent MS/MS experiments were used to select only fully digested peptides and to subsequently identify them. Fully Lys-C-digested lysine unlabeled/labeled pairs were identified by searching the spectra for pairs of isotopic distributions with characteristic m/z differences of 6.02 (singly charged), 3.01 (doubly charged), or 2.0 Th (triply charged), with a tolerance of 0.1 Th. A pair was selected by SWIFT, and simultaneous dissociation was accomplished by applying a SORI waveform of two excitation spikes: one at $\Delta m/z = -3$ Th for the ^{12}C peptide and the other at $\Delta m/z = +4$ Th for the ^{13}C peptide. Because the global Lys-C digestion results in fewer peptides, the LC separation is simplified and the number of MS/MS experiments is reduced. In addition, the peptides are larger and more likely to contain sites for multiple charging. In general, these peptides will also be easier to fragment. A total of 17,857 different peptides (unique mass classes) were observed, and after

each MS sequence the most abundant pair was selected for subsequent MS/MS. Overall, 48% of the components were observed to be paired, whereas 59% of the components in the top 10% of signal intensity were observed as peptide pairs. Of the observed pairs, 89.4% corresponded to fully digested peptides ($\Delta m = 6$ Da), 10.1% to peptides containing one missed Lys-C cleavage ($\Delta m = 12$ Da), and $\sim 0.5\%$ to peptides with multiple missed Lys-C cleavage sites. The LC-MS analysis of the [^{12}C/^{13}C] lysine-labeled digest revealed co-elution of the unlabeled and labeled peptides over the entire course of the experiment. In addition, the peptide pairs eluted with roughly equal intensity. The average AR was 1.0 ± 0.2 for the $\Delta m = 6$ Da pairs. The MS/MS spectra of the selected pairs contained fragment ions from both the N- and C-terminus. Only C-terminal fragments (y-type ions) contained lysine isotope information and showed up as doublets, whereas the N-terminal fragments (b-type ions) appeared as single peaks. Finally the lysine labeling allowed differentiation between lysine and glutamine. An interesting example was presented in which two peptides, NVPLYKHLADLSK and NVPLYQHLADLSK, originating from the proteins yeast enolase 1 and 2, respectively, could be distinguished (0.0036 Da). Since a pair with $\Delta m = 6$ was observed, it was clear that only one lysine was present, and consequently the protein had to be enolase 2.

Conrads et al. (2001) used a combination of ^{15}N-metabolic labeling and cysteine-reactive biotin affinity tag (PEO-biotin) to isolate and quantitate cysteine-containing peptides from *D. radiodurans* and mouse B16 melanoma cells. PEO-biotin is similar to the ICAT reagent, but less expensive. The ^{14}N/^{15}N PEO-biotin labeled mixtures were combined in a 1:1 ratio and analyzed by LC-FTICR MS. The high resolution afforded by FTICR clearly distinguished the ^{14}N/^{15}N peptides pairs and facilitated assignments. More than 600 pairs of cysteine-peptides were observed for the *D. radiodurans* sample with an average AR of 1.12. In the mouse B16 melanoma cells sample, the spectra revealed that the labeling of proteins grown in the ^{15}N-enriched medium was effectively complete based on the observed isotopic distributions. This analysis demonstrated the use of ^{15}N isotopic labeling of mammalian cells in culture to provide a "reference proteome." In addition, the combination of ^{14}N/^{15}N and PEO-biotin did not induce any significant isotope-dependent chromatographic shifts, which is important for reliably obtaining more quantitative information.

Yao et al. (2001) used ^{16}O/^{18}O labeling of tryptic peptides from two virion proteins for two serotypes (Ad2 and Ad5) of adenovirus. The ^{18}O atoms were incorporated universally into the carboxyl termini of all tryptic peptides during the proteolytic cleavage. HPLC separated fractions were collected, and each fraction was investigated by MALDI/FTICR. The high-resolution allowed many peaks to be observed as pairs with a characteristic mass difference of 4 Da. An average AR of 3.3 ± 0.6 was determined, in close agreement with the arbitrary set ratio of 3:1. The dynamic range in protein concentrations in andovirus is ~ 600-fold, easily obtainable by FTICR; however, not all proteins could be identified and quantified because of low corresponding peptide signal intensities. Some unpaired sets of isotopic peaks were observed, and this may indicate mutations, post-translational modifications, or extreme changes in expression levels of corresponding proteins from different pools.

G. Post-Translational Modifications

Biological activity of many proteins is regulated by the extent and position(s) of post-translational modifications. An important and common post-translational modification is phosphorylation of serine, theorine, and tyrosine, which regulate cellular processes such as metabolism, growth, and reproduction. Other important post-translational modifications are glycosylation, sulfonation, and carboxylation of glutamic acid residues. In general, using SORI-CID on an FTICR leads to loss of the modification, and the resulting spectra may become very complicated (Kelleher et al., 1999a). ECD has been shown to often retain the modification, thus allowing determination of the position of the modification.

McLafferty and co-workers investigated a series of model phosphopeptides, atrial natriuretic peptide substrate, platelet-derived growth factor, and PY8, a synthetic peptide corresponding to a sequence in the cytoplasmic domain of insulin-like growth factor I receptor, using SORI-CID and ECD (Shi et al., 2001). As expected, SORI-CID gave extensive loss of H_3PO_4 and HPO_3, whereas these losses were almost absent in the ECD spectra. In general, both SORI-CID and ECD gave complementary backbone cleavages for identifying the modification sites.

Zubarev and co-workers used nano-ESI/FTICR ECD and SORI-CID to study singly and multiply phosphorylated peptides from bovine β-casein or synthetic origin (Stensballe et al., 2000). Extensive fragmentation by ECD was observed for the synthetic phosphopeptide TRDI(pY)ETD(pY)(pY)RK, producing a complete c_2–c_{11} ion series as well as complementary $z_2^{\bullet}$–$z_{10}^{\bullet}$ ion series. The results confirmed both the complete peptide sequence and all three phosphorylation sites. The corresponding SORI-CID spectrum appeared to be more complex. Nano-scale Fe(III)-immobilized metal ion chromatography (IMAC) purified phosphopeptides from bovine β-casein produced singly and quadruply phosphorylated peptides. For the latter, relatively low ECD efficiency was observed, but nonetheless, 23 out of 27 peptide bonds were observed cleaved. Two phosphoserines were directly localized, whereas the other two phosphoserines were indirectly localized.

Kelleher et al. (1999b) performed ECD, IRMPD, and SORI-CID on three, 28-residue synthetic peptides of human proprothrombin (proPT28; residue sequences of -18 to $+10$), with differences between each peptide occurring at the following sites: Glu^{24}-Glu^{25}, Glu^{24}-Gla^{25}, and Gla^{24}-Gla^{25} (Glu = glutamic acid, Gla = γ-carboxyglutamic acid). The synthetic peptide KVACLLIKDDKADPNSVTREY, sulfated at cysteine, was investigated as well. Both IRMPD and SORI-CID resulted in CO_2 loss as the most facile fragmentation pathway for the Glu-Gla and Gla-Gla proPT28 4+ ions. In the ECD spectra, over 90% of the CO_2 group(s) were retained in the c_{25}^{3+}–c_{27}^{3+} ions. This demonstrated the application of ECD for identifying carboxylation of glutamic as a post-translational modification. For the sulfated peptide [M + 4H]$^{4+}$ ions, small amounts of H_2O and H_2SO_3 loss were observed, but no such losses from the 12 c-type and 9 z-type fragment ions.

To date, no successful application of a chromatographic separation technique with ECD to identify post-translational modifications in a real biological sample has been demonstrated; however, it may be too early and likely unwise to rule out the possibility of ECD being valuable in high-throughput multiplexed LC-FTICR MS/MS experiments.

Flora & Muddiman (2001), (2002) developed a method to identify phosphopeptides in enzymatic digests using ESI/FTICR MS with IRMPD. In the negative ion mode, IRMPD resulted in distinct product ion formation involving loss of H_3PO_4 (-98 Da) or loss of H_2O for phosphorylated serine and threonine and loss of HPO_3 and H_2O for phosphorylated tyrosine. For equimolar concentrations (100 nM) of HGLDNYR (M) and HGLDN(pY)R (M_p), selective phosphopeptide signal enhancement was accomplished when the pH was adjusted with 20 mM piperidine (pH = 11.8). The average ratio of M_p to M was 3.28 ± 0.33. Using a similar method for a tryptic digest of β-casein resulted in a detection limit of phosphopeptides of 500 pM. Irradiation of ions with infrared light at 11.1 W for 4 s led to H_3PO_4 and $H_2O + HPO_3$ loss for phosphorylated peptides, whereas the unmodified peptides were unaffected primarily because of the lower activation energy for dissociation of phosphorylated peptides. For the tryptic digest of α-casein, various phosphorylated peptides were identified in a single-acquisition mass spectrum at a concentration of 10 nM.

The fragment ions corresponding to five potential H_3PO_4 losses for a pentaphosphorylated peptide, QMEAE(pS)I(pS)-(pS)(pS)EEIV PN(pS)VEQK, were not observed above the spectral noise because of low signal intensities. For the tryptic digest of β-casein, a tetraphosphorylated peptide (M_{pppp}) was observed, corresponding to amino acid positions 17−40 with phosphorylation of serines 30, 32, 33, and 34. Upon IR irradiation, these $[M_{pppp}-2H]^{2-}$ ions lost one water molecule and one to four H_2PO_4 groups. All fragment ions were identified with high MMA (-3.5 to $+10.8$ ppm error spread) using internal calibration with poly-ethylene glycol ions.

Finally, hot ECD (HECD), which is similar to conventional ECD except that the average kinetic energy of the electrons is 3−13 eV instead of 0−1 eV, was shown to distinguish leucine from isoleucine residues. Under HECD conditions, these residues fragmented by $C_3H_7^\bullet$ and $C_2H_5^\bullet$ loss, forming *w*- and *d*-type fragment ions, respectively (Kjeldsen et al., 2002). Zubarev and co-workers applied HECD to the bovine milk protein PP3 (135 amino acid residues, 8 C-Arg peptides separated by HPLC) (Kjeldsen et al., 2003). They identified 20 out of 25 Ile/Leu (Xle) residues, and 2 were correctly inferred from the absence of *w*-type ions. Three Xle residues could not be assigned.

H. Peptide and Protein Mass Spectra Analysis Software for FTICR

Mass spectra of peptide and protein mixtures and fragmentations can be very complicated, such that, manual interpretation becomes impractical. The high-resolution capability of FTICR allows for characterization of more complex mixtures that would otherwise be extremely challenging. The high MMA provides increased confidence in peptide and protein identifications by computational database searching methods. When enormous amounts of data are generated, especially in a high-throughput environment, data should be processed automatically to extract as much useful information as possible, and ideally assigned with levels of confidence.

Horn et al. developed "thorough high-resolution analysis of spectra by Horn" (THRASH), a fully automated computer algorithm especially designed for application to complex mass spectra of peptides and proteins (Horn, Zubarev, & McLafferty, 2000a). The main features of this algorithm are isotope cluster identification, automated charge determination, generation of a theoretical isotopic abundance distribution, matching the theoretical and experimental isotopic abundance distributions, identification of other overlapping isotopic peak clusters, determination of improved accuracy of averaged centroid mass values, matching of assigned mass values to proteolytic or MS/MS fragment ions, and baseline plus S/N calculations using data between the isotopic peaks. Application of THRASH to the ECD spectrum of the SWIFT-isolated cytochrome *c* 15+ molecular ion resulted in identification of 282 isotopic clusters, of which 231 were assigned by the program as a peptide fragment in the *m/z* range of 550–1,800 Th, using an algorithm initiation threshold S/N of 1.5 in 19 min. In comparison, a whole day of manual and computer assisted interpretation resulted in identification of 220 isotopic clusters and 183 peptide fragments. The algorithm has been implemented in both the ICR-2LS and MIDAS FTICR data analysis programs.

Identifying proteins by traditional database searching of MS/MS data is not an option for genomes that have not been sequenced. In such cases or when proteins have undergone translational modification, *de novo* protein sequencing becomes necessary. Horn et al. developed a program that requires input of an accurate mass value for the monoisotopic mass of the molecular ion (M_m) being sequenced and a separate list of monoisotopic masses from CID and ECD spectra (by THRASH), and that allows a user defined MMA tolerance (Horn, Zubarev, & McLafferty, 2000b). Both CID (nozzle-skimmer in this case) and ECD are needed to maximize the number of overlapping sequences. The program first identifies complementary fragments within the MMA tolerance specified from the list of masses; for ECD these are $c + z^\bullet = M_m + H^\bullet$ and $a^\bullet + y = M_m + H^\bullet - 27.99$ (CO) and for CID, $b + y = M_m$. Application of this program to ubiquitin (8.6 kDa, 76 amino acids) resulted in correct prediction of the sequence, except for Leu/Ile. However, no unambiguous assignment could be made for cytochrome *c* (12.4 kDa, 104 amino acids) because of the lack of information in certain residue regions.

Zhang & Marshall (1998) developed a universal algorithm for charge state determination and deconvolution ("zero-charge" mass spectrum) of ESI/FTICR *m/z* spectra (Zscore). Both high- and low-resolution spectra, including overlapping charge state distributions, can be processed automatically and rapidly. The main applications of this algorithm are for interpreting complex mixture spectra and identifying non-covalent adducts, protein CID fragment ions, and chemical modification of electrosprayed bio-macromolecules.

Kelleher and co-workers developed a quantitative model for probability-based identification of intact proteins using only 3−4 nonadjacent fragment ions and no intact mass bias for protein identification with 99% confidence from genomes of <5 megabase size (Meng et al., 2001). This methodology was developed for use with the top down approach in combination with ESI/FTICR. They showed that to uniquely identify proteins, three out of four fragment ions must be measured at ±0.1 Da MMA, or four out of four at ±1 Da MMA. They identified 16 proteins from *Methanococcus jannashii* using a 3,592 proteins database.

In our laboratory, software has been developed to facilitate proteomics analyses using LC or CE separations with ESI/

FTICR. ICR-2LS software is used for database searching, data processing, fragment identification, and calibration, as well as for many more options. Other linked software produces 2-D displays of molecular weight vs. elution time, spectrum number, or pI. This analysis tool can be coupled to the AMT tag and PMT tag databases for automatic peptide identification, AR determination, and stable isotopic pairs (ICAT, ^{14}N/^{15}N) identification, and uses data files automatically generated by ICR-2LS. Further efforts are under way to abstract new and useful information from the many, rich mass spectra recorded and processed every day.

III. INTACT PROTEIN IDENTIFICATION: THE ''TOP DOWN'' APPROACH

Intact protein level analyses are generally less effective for protein identification than peptide level measurements, but offer the attraction of insights unobtainable at the peptide level. Using the bottom up strategy, complete sequence coverage of proteins is rarely achieved in proteome analyses, thus limiting the ability to examine site-specific mutations and post-translational modifications of individual proteins, which may be important in biological function. Given these limitations, there has been increasing interest in ''top down'' proteome characterization strategies, where individual proteins are selected for analysis by MS, without the need for prior chemical or enzymatic proteolysis. This section covers the various aspects of the top down approach required to obtain useful information on protein masses, structures, and modifications for successful identification.

A. Protein Mass Measurements and Characterization

Confident protein identification by measuring accurate molecular masses has been the focus of much FTICR research over the past decade. This unique capabilitiy of FTICR has allowed for many proteins to be identified. However, there are many aspects that need to be considered and limitations that need to be recognized prior to successful application of FTICR for whole protein identification and for obtaining additional information that cannot be obtained by other techniques.

1. Mass Limits

The highest mass molecule for which unit (isotopically) mass resolution has been achieved has not increased in the last 5 years from the previous record achieved for chondroitinase I, a 997 amino acid protein with a molecular weight of 112,508 Da (measured 112,509 Da) at a resolving power of 170,000 (Kelleher et al., 1997). As stronger magnetic fields (e.g., 12 T) become available for commercial FTICR instruments, attempts most likely will be made in the coming years to break that record. Unit mass resolution is useful to identify, e.g., disulfide bridges [-S-S- vs. 2-SH (2 Da)], deamination [$-NH_2$ vs. -OH (1 Da)], or other post-translational protein modifications, as well as adduct formation. Additionally, if ^{13}C, ^{15}N double-depletion of proteins is used, then even higher mass limits can be expected, perhaps >400,000 Da. Unfortunately, as the molecular mass of proteins increases, using ^{13}C, ^{15}N double-depleted protein will no longer have a simple isotopic distribution in which the mono-

isotopic mass can easily be determined, as shown for proteins up to ~20,000 Da. However, accurate masses of proteins exceeding 200,000 Da have been determined using charge state distributions that did not exhibit unit mass resolution, so unit mass resolution is not absolutely necessary.

2. Mass Accuracy

FTICR has the potential for ultra-high MMA, but even after careful calibration, relatively large errors may be obtained for several reasons. Such errors are often because of frequency shifts caused by space-charge effects, similar to mass analysis of low mass ions. The extent of these effects is dependent on the ion population in the FTICR cell. To quantify this effect, Easterling et al. used MALDI/FTICR to show that a linear correlation exists between the observed frequency and the ion intensity (Easterling, Mize, & Amster, 1999). They obtained excellent linearity by pulsing nitrogen gas prior to detection, which allowed damping of the trapping motion, and by axialization via quadrupolar excitation, which dampened the magnetron motion and centered the ions in the magnetic field axis. In addition, they reported that either the space-charge-induced frequency shift is largely independent of the mass (frequency) separations of ions in the mass spectrum or the influence of the space-charge effect on the observed frequency shift is a collective effect of the total charge in the cell. Thus, high MMA could be obtained by either (1) establishing a calibration curve using a mass spectrum with a total ion intensity that is close to that of the analyte or (2) accounting for the total ion intensity explicitly in the calibration procedure. Using both of these approaches and a PEG 3400 sample for external calibration for the insulin B-chain, average errors of 1.6 and 2.3 ppm were obtained for the two approaches, respectively. The authors suggested that use of a large diameter cell could reduce space-charge effects, but considered the negative consequence of magnetic field inhomogeneity as well.

Taylor & Amster (2003) used a similar approach to show that space-charge-induced frequency shifts do not change with increasing charge on the ion. By applying internal calibration, using the 2+ and 3+ charge states of insulin B-chain, the 8+ to 13+ charge states of hemoglobin chain A were measured with approximately 1 ppm error, mainly because of the fact that the calibrant and analyte peaks experience the same total ion density and, therefore, a very similar space-charge-induced frequency shift. For larger proteins, using an external calibration that takes the total ion intensity explicitly into account appeared to be less successful than for peptides.

Chan and co-workers obtained baseline-resolved broadband MALDI/FTICR mass spectra covering mass ranges as large as 1,000–5,700 Da (Chan, Duan, & Sze, 2002). Using a combination of scanning FTICR, multiple-ion filling, and potential ramping, accurate molecular mass measurement for peptide and protein mixtures could be obtained. A parabolic calibration curve and the above-mentioned techniques were used to detect bovine insulin (MW = 5,734 Da) in its dehydrated (or diaminated) form, isotopically resolved (mass resolution full-width half height of ~14,000).

Lee et al. (2002) applied internal calibration to LC-FTICR analysis of the set of intact proteins from the yeast large ribosomal subunit. Using a simultaneous dual-ESI setup, one

from the LC column and one from a syringe with calibrant, low mass errors were obtained that helped to constrain the database search criteria. γ-Endorphin was used as the internal standard and the $[M + 2H]^{2+}$ ion ($m/z = 929.9664$ Th) was used as a "lock mass" during the data processing.

Hannis & Muddiman (2000) designed a dual-ESI source with sequential hexapole accumulation for internal calibration. The analyte and internal standard were sequentially sprayed (by moving the two ESI tips back and forth in front of the heated metal capillary), accumulated, transferred, and trapped in the FTICR cell, and then simultaneously detected. Application of this method on a polymerase chain reaction (PCR) product with a theoretical average mass of 50,849.20 Da resulted in an absolute average error of -8.9 ppm, using $d(pT)_{18}$ as the internal standard. For an external calibration, this was 44 ppm (five measurements). Recent developments in our laboratory in this field have led to the incorporation of a dual ion funnel on FTICR instruments to allow internal calibration (Belov et al., 2003b). Internal calibration can also be incorporated into MS/MS experiments, as demonstrated by Flora, Hannis, & Muddiman (2001) and Kruppa and co-workers (Kruppa et al., 2002), to obtain high MMA for product ions.

Smith and co-workers developed a method to correct for space-charge effect induced frequency shifts without the need for internal calibration to obtain accurate masses (Bruce et al., 2000). The "deconvolution of Coulombic affected linearity" (DeCAL) process uses multiple charge states of the same neutral mass to calculate the frequency shift to correctly align at least two deconvoluted isotopic envelopes. This process is repeated for all charge state pairs present in the mass spectrum. The final frequency shift applied to all data is based on a weighted average of all the charge state frequency shifts, with the ion signal intensity being an important factor. DeCAL has been successfully applied to peptides. For proteins with many possible charge states, the MMA errors were shown to increase for lower charge states (e.g., higher m/z values).

Williams and co-workers showed that proteins and peptides could obtain more charges if m-nitrobenzyl alcohol was added to the ESI solution (Iavarone, Jurchen, & Williams, 2001). The addition of the alcohol caused the charge states to move to lower m/z regions (where higher resolution is obtained) by increasing the number of attached protons (proportional to z) (Marshall & Hendrickson, 2001). The higher charge states should give rise to lower MMA errors. On the other hand, higher isotopic density may give rise to peak coalescence.

Masselon, Tolmachev, and Smith found systematic errors from the commonly used FTICR calibration law that could not be accounted for by a global space charge correction (Masselon et al., 2002). By changing the ion population and excitation radius, they concluded that each ion cloud experienced a different interaction with other ion clouds, and no identical shift occurred. A correction for each m/z ion cloud based on the total ion population minus its own ion population improved the MMA considerably.

3. Elemental Composition

Ultrahigh mass resolution and accurate ion relative abundances are the key elements required to determine elemental composition. However, in general this is not relevant, especially when working with (complex) mixtures. Marshall and co-workers observed isotopic fine structure for ubiquitin at a mass resolving power of 8,000,000, with complete elimination of peak coalescence. This was accomplished by combining several techniques: high magnetic field (9.4 T) external accumulation ESI/FTICR MS, digital quadrature heterodyne detection, minimization of the number of trapped ions, frequency-domain averaging of multiple data sets, and use of a large diameter (10.2 cm) ion trap. The number of ions was reduced by SWIFT isolation of one charge state and by stepwise reduction of the trapping potential from 2 (during ion trapping) to 0.12–0.20 V over a 60 s period. The high magnetic field and large diameter ion trap should reduce space-charge effects, and consequently, should reduce peak coalescence. By applying these techniques on the ^{13}C, ^{15}N doubly depleted p16 tumor suppressor protein (15.8 kDa), they were able to determine from the resolved isotopic fine structure that the number of sulfur atoms present in the protein was five (Fig. 5) (Shi, Hendrickson, & Marshall, 1998) using only micrograms of protein sample. Even though no practical application (even with a simple complex mixture) can be expected using this method, at least the combination of accurate mass and some elemental composition information may be helpful to identify a protein derived from a relatively simple genome.

4. Metal Atom Oxidation States

Metal ions play an essential role in orienting and stabilizing protein conformations in metalloproteins that fulfill a catalytic or enzymatic function (Loo, 1997). Changes in metal oxidation states can occur when the metal center acts as a mediator in redox reactions. Because metals in metalloproteins may carry charge that contributes to the total charge of an ESI generated ion, knowledge about the oxidation state of the metal or metal-containing center is critical for determining accurate molecular masses. Unfortunately, the ESI process can induce redox reactions, so the oxidation state of a metal in the gas-phase may be different from the native, condensed phase.

With its high resolving power and mass accuracy, FTICR in combination with ESI has been used to study the oxidation states of a number of metalloproteins. Amster and co-workers investigated a series of iron-sulfur proteins to probe the stoichiometry and oxidation states of the metals centers (Johnson et al., 2000). For the cuboidal [4Fe-4S] cluster from *Pyrococcus furiosus* ferredoxin and its mutant D14C, an oxidation state of 2+ was found in combination with a disulfide bond. Sulfur adducts in the disulfide bonds were also observed. The [2Fe-2S] cluster in the hydrogenase from *Thermotoga maritime* had an oxidation state of 2+, the highest state that has been reported under biological conditions.

In another study, Johnson et al. (2001a) used ESI/FTICR to record the positive and negative ion mode mass spectra of iron and zinc rebredoxin from *Clostridium pasteurianum* and ferrodoxins from *P. furiosus* containing [3Fe-4S], [4Fe-4S], [3FeNi-4S], and [3FeMn-4S] clusters. All clusters were stable as negative ions, whereas only the [4M-4S] (M = Fe, Mn, Ni) clusters were stable as positive ions. The [3FeMn-4S] cluster had never been observed before. Figure 6 shows the positive and negative mass spectra of the [3FeMn-4S] cluster from *P. furiosus*.
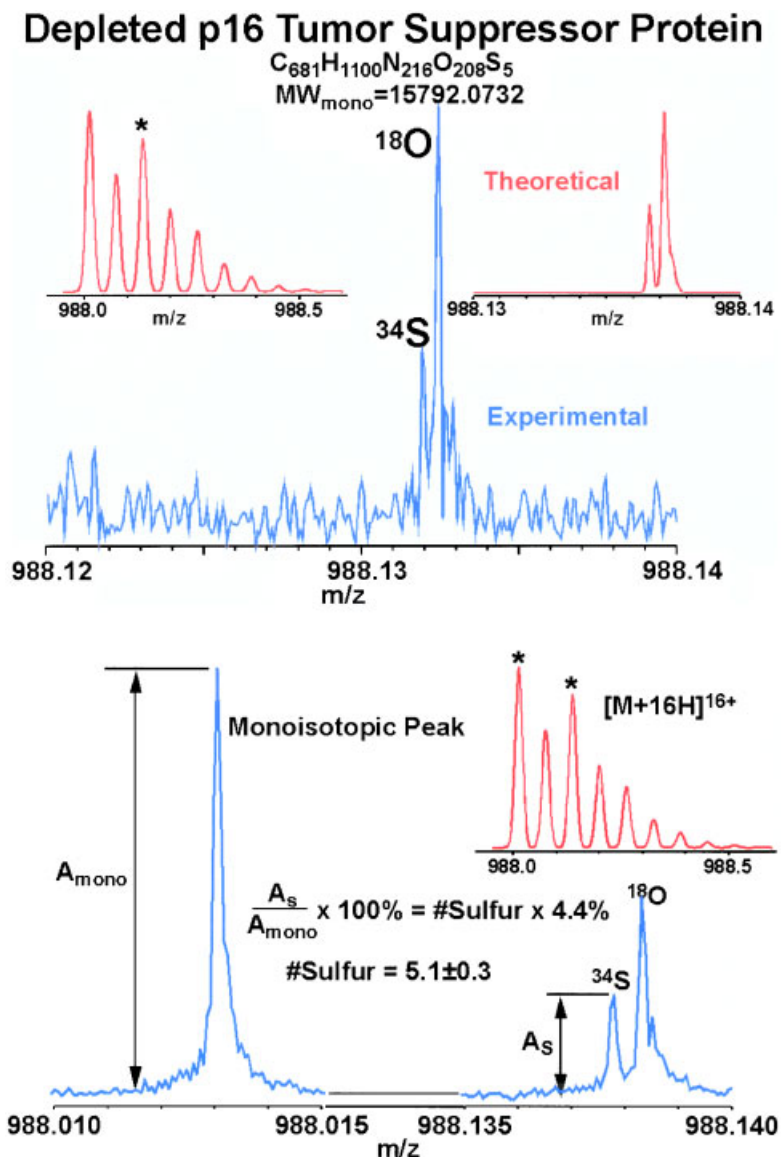
FIGURE 5. Top: Mass scale-expanded spectra of one isotopic peak for ^{13}C, ^{15}N doubly depleted p16 tumor suppressor protein, from a single electrospray ionization (ESI) FTICR mass spectrum. The theoretical isotopic distribution is shown at upper left, in which (*) denotes the isotopic peak ≈ 2 Da above the monoisotopic mass. The theoretical isotopic fine structure for the starred peak is shown at upper right. Double-depletion improves the S/N of the monoisotopic, ^{34}S and ^{18}O species, by reducing the abundances of species containing ^{13}C and/or ^{15}N; (**bottom**) similar spectra as on top obtained by co-adding nine frequency-domain ESI-FTICR spectra. From the abundance ratio (AR) of the (resolved) ^{34}S to the monoisotopic species, the number of sulfur atoms in the protein was determined to be 5.1 ± 0.3, without use of any other information about the protein. (Adapted from Shi, Hendrickson, & Marshall, 1998, with permission from National Academy of Sciences, USA, copyright 1998.)

In this figure, the negative cluster ions exhibit a wider charge distribution than the positive ones. In the positive ion spectra, [3Fe-nS] (n = 1–4) products can be observed. An oxidation state of 1+ was determined for this cluster. In general, in the gas-phase, similar chemistry may be observed, e.g., the loss of manganese from the $[3FeMn-4S]^{1+}$ cluster to form the $[3Fe-4S]^{1+}$ cluster. Johnson et al. (2001b) also were able to observe

reduced forms of several metalloproteins by both chemical and electrochemical reduction. The latter method is less prone to reoxidation. The oxidized and reduced forms were clearly distinguishable because of a 1/z shift in the mass spectrum.

He et al. showed that the iron atom oxidation state in myoglobin, cytochrome c, cytochrome b5, and cytochrome b5 L47R is 3+ for all proteins, similar to the condensed-phase, and is
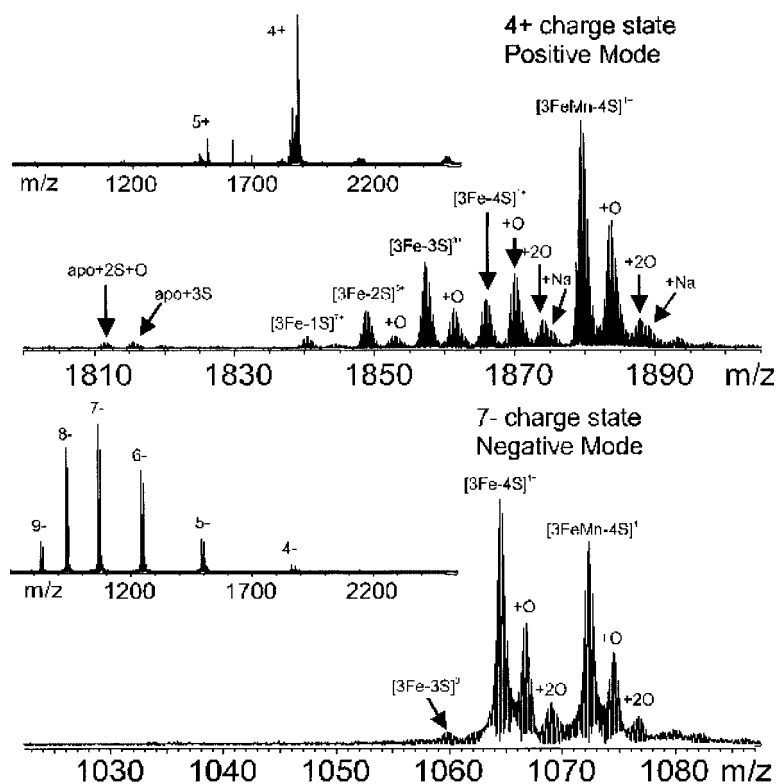
FIGURE 6. ESI-FTICR mass spectra obtained for [3Fe-4S]$^-$ containing ferredoxin in positive ion mode (**top**) and negative ion mode (**bottom**). A stable cluster is observed only for the negative mode spectrum. The insets show an expansion of the 4+ and 7− charge states. The peaks corresponding to protein with an intact cluster are labeled. Major peaks represent degradation products that consist of [3Fe-3S], [3Fe-2S], [3Fe-1S], and [3Fe] are found only in the positive ion mass spectrum. Additional peaks corresponding to sodium and oxygen adducts are present at lower intensity. (Adapted from Johnson et al., 2001a, with permission from Elsevier, copyright 2001.)

conserved during the ESI process (He, Hendrickson, & Marshall, 2000). For these experiments it was important not only to have high resolving power and mass accuracy, but also to have accurate relative isotopic abundances. This was particularly important when fitting experimental to theoretical abundances, expressed in the form of a standard deviation; the smaller the fitting standard deviation, the closer the resemblance of the two abundances. Because space-charge effects could potentially distort the relative abundances, experiments were performed using small ion populations of one charge state. Co-addition of many scans improved the ion isotopic statistics.

5. Isotopic Enrichment

Stable-isotope depletion of proteins using FTICR has been shown to be a useful approach for extending both the mass range for accurate mass determination of intact proteins and the sensitivity. Marshall and co-workers demonstrated that by depleting the content of ^{13}C and ^{15}N to less than 0.01%, the mass range increased from 10 to 100 kDa, for which accurate (<1 Da) mass determination was achievable (Rodgers et al., 2000).

Demirev & Fenselau (2002) used ESI/FTICR to analyze the isotopic distribution of doubly depleted (^{13}C and ^{15}N) human ubiquitin. Both the mass and intensity ratios of the observed isotopic peaks in the multiplet were used to confirm the ratios of ^{12}C:^{13}C and ^{14}N:^{15}N. By comparing the average mass and the overall profile of the experimental distribution with the average mass and isotopic distribution of combinations of ^{12}C:^{13}C and ^{14}N:^{15}N (varying from 5:95 to 2:98 for both elements), the best fit was found for a 3:97 enrichment ratio for both elements, with an error of ±0.5% for C and ±1% for N.

Smith and co-workers used accurate mass measurements and partial amino acid content information to unambiguously identify intact proteins (Veenstra et al., 2000; Martinović et al., 2002). Proteins were extracted from *Escherichia coli* and *Saccharomyces cerevisiae* grown in either minimal medium or minimal medium to which isotopically labeled amino acids (Leu-D$_{10}$, Ile-D$_{10}$, Phe-D$_8$, Arg-^{13}C$_6$, His-^{13}C$_6$, and Lys-^{13}C$_6$) had been added at a concentration of 0.1 mg/mL. The two protein extracts were mixed and analyzed by capillary isoelectric focusing (CIEF) coupled with ESI/FTICR. The two forms of the protein focused together (same pI) during the CIEF separation
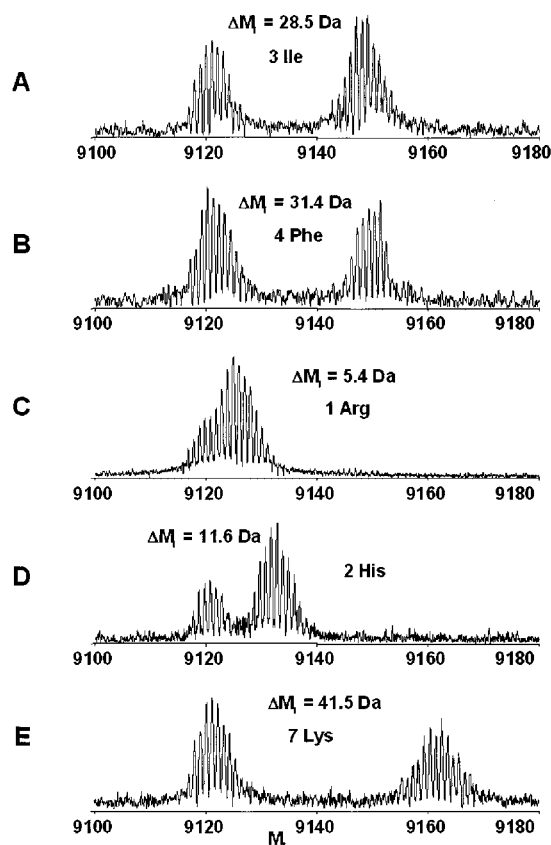
FIGURE 7. Zero charge state spectra of the *Escherichia coli* PTS system protein HPr (Mr ~ 9,119.4 Da) detected during on-line CIEF-FTICR analysis from *E. coli* grown in minimal medium combined with cells grown in minimal medium containing 0.1 mg/mL of (**A**) Ile-D$_{10}$, (**B**) Phe-D$_8$, (**C**) Arg-^{13}C$_6$, (**D**) His-^{13}C$_6$, or (**E**) Lys-^{13}C$_6$. (Adapted from Martinović & Smith, 2002, with permission from John Wiley Ltd., copyright 2002.)

and were observed together in the same mass spectra. In a single run, several hundred proteins could be observed and the unlabeled and labeled pairs were easily recognized in a 2-D display. The constraints provided by three or more amino acids were enough to uniquely identify each protein. Figure 7 provides an example of the identification of the *E. coli* PTS system protein HPr. It was observed that in approximately half the proteins, the initiating Met residue had been removed from the nascent chain by a ribosome-associated Met-aminopeptidase. This method should aid in identification of post-translational modifications if, once the protein has been identified, there is a mass discrepancy between the theoretical and experimental mass.

6. Detection Limits

To date, the lowest reported FTICR detection limit for whole proteins was 0.7–3.0 attomole for 8–29 kDa proteins by McLafferty and co-workers (Valaskovic, Kelleher, & McLafferty, 1996). Here, samples were introduced by CE, and the

amounts reflect the total sample size introduced to the column. Baykut et al. used MALDI/FTICR (4.7 T) with hexapole ion accumulation to acquire a spectrum from 200 attomole of substance P with a resolution of 300,000 and an S/N of 133 (Baykut, Jertz, & Witt, 2000). A mass spectrum with S/N 16 could be obtained from 10 attomole. Using a 7 T FTICR, a single scan mass spectrum from two laser shots with a resolution of 900,000 could be obtained for insulin B chain. Belov et al. (2000) coupled an electrodynamic ion funnel to a 3.5 T FTICR, and by optimizing ionization, ion transmission, and ion trapping in both the accumulation quadrupole and the FTICR, detection limits of approximately 30 zeptomole were obtained for 8–20 kDa proteins. The 30 zeptomole is based on an estimation of the total amount of protein used during the experiment. It seems likely that many of the *E. coli* proteins detected in the CIEF-FTICR studies by Smith and co-workers were present at low attomole levels (Veenstra et al., 2000; Martinović et al., 2002). With the introduction of higher magnetic field FTICRs ($\geq$12 T) and other such improvements, further improvements to detection limits are expected.

B. Protein Structures and Fragmentation

Structure determination of proteins by FTICR MS/MS is benefited by the high resolving power ($>10^5$) and MMA (<10 ppm) for both the parent and fragment ions. SORI-CID, nozzle-skimmer CID, IRMPD, BIRD, and ECD have all been applied as fragmentation techniques. In general, for a given set of conditions and instrumental constraints, the dissociation efficiency reduces as the protein size increases. In addition, the charge state distribution under denaturing conditions becomes wider as the protein size increases, and the different charge states may show different fragmentation patterns and dissociation efficiencies. Fragmentation of one selected charge state may cause loss in sensitivity, because only a fraction of the protein molecules are in that specific charge state, whereas collective fragmentation of all charge states may produce a fragment ion spectrum with abundant and possible overlapping of fragment ions. To overcome some of these problems, McLuckey and co-workers developed techniques that use QIT mass spectrometers (Reid & McLuckey, 2002) and positive–negative ion reactions to manipulate charge states of protein ions generated by ESI. Protein ions can be subjected to gas-phase concentration ("ion parking" into one specific charge state), isolation, and purification, before undergoing MS/MS followed by ion–ion reaction. All these techniques are possible because of the ability to simultaneously trap positive and negative ions, and the large m/z range on QITs. In general, spectra look very clean and provide structural information. To date these techniques have not been applied with or to FTICR, and are of limited relevance because of the drop-off in FTICR performance with increasing m/z.

ECD has been the subject of much research over the last 5 years, and successful applications have been demonstrated on both peptides and proteins. Because of its nature, ECD provides information on the locations of post-translational modifications for both peptides and proteins. Improvements have mainly focused on increasing the sensitivity and rate of ECD and include modification of the electron injection system (Tsybin et al., 2001), collisional activation of ions during ECD or so-called

activated ion (AI)-ECD, plasma ECD (Sze et al., 2003), HECD (Kjeldsen et al., 2003), and IRMPD/ECD. McLafferty and co-workers characterized the post-translational modifications of carbonic anhydrase to within one residue (Sze et al., 2002). By using a variety of ESI solutions (to manipulate charge state distributions), ion activation (CID, IR laser, cell temperature), and ECD (electron energy and current) conditions that provide various fragmentation efficiencies and patterns, 250 out of the possible 258 cleavages of inter-residue locations were eventually determined; however, the highest number of cleavages in one particular set of conditions was only 138. For the same protein, plasma ECD gave rise to cleavage of 183 the inter residue bonds. Additionally, plasma ECD provided *de novo* sequencing of ubiquitin, except for two residues, whereas for casein all of the 26 Ser/Thr/Tyr phosphorylation sites where characterized (Sze et al., 2003).

Apparently, there is not one set of ECD conditions that can provide all the information needed to obtain such a detailed map as presented here. Clear understanding of the effects of each parameter may lead to establishing a set of ECD experiments that may be used as default in whole protein analysis. Unfortunately, this approach will substantially limit high-throughput applications. The highest mass protein on which ECD has been successfully applied to date is phosphopantothenoylcysteine synthetase/decarboxylase, a 45-kDa enzyme involved in Coenzyme A biosynthesis, although information from SORI- and nozzle-skimmer CID were used as well (Ge et al., 2002). Costello and co-workers studied the 11-kDa proline-rich protein (PRP)-3 from human saliva by ECD (Leymarie et al., 2002). The fragment ion spectra of ECD were much easier to interpret than the CID spectra obtained on a Q-orthogonal-TOF instrument. Both techniques provided complementary information and yielded >93% sequence coverage.

Kelleher and co-workers took a different path to characterize proteins above ~70 kDa, a rough limit above which the top down method becomes less effective, even at 9.4 T (Forbes et al., 2001). Their approach employed proteolysis and used, for instance, Lys-C to generate 5–10 kDa peptides. For a >90% pure Pyochelin chelator E sample (159 kDa), complete sequence coverage was obtained from six Lys-C peptides of 5, 8, 26, 32, 40, and 48 kDa, all within 20 ppm mass accuracy in one mass spectrum. Figure 8 exemplifies this strategy for obtaining high percent sequence coverage for mixtures of wild type protein with masses >70 kDa. Note that subsequent MS/MS to identify post-translational modifications is easier with lower mass peptides. Automation of protein MS/MS, including data processing and identification, has been demonstrated by Kelleher and co-workers (Johnson et al., 2002).

Marshall and co-workers introduced the "focused radiation for gaseous multiphoton energy transfer" (FRAGMENT) technique to obtain activation energies for unimolecular dissociation of biomolecular ions (Freitas et al., 1999a). This method measures the unimolecular rate constant as a function of a continuous-wave CO_2 laser intensity. The method is similar to the BIRD technique developed by Williams and co-workers (Price, Schnier, & Williams, 1996), and there is close agreement between data obtained by both methods. However, FRAGMENT requires much less time to obtain the same amount of data and covers a larger temperature range. Unfortunately, no data on
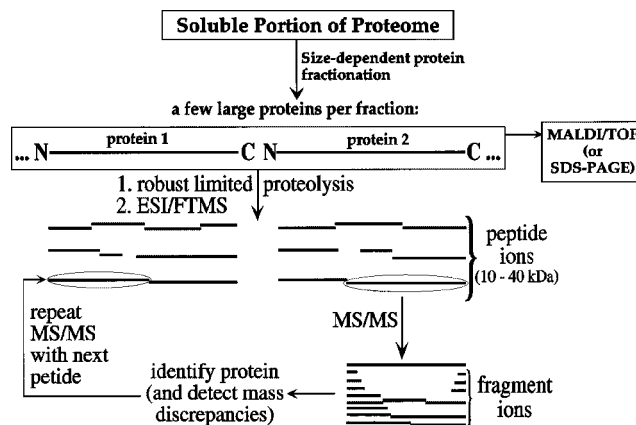
FIGURE 8. A strategy for obtaining high percent sequence coverage for mixtures of wild type proteins >70 kDa. (Adapted from Forbes et al., 2001, with permission from John Wiley Ltd., copyright 2001.)

protein ions or protein–protein complexes have been published to date.

Kelleher et al. (1998) investigated six gene products from the *E. coli* thiamin biosynthetic operon (ThiX; X = C, E, F, G, H, S) by FTICR MS/MS using several techniques (nozzle-skimmer dissociation, SORI, IRMPD). Various errors and modifications to the DNA-derived sequences were detected and localized. No modifications were observed for ThiE ($M_r = 23,015.1$-*14* Da), and the measured M_r of 23,014.8-*14* Da agreed within 13 ppm of the predicted value. ThiF was identified with the amino terminal methionine removed. For ThiF, it was noted that compared to a previously determined DNA sequence, the Cys242 had been replaced with Arg and Trp, whereas the N-terminus start methionine had been deformylated. For ThiC with a predicted $M_r = 74,528$ Da (627 amino acids plus a 37 amino acids N-terminal histidine tag), the N-terminal Met had been removed. Two components for purified ThiG were identified, one with $M_r = 26,896.5$-*17* Da (ThiG) and the other with $M_r = 7,310.74$-*4* Da (called ThiS). These components were believed to co-purify and associate non-covalently. The M_r value for ThiG was confirmed with $\Delta M_r = +0.4$ Da. ThiS was consistent with the predicted 12 N-terminal residues of ThiG with the deformylated start Met still attached. It was stated that a TGA stop codon for ThiS can be formed by a reading frame shift of −1.

Fridriksson et al. (2000) used SORI-CID and ECD to characterize the glycosylation and oxidation of three constructs of the Fc segment of immunoglobin E (190 kDa): Fcε $(3-4)_2$ (52 kDa), Fcε $(3-4)_2$ R334S, and Fcε $(2-3-4)_2$ (76 kDa) disulfide-bonded homodimers. For the Fcε $(2-3-4)_2$ DNA-predicted sequence, an M_r of 72,604 Da can be determined before addition of any post-translational modifications. However, after reduction of the eight S-S disulfide bonds, no monomer with $M_r = 72,604$ Da was detected. Instead, various glycoforms were identified, indicated as 2 hexNAc + n hex (n = 5–9). Similar glycosylation patterns were identified for Fcε $(3-4)_2$ and Fcε $(3-4)_2$ R334S. Samples at both the protein (reduced) and peptide (non-

reduced and reduced) levels were analyzed. In a proteolytic Asp-N digest of Fcε (2-3-4)$_2$, the glycosylation site Asn-394 was confirmed, as well as the N371Q mutation. As expected, MS/MS of relatively large proteolytic peptides combined with accurate molecular masses can provide very detailed information on complex modification patterns.

Uchiki et al. (2002) characterized the monomeric and dimeric forms of recombinant Sml1p-histag protein by FTICR MS/MS, using both SORI-CID and extended hexapole accumulation (3,100 ms) dissociation. The deconvoluted mass spectrum revealed several variants. Both dissociation methods gave substantial fragmentation with similar fragmentation patterns. The average masses of all the observed fragment ions corresponded to within 0.075 Da of the calculated fragment ion masses expected from the proposed sequence. Unfortunately, no prior isolation of single protein forms for the extended hexapole accumulation dissociation was provided, and all forms dissociated simultaneously. The results indicated that Sml1p-histag protein strongly binds alkali metal ions and that the various observed forms correspond to different levels of Na$^+$ and K$^+$ adduction. Interestingly, it appears that metal ions bonded to the proteins were lost once the protein ions dissociated. The dimer did not show much metal ion binding compared to the monomer. A disulfide bridge between cysteine residues at position 17 might be responsible for the dimer formation. Using extended hexapole accumulation times of 1–2.5 s generated abundant y-type fragment ions from the C-terminal region of the dimer, indicating an open structure at the C-terminal end and possibly a higher order structure at the N-terminal end.

Ge et al. (2002) characterized secreted proteins and glycoproteins from *Mycobacterium tuberculosis* by SORI-CID, IRMPD, and ECD FTICR MS/MS. Secreted proteins were precipitated from cell culture filtrate, resuspended, and examined directly by ESI/FTICR MS. In the ESI spectra of three different precipitates, a total of 689 different and accurate M$_r$ values were observed, compared to 620 by 2-D gel electrophoresis. Ten proteins were identified by MS/MS, for which ECD provided the most valuable results. The low success for assigning molecular weights was speculated to be because of the high degree of protein modification that accompanies its secretion. Extensive proteolysis, loss of signal sequences and N-terminal residues, oxidation, as well as extensive glycosylation (up to eight hexose attachments) presumably make protein characterization extremely challenging; a general problem for such analyses.

Ge et al. (2003) determined the oxidation sites in viral prolyl 4-hydroxylase by top down MS. A mixture of both non-oxidized and oxidized molecular ions was dissociated with IRMPD and "in-beam" activated ECD. Alternatively, specific ions were isolated by SWIFT followed by IRMPD or SORI-CID. Four oxidation sites were identified by MS/MS analysis (R^{28}-V^{31}, E^{95}-F^{107}, V^{125}-D^{150}, and K^{216}). From the MS spectra, relative activities of these sites were determined to be 5.5:7:1:8.5, respectively. In addition, sequence comparison between various prolyl 4-hydroxylase enzymes demonstrated that K216, H137, and 139 are conserved residues and are likely to be at the active site of prolyl 4-hydroxylase, suggesting that the oxidation of V^{125}-D^{150} and K^{216} is occurring at the active site. Using the bottom up approach would have complicated identification of the V^{125}-D^{150} and K^{216} oxidation sites.

C. Protein Confirmations and (un)Folding

The biological activity of a protein strongly depends on its tertiary structure or conformation. ESI is a "soft" ionization technique that can retain protein conformations as they are transferred from the condensed phase into the gas-phase, and various ESI-FTICR-based approaches are potentially useful for investigating protein confirmations. For instance, Wu et al. (1995) showed that for the 15+ charge state of various cytochrome *c* ions (bovine, tuna, rabbit, horse) dramatic SORI-CID fragmentation patterns could be observed that were related to changes in the higher-order structure because of small variations in amino acids at certain locations. Kaltashov and co-workers probed the conformational stability of cellular retinoic acid-binding protein (CRAPB) I, a predominantly β-sheet protein, using H/D-exchange in solution (Eyles et al., 2000) and FTICR. Species in which secondary structure elements become labile to exchange with the solvent can be populated by altering the solution pH. All labile hydrogen atoms were first replaced with deuterium and then back-exchange was initiated at certain times. This process was followed by nozzle-skimmer CID (80–300 V) in an FTICR. The exchange occurring in solution was measured in fragment ions as a function of exchange time. Certain exchange patterns indicated that regions of the protein became unprotected simultaneously and all of the exposed amides then exchanged cooperatively, suggesting that folding and unfolding in those regions was a highly concerted process. Some fragment ions from certain β-sheets exhibited bimodal isotopic distributions that indicated cooperative exchange of all amide deuterium atoms on these sheets. Under the CID conditions used, the lifetime of the activated protein ions was believed to exceed the microsecond timescale, making the dissociation more likely to take place in the hexapole ion guide than in the ion source. However, no clear answer was obtained as to whether or not hydrogen scrambling took place or whether the secondary structure was retained while trapping the ions. Unfortunately, the extent of fragmentation for some regions was too low to obtain more information on the conformations of these regions.

Kruppa and co-workers recently described a new top down approach to study protein tertiary structures using chemical cross-linking and FTICR (Kruppa, Schoeniger, & Young, 2003). Ubiquitin was reacted with the homo-bifunctional lysine-lysine cross-linking reagent dissuccinimidyl syberate. FTICR was then used to isolate singly internally cross-linked species by multi-CIEF, eliminating the need for pre-ESI separation. In addition, gas-phase fragmentation by SORI-CID (or IRMPD and/or ECD) replaced proteolytic digestion, and MS/MS analysis of the fragments was used to localize the cross-links. The resolution and mass accuracy of FTICR made identification of the cross-linked fragments unambiguous. A shift to lower charge states was observed for the cross-linked ubiquitin, because two basic primary amino groups reacted with the cross-linker, forming amide bonds, and thus eliminated the availability for protonation of these sites during the ESI process. Both fully and partial ("hanging") singly cross-linked products were observed. The technique of multi-correlated harmonic excitation fields was used to isolate multiple charge states of the fully singly cross-linked ubiquitin ions. The MS/MS spectra for the unmodified and modified ubiquitin ions showed larger differences, with the latter being more complex. In
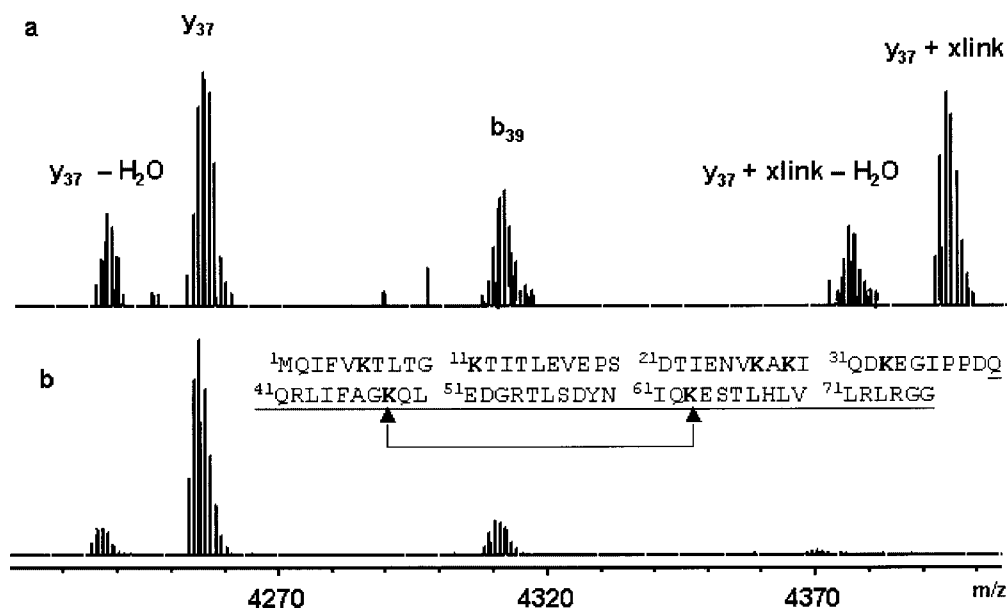
FIGURE 9. Expanded regions of an MS/MS spectrum showing the region around the y_{37} ion for (**a**) the cross-linked ubiquitin and (**b**) the unmodified ubiquitin. The sequence of ubiquitin is shown in (b), the part of the sequence corresponding to the y_{37} ion is underlined, and the arrows indicate the lysine residues that are cross-linked. (Adapted from Kruppa, Schoeniger, & Young, 2003, with permission from John Wiley Ltd., copyright 2003.)

Figure 9, expanded regions of the SORI-CID deconvoluted MS/MS spectra around the y_{37} fragment ion clearly indicate the presence of a new peak corresponding to y_{37} with a singly internal cross link ($y_{37} + x$link). The presence of these fragment ions, the cross-link distance constraint of 24 Å, and the presence of only two lysine residues in the observed fragment chain unambiguously identified the K48-K63 cross-link. The possible K6-K11 cross-link could not be identified unambiguously. For future study the authors indicated using MS[3] to localize cross-links that remained ambiguous after MS/MS. They also indicated that this method might be restricted to smaller proteins (up to 15 kDa), and that for larger proteins, proteolysis by non-restrictive enzymes may result in larger fragments that are more suited for subsequent MS/MS.

McLafferty and co-workers studied gaseous conformational structures of cytochrome c by various techniques using FTICR, including H/D exchange before and after IR heating, collisional heating, and charge stripping (McLafferty et al., 1998). For the 7+ ions, spectra were consistent with as many as seven (I–VII) stable confirmations, and for the 8+ ions similar results were obtained. A linear correlation of H/D exchange values (protein sites exchanged) versus the charge state was observed for the five distinguishable conformational states I, II, IV–VII, and they argued that decreasing the charge state only caused a gradual nonspecific opening of each conformational state, unaccompanied by local folding to form a new structure. Quadrupolar axialization for 30 min to heat the ions by collisions with D_2O exchanged 193 of the 198 possible exchangeable condensed phase H atom sites. The proportions of conformers formed at a specific charge value depended on the method of preparation,

supporting relative similar free energies for states I, II, IV, and V, with strong entropic control on their formation. To locate D atoms, MS/MS was performed on state II 13+ ions that had exchanged 132 H atoms. For SORI-CID and IRMPD, extensive, but incomplete, H/D scrambling was observed, consistent with the unfolding of the conformers before dissociation. Negligible amide H atoms exchange was observed for the residues 22–25, 48–53, 65–75, and 91–93, indicating relative protection. This negligible exchange might be indicative of an α-helix structure in these regions. In addition, results indicated that β-sheet secondary structures may also be present and that the I, II, IV, and V states are most likely mixtures of these secondary structures.

Freitas et al. (1999b) investigated the H/D exchange reactions of various bovine ubiquitin ions, $[M + nH]^{n+}$ ($n = 5$–13). The odd and even numbered charge states were separately selected by SWIFT and reacted with D_2O at a pressure of 2×10^{-7} Torr from 1 s to 1 hr. Applying these techniques was possible because of using a high field (9.4 T) FTICR. The 5+ and 6+ charge states displayed broad conformational distributions, ranging from 0 to 80% D incorporation. The 7+ to 11+ and 13+ charge states displayed single major distributions, whereas the 12+ charge state showed two distinct distributions of comparable abundance (Fig. 10). The extent of D incorporation decreased as the charge state increased.

Breuker et al. (2002) investigated the unfolding and folding of ubiquitin ions by ECD. In this very detailed study, various charge states were investigated under a number of experimental conditions that included temperature (25–175°C) and laser irradiation. A three-state process for unfolding, $C \rightarrow P \rightarrow U$, as originally proposed by Clemmer and co-workers from ion
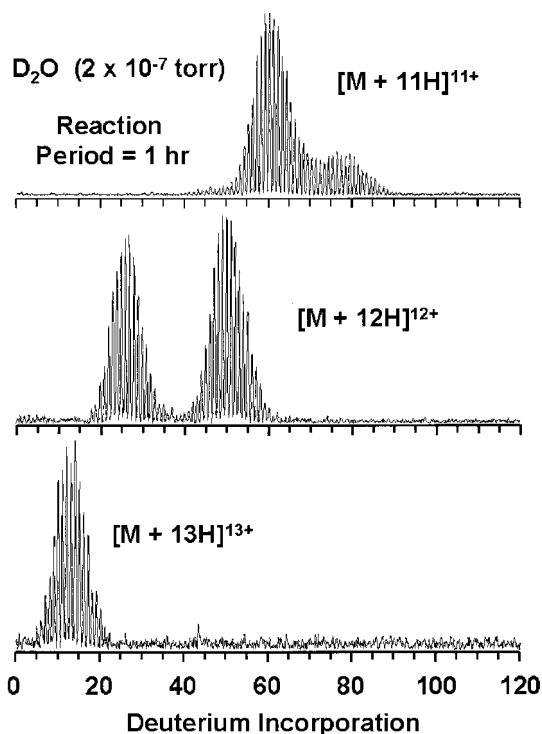
FIGURE 10. FTICR mass spectra (the abscissa represents a convolution of the deuterium incorporation distribution after 1 hr and the initial natural-abundance isotopic distribution of the fully protonated ion) following gas-phase H/D exchange for 1 hr at 2×10^{-7} Torr D_2O for the 11–13+ charge states of bovine ubiquitin. Note the presence of two comparably abundant gas-phase conformers of the 12+ charge state, baseline-resolved by extent of deuterium incorporation. (Adapted from Freitas et al., 1999, with permission from Elsevier, copyright 1999.)

mobility measurements (Li et al., 1999a), was confirmed by the ECD spectra. Folded conformers were held together by non-covalent bonds and should not have been affected by ECD. The observed $[M+nH]^{(n-1)+\bullet}$ ions (from electron capture by $[M+nH]^{n+}$ ions) were considered to be "folded" product ions. At 25°C the relative yield of unfolded product ions increased with increasing charge state, explained as an increase in the secondary structure. The 12+ ion population was believed to be a mixture of 11+ and 13+ conformers, and the 13+ ions were believed to be a near linear structure with no tertiary structure. The ECD results indicated that the 5+ charge state was in a folded conformation, whereas the 6+ and 7+ charge states started to unfold at the N- and C-termini. The 8+ and 9+ charge states were even more unfolded, a trend that continued for the 10+ ions. For the 6+ to 9+ ions, Van't Hoff plots for the unfolded ⇌ folded equilibrium were constructed for the temperature range 25–175°C (at 10°C intervals). Ions were stored in the FTICR cell for 40 s to establish thermal equilibrium using blackbody radiation prior to ECD. ΔH values of ~30 kJ/mol were obtained for the 6+ and 7+ (up to 100°C) charge states, whereas ΔH values of 11, 11, and 6 kJ/mol were obtained for the 7+ (up from 100°C), 8+, and 9+ charge states, respectively. Irradiation of ions with a continuous IR laser

(0.25 s) prior to ECD showed unfolding and refolding at short and long ECD delay times. Melting temperatures (T_m) were determined from site-specific normalized intensities of separated ECD dissociation products of the 6+ to 9+ ions at 25–175°C. T_m values of 156 and 87°C were determined for the C→P transition for the 6+ and 7+ charges states, respectively. No direct correlation with condensed phase native structures was found. It was suggested that major portions of the backbone underwent cooperative unfolding. No close structural relationship between the gas and condensed phase conformers of any proteins was assumed to exist. The C, P, and U states represent many individual conformers with similar tertiary non-covalent structures of comparable energies, and each charge state may consist of different protonation site isomers. A reasonable assumption is that this kind of detailed approach will be very complicated for larger (>12 kDa) proteins.

Hofstadler et al. performed H/D exchange reactions on ESI generated ions in an ion accumulation hexapole (Hofstadler, Sannes-Lowery, & Griffey, 2000). Using a leak-valve and an estimated D_2O pressure of ~10^{-4} mbar for the $[M+10H]^{10+}$ charge state of ubiquitin, 65 H/D exchanges were observed after 20 s. For the $[M+11H+Fe(III)]^{12+}$ charge state of bovine cytochrome c, two distinct populations of 77 and 96 H/D exchanges were observed after 5.1 s, whereas after 30 s these had gone up to 92 and 115, respectively. These results were 100-times faster than data reported in literature (Suckau et al., 1993) using in-cell reactions at a pressure of ~10^{-7} mbar.

D. Chromatographic Coupling

Most of the protein identification and characterization research discussed so far is inherently low throughput, and one can argue that it is not proteomics in that the studies do not aim at "global" measurements. Such measurements generally require additional methods for more detailed protein identification, as discussed previously. To the extent that proteomics implies working with complex mixtures, there is a need for separations. Some of the above-discussed work applied extensive off-line separations. Unlike peptide identification by LC-FTICR MS or LC-FTICR MS/MS, relatively few articles on separation and identification of whole protein mixtures using FTICR have been published to date, although this is changing rapidly. One might even envision an AMT tag approach for whole protein characterization to perform high-throughput global intact proteomics measurements.

Identification of whole proteins eliminates time-consuming digestion and sample workup steps, plus possible post-translational modifications can be identified more easily after determining an accurate molecular mass. CIEF and RPLC in combination with FTICR have been demonstrated for analyzing protein mixtures, and results show that these methods are effective for proteome characterization for at least soluble proteins.

Jensen et al. (1999) used CIEF-ESI/FTICR MS to study the cell lysate of *E. coli* cultured in normal and isotopically depleted media. In a single analysis 400–1,000 putative proteins in the mass range 2–100 kDa were observed from a total injection of ~300 ng protein (Fig. 11). For proteins with $M_m > 30$ kDa, no isotopic resolution could be obtained, but accurate average molecular masses could still be obtained by deconvoluting
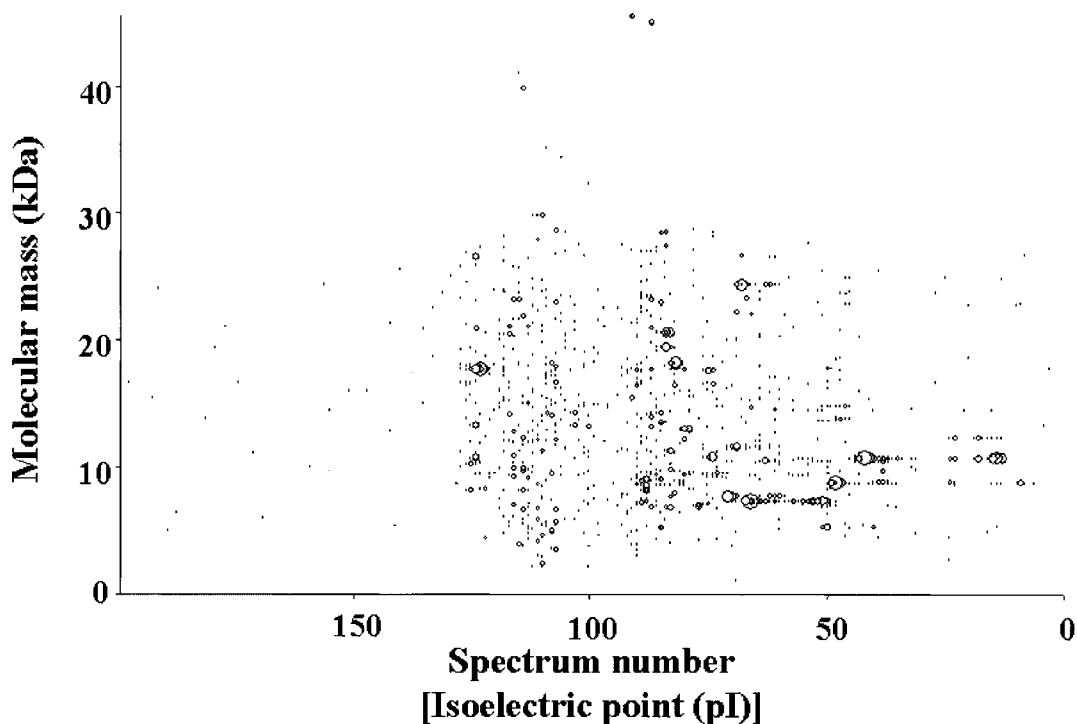
FIGURE 11. 2-D display of the CIEF-FTICR analysis for a lysate from *E. coli* grown in isotopically depleted media. (Adapted from Jensen et al., 1999, with permission from American Chemical Society, copyright 1999.)

charge state distributions. The use of isotope depletion extended the achievable accuracy, sensitivity, dynamic range, and detection limit. In some spectra up to 10 proteins could be detected. The large dynamic range and the high-resolution of the CIEF separation (up to 0.013 pI units has been obtained) resulted in some low abundance proteins being detected in only one spectrum. Therefore, for these kinds of experiments, a fast duty cycle is very important. The authors argued that accurate pI and mass measurements might still be insufficient to identify proteins, and they showed the possibility of performing MS/MS by SORI-CID after SWIFT isolation of the 5+ charge state of ^{13}C, ^{15}N double depleted cold shock-like protein CspC. The partial sequence information allowed for unambiguous identification of this protein. A similar approach was taken to study the cell lysate of *D. radiodurans* and of *E. coli* before (grown on normal media) and after being stressed by exposure to Cd^{2+} stress for 45 min (grown on isotopically depleted media) (Jensen et al., 2000). The differential expressions for various proteins showed which proteins were sensitive to external perturbation. Ideally, an AMT tag-like approach for whole proteins would benefit from a high-throughput CIEF-FTICR methodology.

Marshall and co-workers used an alternating LC-FTICR MS and LC-FTICR IRMPD MS/MS approach to separate and unambiguously identify a simple mixture of standard proteins (ubiquitin, GroES, melittin, myoglobin, and carbonic anhydrase II) (Li et al., 1999b). IRMPD requires no gas injection and shortens the duty cycle, and on-axis fragmentation provides more efficient observation of product ions. A "hit" and "tag" scoring

system was introduced for protein identification. The authors claimed their method should be suitable for identifying mixtures of as many as 30–50 proteins.

Lee et al. (2002) used RPLC-FTICR with a dual ESI setup to characterize intact proteins from the large subunit of yeast ribosome. High measurement accuracy was obtained by using one of the ESI entities to introduce an internal standard so as to enable "mass locking." The characterization of yeast large subunit proteins was challenging because of the number of proteins to be characterized, the possible patterns of protein modification, and the existence of protein isoforms with more than 99% sequence identity. Out of the 43 possible core large ribosomal subunit proteins, 42 were identified, and 58 (out of 64 possible) core large subunit protein isoforms also were identified in one single analysis. The identified co- and post-translational modifications included loss of the initiating methionine, acetylation, methylation, and proteolytic maturations. The high-resolution separation provided differentiation of protein isoforms that have high structural similarity, as well as proteins in their modified form. In Figure 12, the total ion current chromatogram was reconstructed from FTICR mass spectra acquired from the RPLC-FTICR experiment, and a 2-D display of some of the proteins identified in spectra 460–620 are shown as representative data. This example clearly shows the utility of this method for high-throughput analyses of whole protein proteomes, and also shows that MS/MS is not necessarily required for unambiguous protein identification, even if the proteins have been modified. While RPLC is less sensitive and more time-consuming than CIEF, no
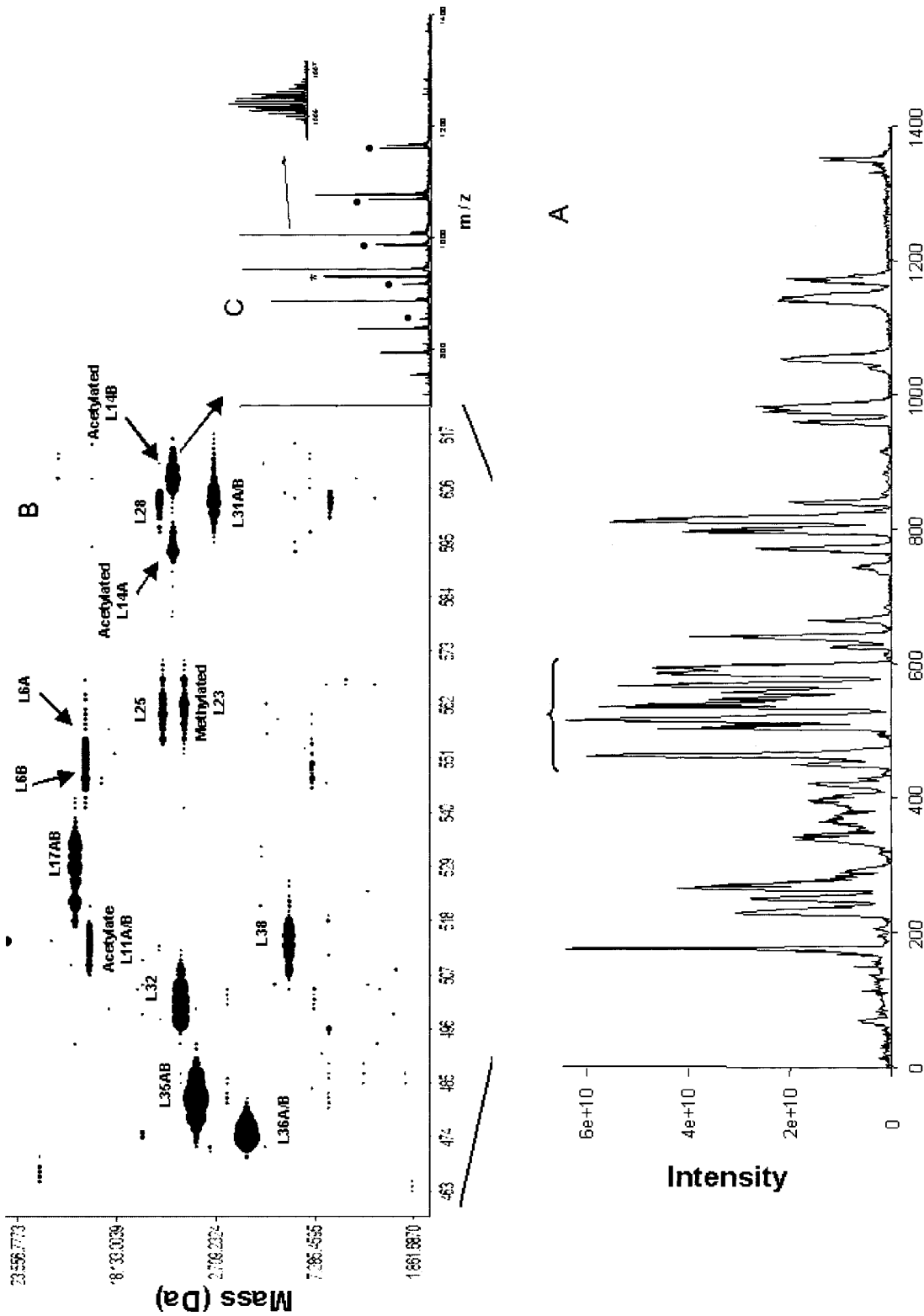
FIGURE 12. Total ion current chromatogram reconstructed from FTICR mass spectra acquired during a whole capillary RPLC-FTICR experiment. **B**: A 2-D display of proteins in spectra 460–620. The size of the black solid circle is scaled to the intensity of a protein ion signal. The designators AB and A/B indicate protein isoforms with identical mass and ones with different masses that eluted simultaneously, respectively. C: Mass spectrum of proteins from spectrum number 609. The major peaks correspond to different charge states of acetylated rpL14B. rpL14B coeluted with rpL31A and rpL31B, which are denoted by ● . * Indicates γ-endorphin internal standard. (Inset) An expanded mass spectrum of +15 charge state of acetylated rpL14B. (Adapted from Lee et al., 2002, with permission from National Academy of Sciences, USA, copyright 2002.)
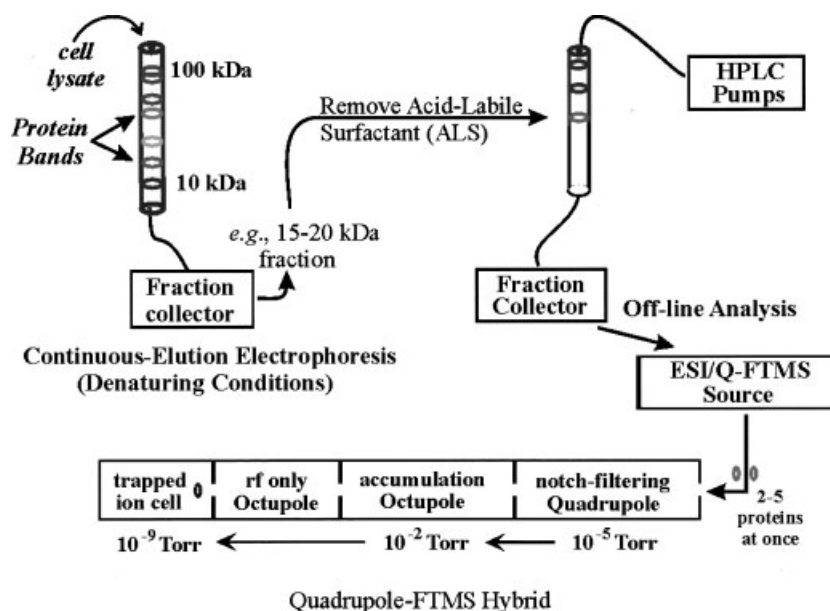
FIGURE 13. Schematic of ALS–PAGE coupled with RPLC as a proteome fractionation method before characterization of intact protein ions with a 9.4 T quadrupole-FTMS hybrid instrument. (Adapted from Meng et al., 2002, with permission from American Chemical Society, copyright 2002.)

significant chemical noise from ampholite ions was present to complicate spectra.

Kelleher and co-workers developed a protocol for efficient solubilization, fractionation, identification, and characterization of whole proteins (Fig. 13) (Meng et al., 2002). They used an acid-labile surfactant (ALS) to enhance solubilization and fractionation of a whole protein *S. cerevisiae* cell lysate by continuous-elution electrophoresis into 40 fractions. The ASL was removed, and the fractions were further separated by RPLC and collected. Approximately 5–13 proteins per RPLC fraction could be detected. MS analysis was performed using a 9.4 T Q-FTICR hybrid instrument. Direct infusion allowed for the summation of transients and thus better resolution, MMA, and lower S/N, unlike on-line coupling, which is regularly used in high-throughput experiments. MS/MS by SWIFT isolation followed by IRMPD or quadrupole selection of a $\sim$30 Th m/z window followed by IRMPD, depending on the spectra overlap, indicated that the use of the quadrupole increased the dynamic range by a factor of $\sim$30. Direct fragmentation of intact proteins, combined with automated data processing and database search analysis identified a number of proteins. Multi-dimensional fractionation to reduce the complexity of protein mixtures, combined with FTICR MS and MS/MS, when necessary, and/or further breakup of large ($>$70 kDa) proteins will become increasingly more important in proteome analyses, and there is still a lot of room for data analysis improvement.

E. Single Ion Measurements

The high-resolution and MMA obtainable only with FTICR is mainly because of the long time in which signals, resulting from the coherent cyclotron motions of the ions, can be recorded. Extended detection times also improve the sensitivity and S/N; however, large ion populations in the FTICR cell may give rise to phenomena like frequency shifts, peak coalescence, and loss of ion packet coherence.

Smith and co-workers developed a method based on the summation of sparsely to moderately populated individual ion spectra either to define charge states or to provide isotopic spacing data and thus allow mass determination (Bruce et al., 1998). The large numbers of charges required to define a charge state distribution were spread out temporally over many different spectra. The minimal Coulomb-related effects and maximal time domain signal provided for the possibility of obtaining high-resolution and MMA. In addition, high ion cloud stability allowed for an increased upper molecular mass. Data processing was critical to the success of this method and the following steps were performed: (1) "scanning" each spectrum with a peak-finder routine to locate peaks with a magnitude above a predefined threshold and (2) filtering each spectrum with a window function that forces to zero all regions of the frequency domain spectrum with magnitude less than the predefined threshold (in general 3-times the average magnitude value of the entire spectrum). Each resultant filtered spectrum was then summed with all other similarly processed spectra, followed by smoothing with a box-car integrator that summed all contributions within a user-defined window size. Using BSA, a 7 T FTICR, and conventional methods, a molecular mass of $66,429 \pm 1.7$ Da (95% confidence) was obtained.

Selective ion accumulation of a certain m/z range can be used to reduce the number of ions and allow long-term ion cloud stability. Unfortunately, this method may lead to decreased sensitivity and will fail to work if there are too many m/z values in
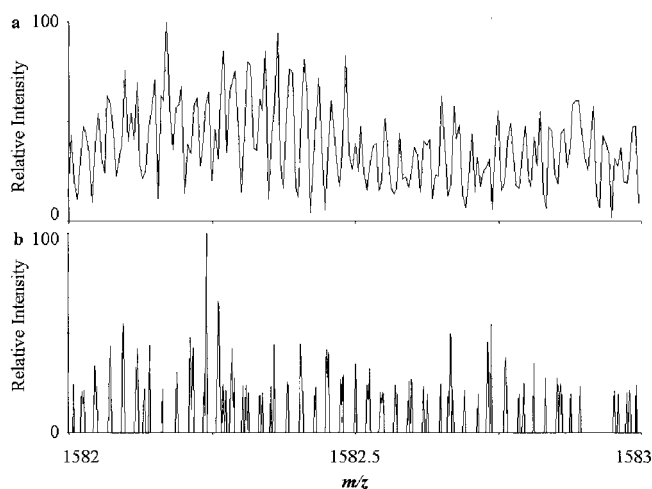
FIGURE 14. Comparison of isotopic peaks resulting from (**a**) single ion accumulation, and (**b**) summation of 100 individual ion spectra. Isotopic peaks from the individual ion summation approach appear narrower and at slightly higher frequency, presumably because of the reduced ion population present during detection of the individual ion spectra. (Adapted from Bruce et al., 1998, with permission from American Chemical Society, copyright 1998.)

a single charge state. By restricting the number of ions in the FTICR cell and representing a larger *m/z* range, space charge effects should be reduced. In the individual ion approach the extended time domain signals allow high-resolution of each individual ion's cyclotron frequency, with the possibility of defining the charge state or isotopic distributions. Summation of processed spectra does not degrade weak signals into the noise and results in increased spectral density in regions of the charge states. Application of boxcar integration resulted in some loss of resolution, yet still yielded a spectrum of conventional appearance. Using the individual ion approach to analyze BSA, Bruce and Smith obtained an average molecular mass of $66,429.9 \pm 0.7$ Da (95% confidence), in excellent agreement with the theoretical value (66,430.3 Da) and the value obtained with the conventional method. This approach can provide spectra of extremely high-resolution (Fig. 14). The upper spectrum was obtained by selective ion accumulation, whereas the lower spectrum was obtained by summing 100 individual peak-picked ion spectra, inverse-Fourier transformed, zero-filled twice, and Fourier transformed, leading to very narrow peak widths and baseline-resolved peaks. The primarily limitation of the SIA methods is the longer experiment time compared to conventional methods. Future applications could include resolvinig isotopic envelopes and isotopic "fine-structure" of very high mass biopolymers, but any practical high-throughput application seems unattainable for the time being.

IV. PROTEIN–PROTEIN COMPLEXES

Protein–protein complexes are very important in cellular functions, and their structural characterization remains a major analy-

tical challenge. Protein–protein complexes could be considered part of supra-molecular proteomics, a potentially new field of study in the post-proteomics era. A MS-based approach for studying protein–protein complexes seems to have more flexibility compared to X-ray crystal diffraction and high-resolution solid-phase nuclear magnetic resonance, two traditional methods. ESI in combination with FTICR has been used to study protein–protein complexes since ESI can retain non-covalent interactions while transferring protein–protein complexes from the condensed phase into the gas-phase. Unfortunately, protein–protein complexes tend to have lower charge states and consequently relatively high *m/z* values because of their more compact structures. This may shift protein–protein complex ion signals into *m/z* ranges where the FTICR resolution has degraded, thus losing one of the key attractions of FTICR. Nevertheless, the sensitivity and MMA afforded by FTICR still make this technique a popular choice, and a variety of different FTICR techniques have already been used to investigate protein–protein complexes.

Martinović et al. (2000) detected non-covalent complexes for a variety of proteins under CIEF-ESI/FTICR conditions. By using different sheath liquid compositions for the ESI interface, either the preservation of the non-covalent complexes or the dissociation into their sub units could be favored. Creatine phosphokinase from rabbit muscle was detected as a homotetramer, whereas homedimers were observed for glyceraldehydes-3-phosphate dehydrogenase from rabbit muscle. Low-energy SORI-CID was used to effectively remove adducts without significant dissociation of covalent or specific non-covalent interactions (Paša-Tolić et al., 1998). In addition, enhanced S/N and improved MMA were observed. Homotetramers and homodimers were detected for lentil lectin, horse myoglobin, and β-lactoglobulin, respectively. By obtaining the masses of the subunits and the non-covalent complex, the stoichiometry could be determined, as well as the presences of metal ions or other adduct species.

Derrick and co-workers investigated the calmodulin (CaM)–RS20-Ca_4 complex in the gas-phase with ESI/FTICR (Hill et al., 2000; Nousiainen et al., 2001). CaM is a small, calcium binding protein that regulates the activation of a number of enzymes and proteins in eucaryotic cells. This protein plays an important role as an intracellular receptor for calcium ions during, e.g., muscle contraction. RS20 is a synthetic peptide derived from the phosphorylation site of myosin light-chain kinase, a CaM binding peptide that inhibits CaM activation. The $[CaM + RS20 + 4Ca - 8H + nH]^{n+}$ ($n = 7$ and 8) complexes indicated that CaM was in a highly folded structure. The stoichiometry of the $[CaM + RS20 + 4Ca]$ complex was speculated to be indicative of two globular domains that functioned cooperatively. Formation of the complex was initiated by the binding of CaM and RS20, followed by the calcium ions occupying the four calcium-binding sites. SORI-CID of the $[CaM + RS20 + 4Ca]^{8+}$ complex took place by two pathways: formation of $[CaM + 4Ca - 2H]^{6+} + [RS20 + 2H]^{2+}$ and $[CaM + 4Ca - 3H]^{5+} + [RS20 + 3H]^{3+}$.

Yamada et al. used FTICR IRMPD MS/MS to study the binding sites of recombinant human interleukin-6 (IL-6, 21 kDa) with anti-human IL-6 mouse monoclonal antibody MH166 (mAb, 150 kDa) (Yamada, Suzuki, & Hirayama, 2002). IL-6 is a

globular protein with a four-helix (A–D) bundle structure consistent with up-up-down-down topology. An additional, small helix (E) was located between helices C and D. IL-6 and mAb underwent H/D exchange with D_2O in an immobilized column prior to mixing and complex formation. This step was followed by D/H exchange and removal of free IL-6. After quenching the D/H exchange, the partially deuteronated IL-6 was eluted from the column and transferred by ESI into the gas-phase. All IL-6 charge states (12+ to 23+) underwent IRMPD simultaneously and 70 fragment ions were observed, covering the complete IL-6 sequence. Calculated percentages for deuterium incorporation of the assigned fragment ions indicated relatively high levels for the Leu 126-Lys 131 (17%) and Asp 160-Met 184 (percentage not mentioned) regions compared to the average level of 3.5% for the whole protein. This high deuterium incorporation suggested that the antibodies were bound to the antigen IL-6 on the C-terminal area, where the edges of the helices had gathered. The relatively large sizes of the fragment ions made more detailed location unlikely, although MS[3] on fragment ions with a high percentage of deuterium incorporation could refine the location.

Lam et al. (2002) studied the protein–protein contact interface(s) for *in vitro* assembled human immunodeficiency virus (HIV-1) capsid protein (CA, $M_m \sim 25$ kDa). The monomeric and assembled CAs were exposed to D_2O for various periods of time (from 30 s to 68 hr, 14 time points in total), followed by digestion with pepsin. A fast (4 min) HPLC front-end separation followed by ESI and mass detection using a 9.4 T FTICR allowed for detection of peptide ions, covering 95% of the CA sequence. The 4 min separation minimized D/H back-exchange. The high-resolution and mass accuracy capabilities of the FTICR helped to determine the deuterium content in the various pepsin peptides after each H/D exchange period. Comparison of the number of exchangeable hydrogens with the H/D exchange period for peptide fragment 169–189 showed that the exchange rate in the monomeric CA was much faster than in the assembled CA. The explanation given was that peptide fragment 169–189, corresponding to helix IX, was protected against H/D exchange in the assembled CA and thus participated in the protein–protein interaction.

Freitas et al. (2000) showed that irradiation with a continuous-wave CO_2 laser at 2.5 W for 1 s removed buffer adducts from a non-covalent complex between oligopeptide-binding protein and either $(Ala)_3$ or LysTrpLys without dissociating the complex. The S/N was increased by one order of magnitude, and the cleaner spectra also made identification easier.

Kruppa et al. (2002) used a multi-correlated harmonic excitation field to isolate many charge states (32+ to 58+) of EGFr protein (62,669.07 Da) with rapamycin in the presence of unreacted protein, followed by IRMPD. EGFr protein has been implicated in the development, progression, and severity of certain human cancers, wheres rapamycin is an important immuno-suppressor antibiotic. High-resolution (>100,000) of fragment ions resolved isotopes of the ions and made charge state determination possible. Fragment ions with the rapamycin bound to them were observed, allowing identification of the binding site in the EGFr protein.

Klassen and co-workers investigated the dissociation of protein homodimer ecotin using BIRD in an FTICR (Felitsyn, Kitova, & Klassen, 2002). Nanoelectrospray of a 2.0 μM aqueous solution of ecotin (E) with 5 mM acetic acid produced a wide distribution of homodimer charge states $[E_2 + nH]^{n+}$ (n = 12–20). For dissociation of the n = 14–17 dimers, approximately 50–60% of the monomeric product ions corresponded to an asymmetric charge distribution. At longer reaction times (up to 300 s) the symmetric distributions became more important. The asymmetric charge distribution between the subunits was believed to be a consequence of the ESI/desolvation process, while some inter-subunit proton transfer was not completely ruled out. Clearly, more understanding is required to deduce additional information about composition and condensed phase structures of protein–protein complexes from gas-phase dissociation experiments, especially since the product ions were observed to be so stable that further dissociation did not occur. The observed stability may hinder the possibility of using MS[3] experiments to obtain structural information on the monomeric units. In addition, monomeric product ions may have *m/z* values that are higher than the set *m/z* range of the FTICR, and so high field instruments may be absolutely necessary for these kinds of systems, especially if high-resolution (>50,000) is required at high *m/z* values (>6,000 Th). Activation energies determined by temperature dependent (126–175°C) unimolecular dissociation rate constant measurements showed that the non-covalent binding in the homodimer ecotin ions is fairly strong (40–60 kcal/mol, depending on the charge state). It was suggested that the ESI produces various charge isomers and that these distributions reflect the product ion charge state distributions. Combinations of IRMPD for MS/MS and ECD for MS[3] of SWIFT- or correlated harmonic excitation field-isolated parent and product ions might provide structural information on these important complexes.

Zubarev and co-workers showed that ECD could be applied in polypeptide polycationic complexes (Haselmann et al., 2002). ECD initiated bond rupture of intramolecular bonds for a series of non-specific peptide-peptide dimers and a specific complex of glycopeptide antibiotics (vancomycin and eremomycin) with their target peptide (diacetyl-L-Lys-D-Ala-D-Ala); intermolecular bonds were not affected. The use of ^{13}C, ^{15}N double depleted proteins will allow high *m/z* ions to be identified.

Wigger et al. (2002) showed that by combining ESI, FTICR, SWIFT, and IRMPD in a single experiment, they could determine formation of non-covalent complexes between the preferred ligands (L) from a 324-member peptide combinatorial library (Ac-GpYEXX-Eda, X ≠ Cys and Leu) and the Hck Src homology 2 domain receptor (M). In the mass spectrum of a solution containing M and the complete library, ion signals corresponding to $[L + H]^+$, $[M + nH]^{n+}$ (n = 11–14), and the non-covalent complexes $[M + L + nH]^{n+}$ could be observed. Figure 15 shows the product ion spectrum after IRMPD of SWIFT-isolated $[M + L + 11H]^{11+}$, which produces $[M + 10H]^{10+}$ and $[L + H]^+$. The high MMA and sensitivity of the 9.4 T FTICR helped assign the two strongest ion signals in the $[L + H]^+$ *m/z* region as either XX = DI/EV or EI, respectively. Comparison of the FTICR MS based method with results from a conventional library screening affinity protocol gave similar results for the identified polar ligands. The observed similarities between the two methods were because of the dominance of polar interactions in gas-phase protein–ligand complexes, whereas the condensed phase binding strength was determined by hydrophobic interaction. The speed,
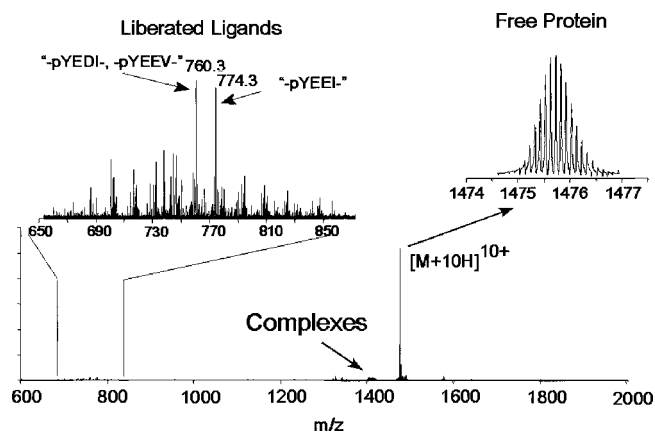
FIGURE 15. FT-ICR mass spectrum following infrared multi-photon dissociation (IRMPD) of tightly-bound non-covalent complexes, $[M+L+11H]^{11+}$, to yield free Hck SH2 domain (10+ charge state) and its (singly-charged) ligands. The insets show mass scale-expanded segments for the released ligands (upper left) and the isotopically resolved free protein $[M+10H]^{10+}$ (upper right). The two major peaks correspond to the peptides, Ac-GpYEXX-Eda, in which XX = DI/EV or EI, respectively. (Adapted from Wigger et al., 2002, with permission from Elsevier, copyright 2002.)

simplicity, and limited amount of protein needed (1–2 nmol) make this an attractive alternative approach for affinity studies.

Xi et al. (2001) identified an acyldisulfide-linked protein–protein conjugate between ThiF and ThiS. Identification was accomplished by a top down approach using alkylated ThiF ions and MS analysis of a Glu-C proteolysis reaction mixture. The Cys-184 of ThiF and C-terminus of ThiS were involved in the formation of this complex.

V. APPLICATIONS

A. Integrated Studies

Both bottom up and top down approaches have their particular advantages and disadvantages, and it seems logical that integration *a priori* should create a method that significantly improves and extends proteome analyses. The first reported integrated study came from McLafferty and co-workers, who studied bovine carbonic anhydrase B (Kelleher et al., 1999a). For the bottom up approach, the protein was digested with α-chymotrypsin and Lys-C (8 hr), and mixtures of unseparated peptides were electrosprayed by direct infusion and detected by FTICR. For α-chymotrypsin, 64 peptides were identified out of 84 fragment masses, covering 95% of the DNA-derived protein sequence. For Lys-C, 17 fragments out of 40 were consistent with fragment sequences, covering 96% of the protein sequence. MS/MS was performed on some of the fragment ions. For the top down approach, a variety of methods were applied: nozzle-skimmer and SORI-CID, Lys-C digestion for 30 min, and CNBr (Met-specific). With the combined approach, 100% coverage was

obtained, and modified protein fragments were easily identified, although an incorrect sequence, CA-X, was identified as $T^{172} \rightarrow S^{172}$. The combined approach proved to be successful, but time-consuming.

VerBerkmoes and co-workers described a preliminary analysis of the *Shewanella oneidenis* proteome (VerBerkmoes et al., 2002). Crude cell lysate was fractionated by strong anion exchange chromatography (36 fractions). A portion of each fraction was proteolytically digested, and the resulting peptides were characterized by LC-MS/MS on an LCQ ion trap mass spectrometer. Various experiments were performed on crude and fractionated cell lysate samples, as well as membrane fractions, and included data-dependent LC-MS/MS over the following 4 *m/z* ranges: (1) 400–2,000 Th; (2) 400–1,000 and 980–2,000 Th; (3) 400–1,000, 980–1,500, and 1,480–2,000 Th; and (4) 400–800, 780–1,200, 1,180–1,600, and 1,580–2,000 Th. Each fractionated cell lysate sample was also analyzed by ESI/FTICR and by SORI-CID. Both the bottom up and top down data were searched against a *S. oneidenis* database, and SEQUEST was used to process the LC-MS/MS data. With the bottom up approach, 868 proteins were identified, which is ∼17% of the theoretical proteome of *S. oneidenis*. Maximization of separation power and dynamic range parameters was considered critical to analyze mixtures of high complexity, and the extensive number of experiments performed was considered worthwhile. In fractions examined with the top down approach, 5–20 proteins were typically observed per fraction. Most of these proteins were measured with isotopic resolution at high-resolution (between 50,000 and 100,000) afforded by use of a 9.4 T FTICR. The determined protein molecular masses were searched against the database created by the bottom up approach (25 ppm MMA) and 22 proteins were identified, both intact and with N-terminal methionine truncation. MS/MS experiments were performed for some abundant unidentified proteins, but limited sequence coverage prevented definitive identification, even when common protein modifications were considered.

Figure 16 provides an example of the top down approach, showing the advantages of high field FTICR and fractionation for analyzing protein mixtures (VerBerkmoes et al., 2002). Only a small overlap of proteins identified with the bottom up and top down approaches was observed, which the authors rationalized by the fact that proteins had been modified (signal peptides, post-translational modifications), that the database used might have had errors in the predicted gene start sites, and that ions might have been because of proteolytic breakdown products. An example showed the sequence of a periplasmic protein from an ABC transporter system found in a certain fraction was covered by 85% (bottom up), but no protein mass corresponding to the predicted mass could be found with the top down approach. Removal of a 28 amino acid N-terminal signal peptide resolved the discrepancy, which was confirmed by protein MS/MS. A protein with mass 61,058.86 Da was observed using the top down approach, but not identified using the bottom up approach, although a possible candidate was available (fumarate reductase). Nevertheless, the integrated approach holds future promise, and automation plus data analysis with advanced bioinformatics software should allow for higher throughput of more complex samples and for obtaining more information from the MS data.
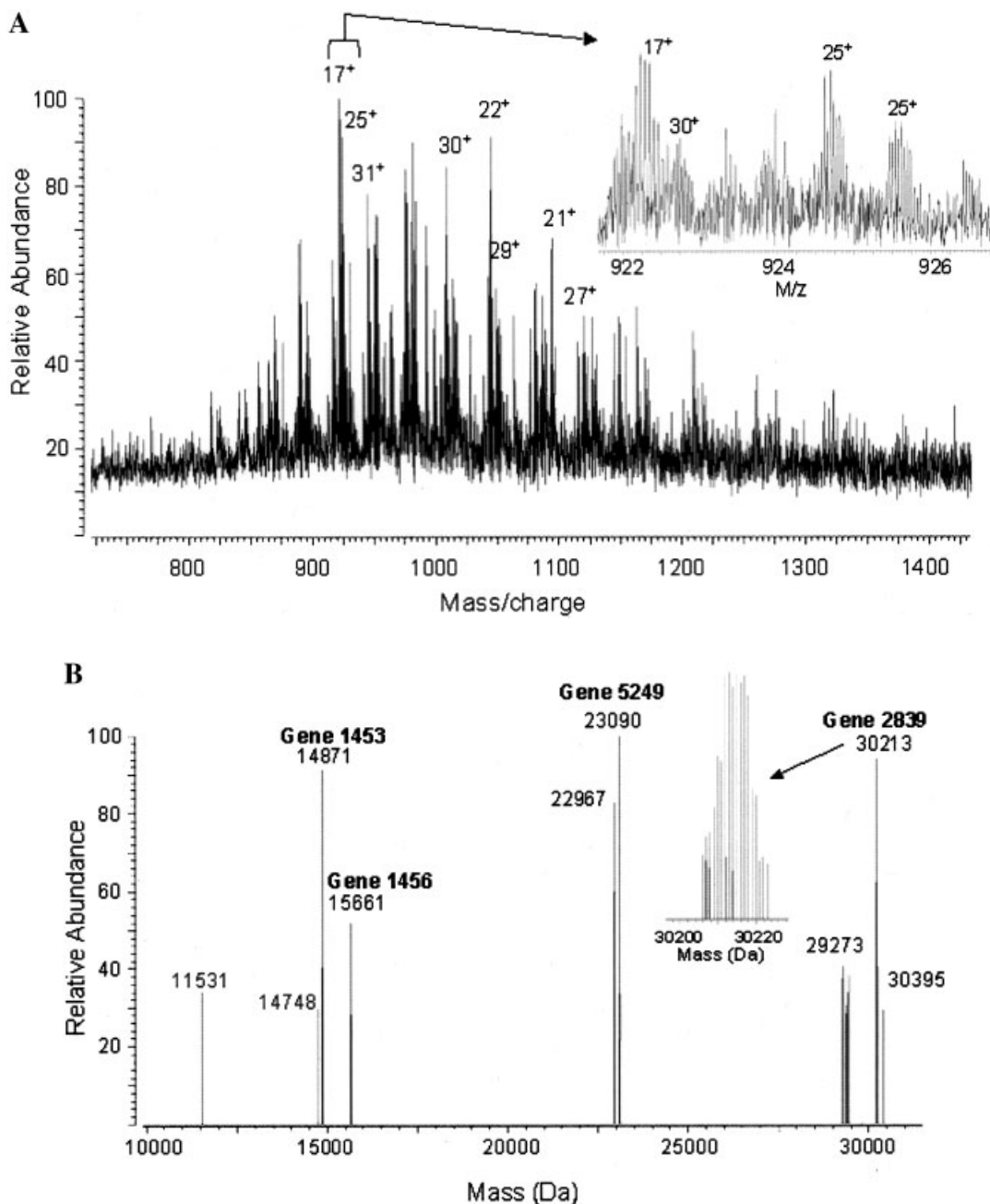
FIGURE 16. Mass spectra of FPLC fraction no. 23 (intact proteins) from *Shewanella oneidensis*. **A**: ESI/ FTICR MS revealing a complex mixture of proteins (close-up of *m/z* 920–927 region shown as an inset). **B**: Deconvoluted molecular mass spectrum revealing the presence of approximately one dozen proteins detected in this sample. In each case, an isotopically resolved cluster was measured (as shown in the inset). High-resolution mass measurement was sufficient for identification of at least four proteins in this sample, as designated by the gene number labels. (Adapted from VerBerkmoes et al., 2002, with permission from American Chemical Society, copyright 2002.)

B. Medical Applications

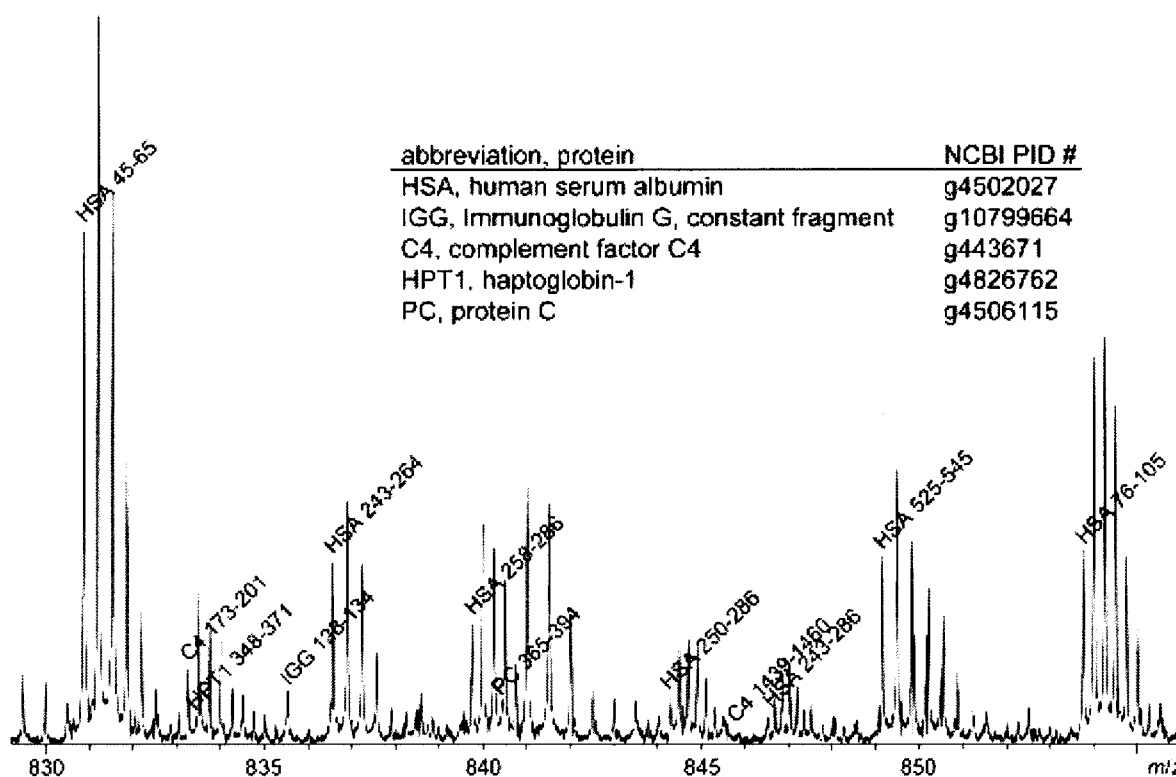The explosion of proteomics research during the last 5 years has not only focused on bacterial organisms, but also on mammalian organisms, including human. A proteomics approach for studying diseases seems attractive, since certain proteins or modified forms may act as a biomarker for a disease. Surprisingly though, very few studies using FTICR to identify human clinical

biomarkers have been reported. Bergquist et al. (2002) used FTICR to identify and characterize proteins in four human body fluids: plasma, CSF, saliva, and urine. Collected samples were tryptically digested and desalted before direct infusion, followed by ESI ion generation and detection with a 9.4 T FTICR. In the mass spectra, 2,745, 1,414, 624, and 1,183 peaks from plasma, CSF, saliva, and urine, respectively, were detected, which corresponded to 1,242, 585, 359, and 654 isotopic clusters, respectively. Approximately 50% of all peaks could be assigned to a specific protein in the body fluid database that contained 150 proteins (mass error <5 ppm). The spectra were obtained by summing 2,048 transients (512 kB); ~1.7 µL of sample was used during collection. Figure 17 shows a small region of the plasma spectrum, including some peptide and corresponding protein identifications. In the plasma, CSF, saliva, and urine samples, 5, 6, 2, and 8 proteins, respectively, could be identified with 99% confidence, whereas 4, 7, 6, and 5 proteins, respectively, could be identified with 95–99% confidence. The most significant matches were human serum albumin (in plasma, CSF, and urine), and amylase (in saliva). Expectations are that the state-of-the-art instrumentation available today will be used more frequently in medical proteomics applications, and that the wealth of new information may shed light on how the medicinal community looks at diseases and searches for improved or new ways to detect

and cure them. In our laboratory the AMT tag approach is being used in a broad range of biomedical collaborations to study the blood plasma proteome, various human cell lines, tissue samples, and pathogens such as *Yersinia pestis* (plague), among others.

VI. OUTLOOK

Even though the field of proteomics has seen an explosive growth over the last 5 years, it is still in its relative infancy. Currently, FTICR along with the AMT tag approach appears to be the most powerful technique for large-scale proteome analyses. However, the AMT tag approach can also be implemented using LC-TOF MS instrumentation, as recently shown for *D. radiodurans* (Strittmatter et al., 2003). Regardless, FTICR provides the most powerful tool for peptide/protein characterization yet available. However, it is not as routinely used and as broadly accepted as other types of MS. This situation is changing, particularly as instrumentation with advanced capabilities becomes more broadly available and as the technology continues to improve. For example, improvements in both resolution of the DREAMS ion ejection step and variable data-directed ion accumulation time (AGC) will further improve dynamic range and MMA (Belov

abbreviation, protein	NCBI PID #
HSA, human serum albumin	g4502027
IGG, Immunoglobulin G, constant fragment	g10799664
C4, complement factor C4	g443671
HPT1, haptoglobin-1	g4826762
PC, protein C	g4506115

FIGURE 17. Detailed view of a region of a plasma tryptic digest spectrum with some of the peaks putatively identified from the small subset of the NCBI protein database. As can be seen, there is much overlap between isotopic clusters in the densest regions of the spectra, sometimes with as many as four to five overlapping clusters of isotopic peaks. (Adapted from Bergquist et al., 2002, with permission from John Wiley & Sons, Inc., 2002.)

et al., 2003a,b). Shorter ion cooling times will also improve the duty cycle, and contribute to the sensitivity and dynamic range of this technique. The recent introduction of the Finnigan LTQ FT MS instrument that incorporates AGC in both the MS and MS/ MS modes is a key advance for commercial instrumentation. Other advances such as improved identification of low abundance peptides by multiplexed-MS/MS will increase the throughput and overall effectiveness of this approach. It seems likely that FTICR will play an important role in proteomics for the near future, and will continue to benefit from advances in other fields.

While a bright future lies ahead for FTICR, especially as magnetic field strengths increase, increased competition from other mass spectrometric techniques such as TOF MS is expected. TOF MS has some advantages over FTICR, such as in spectrum acquisition rate and price. It seems certain that new hybrid FTICRs will be developed that incorporate TOF, ion mobility, or other novel combinations. In addition, development of automated, multi-task FTICR instruments that incorporate data stations that much more effectively control various ion dissociation and decision-making techniques for high-throughput analysis environments can be anticipated. We also anticipate the future introduction of multi-trap instruments that effectively exploit the non-destructive nature of FTICR detection, enabling real time selection of trapped species for sequential dissociation by multi-stage MS^n. Ultimately, such experiments will be multiplexed, perhaps also in multi-cell arrangements, thereby greatly increasing the effective sensitivity and throughput of analyses. Developments in FTICR theory and techniques remain critical to the advancement of this powerful technique to meet future analytical challenges.

ACKNOWLEDGMENTS

REFERENCES

Baykut G, Jertz R, Witt M. 2000. Matrix-assisted laser desorption/ionization Fourier transform ion cyclotron resonance mass spectrometry with pulsed in-source collision gas and in-source ion accumulation. Rapid Commun Mass Spectrom 14:1238–1247.

Belov ME, Anderson GA, Smith RD. 2002. Higher-resolution data-dependent selective external ion accumulation for capillary LC-FTICR. Int J Mass Spectrom 218:265–279.

Belov ME, Gorshkov MV, Udseth HR, Anderson GA, Smith RD. 2000. Zeptomole-sensitivity electrospray ionization Fourier-transform ion cyclotron resonance mass spectrometry of proteins. Anal Chem 72:2271–2279.

Belov ME, Gorshkov MV, Udseth HR, Smith RD. 2001a. Controlled ion fragmentation in a 2D quadrupole ion trap for external ion accumulation in ESI FTICR mass spectrometry. J Am Soc Mass Spectrom 12:1312–1319.

Belov ME, Anderson GA, Angell NH, Shen Y, Tolić N, Udseth HR, Smith RD. 2001b. Dynamic range expansion applied to mass spectrometry based on data-dependent selective ion ejection in capillary liquid chromatography Fourier transform ion cyclotron resonance for enhanced proteome characterization. Anal Chem 73:5052–5060.

Belov ME, Rakov VS, Nikolaev EN, Goshe MB, Anderson GA, Smith RD. 2003a. Initial implementation of external accumulation liquid chromatography/electrospray ionization Fourier transform ion cyclotron resonance with automated gain control. Rapid Commun Mass Spectrom 17:627–636.

Belov ME, Zhang R, Strittmatter EF, Prior DC, Tang K, Smith RD. 2003b. Automated gain control and internal calibration with external ion accumulation capillary liquid chromatography-electrospray ionization-Fourier transform ion cyclotron resonance. Anal Chem 75:4195–4205.

Berger SJ, Lee SW, Anderson GA, Paša-Tolić L, Tolić N, Shen Y, Zhao R, Smith RD. 2002. High-throughput global peptide proteomic analysis by combining stable isotope amino acid labeling and data-dependent multiplexed-MS/MS. Anal Chem 74:4994–5000.

Bergquist J, Palmblad M, Wetterhall M, Håkansson P, Markides KE. 2002. Peptide mapping of proteins in human body fluids using electrospray ionization Fourier transform ion cyclotron resonance-mass spectrometry. Mass Spectrom Rev 21:2–15.

Bossio RE, Marshall AG. 2002. Baseline resolution of isobaric phosphorylated and sulfated peptides and nucleotides by electrospray ionization FTICR MS: Another step toward mass spectrometry-based proteomics. Anal Chem 74:1674–1679.

Breuker K, Oh HB, Horn DM, Cerda BA, McLafferty FW. 2002. Detailed unfolding and folding of gaseous ubiquitin ions characterized by electron capture dissociation. J Am Chem Soc 124:6407–6420.

Bruce JE, Anderson GA, Smith RD. 1996. Colored noise waveforms and quadrupole excitation for the dynamic range expansion of Fourier transform ion cyclotron resonance mass spectrometry. Anal Chem 68:534–541.

Bruce JE, Anderson GA, Udseth HR, Smith RD. 1998. Large molecule characterization based upon individual ion detection with electrospray ionization FT-ICR mass spectrometry. Anal Chem 70:519–525.

Bruce JE, Anderson GA, Brands MD, Paša-Tolić L, Smith RD. 2000. Obtaining more accurate Fourier transform ion cyclotron resonance mass measurements without internal standards using multiply charged ions. J Am Soc Mass Spectrom 11:416–421.

Chan T-WD, Duan L, Sze T-PE. 2002. Accurate mass measurements for peptide and protein mixtures by using matrix-assisted laser desorption/ionization Fourier transform mass spectrometry. Anal Chem 74:5282–5289.

Chorush RA, Little DP, Beu SC, Wood TD, McLafferty FW. 1995. Surface-induced dissociation of multiply protonated proteins. Anal Chem 67:1042–1046.

Conrads TP, Alving K, Veenstra TD, Belov ME, Anderson GA, Anderson DJ, Lipton MS, Paša-Tolić L, Udseth HR, Chrisler WB, Thrall BD, Smith RD. 2001. Quantitative analysis of bacterial and mammalian proteomes using a combination of cysteine affinity tags and ^{15}N-metabolic labeling. Anal Chem 73:2132–2139.

Demirev PA, Fenselau C. 2002. Determination of isotope-enrichment in proteins by high-resolution Fourier-transform ion cyclotron resonance mass spectrometry. Eur J Mass Spectrom 8:163–167.

Dienes T, Pastor SJ, Schürch S, Scott JR, Yao J, Cui S, Wilkins CL. 1996. Fourier transform mass spectrometry—Advancing years (1992-mid 1996). Mass Spectrom Rev 15:163–211.

Easterling ML, Mize TH, Amster IJ. 1999. Routine part-per-million mass accuracy for high-mass ions: Space-charge effects in MALDI FT-ICR. Anal Chem 71:624–632.

Eng JK, McCormack AL, Yates JR. 1994. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spectrom 5:976–989.

Ewing NP, Cassady CJ. 2000. Dissociation of multiply charged negative ions for hirudin (54–65), fibrinopeptide B, and insulin A (oxidized). J Am Soc Mass Spectrom 12:105–116.

Eyles SJ, Speir JP, Kruppa GH, Gierasch LM, Kaltashov IA. 2000. Protein conformational stability probed by Fourier transform ion cyclotron resonance mass spectrometry. J Am Chem Soc 122:495–500.

Felitsyn N, Kitova EN, Klassen JS. 2002. Thermal dissociation of the protein homodimer ecotin in the gas phase. J Am Soc Mass Spectrom 13:1432–1442.

Flora JW, Hannis JC, Muddiman DC. 2001. High-mass accuracy of product ions produced by SORI-CID using a dual electrospray ionization source coupled with FTICR mass spectrometry. Anal Chem 73:1247–1251.

Flora JW, Muddiman DC. 2001. Selective, sensitive, and rapid phosphopeptide identification in enzymatic digests using ESI-FTICR-MS with infrared multiphoton dissociation. Anal Chem 73:3305–3311.

Flora JW, Muddiman DC. 2002. Gas-phase ion unimolecular dissociation for rapid phosphopeptide mapping by IRMPD in a penning ion trap: An energetically favored process. J Am Chem Soc 124:6546–6547.

Forbes AJ, Mazur MT, Patel HM, Walsh CT, Kelleher NL. 2001. Towards efficient analysis of >70 kDa proteins with 100% sequence coverage. Proteomics 1:927–933.

Freitas MA, Hendrickson CL, Marshall AG. 1999a. Gas phase activation energy for unimolecular dissociation of biomolecular ions determined by Focused RAdiation for Gaseous Multiphoton ENergy Transfer (FRAGMENT). Rapid Commun Mass Spectrom 13:1639–1642.

Freitas MA, King E, Shi SD-H. 2003. Tool command language automation of the modular ion cyclotron data acquisition system (MIDAS) for data-dependent tandem Fourier transform ion cyclotron resonance mass spectrometry. Rapid Commun Mass Spectrom 17:363–370.

Freitas MA, Hendrickson CL, Emmett MR, Marshall AG. 1999b. Gas-phase bovine ubiquitin cation conformations resolved by gas-phase hydrogen/deuterium exchange rate and extend. Int J Mass Spectrom 185–187:565–575.

Freitas MA, Hendrickson CL, Marshall AG, Rostom AA, Robinson CV. 2000. Competitive binding to the oligopeptide binding protein, OppA: In-trap cleanup in an Fourier transform ion cyclotron resonance mass spectrometer. J Am Soc Mass Spectrom 11:1023–1026.

Fridriksson EK, Beavil A, Holowka D, Gould HJ, Baird B, McLafferty FW. 2000. Hetrogeneous glycosylation of immunoglobin E constructs characterized by top-down high-resolution 2-D mass spectrometry. Biochemistry 39:3369–3376.

Gauthier JW, Trautman TR, Jaconson DB. 1991. Sustained off-resonance irradiation for collision-activated dissociation involving Fourier transform mass spectrometry. Collision-activated dissociation technique that emulates infrared multiphoton dissociation. Anal Chim Acta 246:211–225.

Ge Y, Lawhorn BG, ElNaggar M, Strauss E, Park J-H, Begley TP, McLafferty FW. 2002. Top down characterization of larger proteins (45 kDa) by electron capture dissociation mass spectrometry. J Am Chem Soc 124:672–678.

Ge Y, ElNaggar M, Sze SK, Bin Oh H, Begley TP, McLafferty FW, Boshoff H, Barry CE. 2003. Top down characterization of secreted proteins from *Mycobacterium tuberculosis* by electron capture dissociation mass spectrometry. J Am Soc Mass Spectrom 14:253–261.

Green MK, Lebrilla CB. 1997. Ion-molecule reactions as probes of gas phase structures of peptides and proteins. Mass Spectrom Rev 16:53–71.

Håkansson K, Emmett MR, Hendrickson CL, Marshall AG. 2001. High-sensitivity electron capture dissociation tandem FTICR mass spectrometry of microelectrosprayed peptides. Anal Chem 73:3605–3610.

Hannis JC, Muddiman DC. 2000. A dual electrospray ionization source combined with hexapole accumulation to achieve high mass accuracy of biopolymers in Fourier transform ion cyclotron mass spectrometry. J Am Soc Mass Spectrom 11:876–883.

Haselmann KF, Jørgensen TJD, Budnik BA, Jensen F, Zubarev RA. 2002. Electron capture dissociation of weakly bound polypeptide polycationic complexes. Rapid Commun Mass Spectrom 16:2260–2265.

He F, Hendrickson CL, Marshall AG. 2000. Unequivocal determination of metal atom oxidation state in naked heme proteins: Fe(III)myoglobin, Fe(III)cytochrome c, Fe(III)cytochrome b5, and Fe(III)cytochrome b5 L47R. J Am Soc Mass Spectrom 11:120–126.

He F, Hendrickson CL, Marshall AG. 2001. Baseline mass resolution of peptide isobars: A record for molecular mass resolution. Anal Chem 73:647–650.

Hill TJ, Lafitte D, Wallace JI, Cooper HJ, Tsvetkov PO, Derrick PJ. 2000. Calmodulin-peptide interactions: Apocalmodulin binding to the myosin light chain kinase target-site. Biochemistry 39:7284–7290.

Hofstadler SA, Sannes-Lowery KA, Griffey RH. 2000. Enhanced gas-phase hydrogen-deuterium exchange of oligonucleotide and protein ions stored in an external multipole ion reservoir. J Mass Spectrom 35:62–70.

Hofstadler SA, Wahl JH, Bruce JE, Smith RD. 1993. On-line capillary electrophoresis with Fourier transform ion cyclotron resonance mass spectrometry. J Am Chem Soc 115:6983–6984.

Holliman CL, Rempel DL, Gross ML. 1994. Detection of high mass-to-charge ions by FTMS. Mass Spectrom Rev 13:105–132.

Horn DM, Zubarev RA, McLafferty FW. 2000a. Automated reduction and interpretation of high resolution electrospray mass spectra of large molecules. J Am Soc Mass Spectrom 11:320–332.

Horn DM, Zubarev RA, McLafferty FW. 2000b. Automated *de novo* sequencing of proteins by tandem high-resolution mass spectrometry. Proc Natl Acad Sci USA 97:10313–10317.

Iavarone AT, Jurchen JC, Williams ER. 2001. Supercharged protein and peptide ions formed by electrospray ionization. Anal Chem 73:1455–1460.

Ijames CF, Wilkins CL. 1990. Surface-induced dissociation by Fourier transform mass spectrometry. Anal Chem 62:1295–1299.

Jensen PK, Paša-Tolić L, Anderson GA, Horner JA, Lipton MS, Bruce JE, Smith RD. 1999. Probing proteomes using capillary isoelectric focusing-electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. Anal Chem 71:2076–2084.

Jensen P, Paša-Tolić L, Peden KK, Martinović S, Lipton MS, Anderson GA, Tolić N, Wong K-K, Smith RD. 2000. Mass spectrometric detection for capillary isoelectric focusing separations of complex protein mixtures. Electrophoresis 21:1372–1380.

Johnson KA, Verhagen MFJM, Brereton PS, Adams MWW, Amster IJ. 2000. Probing the stoichiometry and oxidation states of metal centers in iron-sulfur proteins using electrospray FTICR mass spectrometry. Anal Chem 72:1410–1418.

Johnson KA, Shira BA, Anderson JL, Amster IJ. 2001a. Chemical and on-line electrochemical reduction of metalloproteins with high-resolution electrospray ionization mass spectrometry detection. Anal Chem 73:803–808.

Johnson KA, Verhagen MFJM, Adams MWW, Amster IJ. 2001b. Differences between positive and negative ion stabilities of metal-sulfur cluster proteins: An electrospray ionization Fourier transform ion cyclotron resonance study. Int J Mass Spectrom 204:77–85.

Johnson JR, Meng F, Forbes AJ, Cargile BJ, Kelleher NL. 2002. Fourier-transform mass spectrometry for automated fragmentation and identification of 5–20 kDa proteins in mixtures. Electrophoresis 23:3217–3223.

Kelleher NL, Senko MW, Siegel MM, McLafferty FW. 1997. Unit resolution mass spectra of 112 kDa molecules with 3 Da accuracy. J Am Soc Mass Spectrom 8:380–383.

Kelleher NL, Taylor SV, Grannis D, Kinsland C, Chiu H-J, Begley TP, McLafferty FW. 1998. Efficient sequence analysis of the six gene products (7–74 kDa) of the *Escherichia coli* thiamin biosynthetic operon by tandem high-resolution mass spectrometry. Prot Science 7:1796–1801.

Kelleher NL, Lin HY, Valaskovic GA, Aaserud DJ, Fridriksson EK, McLafferty FW. 1999a. Top down versus bottom up protein characterization by tandem high-resolution mass spectrometry. J Am Chem Soc 121:806–812.

Kelleher NL, Zubarev RA, Bush K, Furie B, Furie BC, McLafferty FW, Walsh CT. 1999b. Localization of labile posttranslational modifications by electron capture dissociation: The case of γ-carboxyglutamic acid. Anal Chem 71:4250–4253.

Kjeldsen F, Haselmann KF, Budnik BA, Jensen F, Zubarev RA. 2002. Dissociative capture of hot (3–13 eV) electrons by polypeptide polycations: An efficient process accompanied by secondary fragmentation. Chem Phys Lett 356:201–206.

Kjeldsen F, Haselmann KF, SØrensen ES, Zubarev RA. 2003. Distinguishing of Ile/Leu amino acid residues in the PP3 protein by (hot) electron capture dissociation in Fourier transform ion cyclotron resonance mass spectrometry. Anal Chem 75:1267–1274.

Kruppa GH, Schoeniger J, Young MM. 2003. A top down approach to protein structural studies using chemical cross-linking and Fourier transform mass spectrometry. Rapid Commun Mass Spectrom 17:155–162.

Kruppa G, Schnier PD, Tabei K, Van Orden S, Siegel MM. 2002. Multiple ion isolation applications in FT-ICR MS: Exact-mass MS^n internal calibration and purification/interrogation of protein–drug complexes. Anal Chem 74:3877–3886.

Lam TT, Lanman JK, Emmett MR, Hendrickson CL, Marshall AG, Prevelige PE. 2002. Mapping of protein:protein contact surfaces by hydrogen/deuterium exchange, followed by on-line high-performance liquid chromatography-electrospray ionization Fourier-transform ion cyclotron resonance mass analysis. J Chromatography A 982:85–95.

Laskin J, Denisov E, Futrell JH. 2000. A comparative study of collision-induced and surface-induced dissociation. 1. Fragmentation of protonated dianaline. J Am Chem Soc 122:9703–9714.

Laskin J, Denisov E, Futrell JH. 2001. A comparative study of collision-induced and surface-induced dissociation. 2. Fragmentation of small analine-containing peptides in FT-ICR MS. J Phys Chem B 105:1895–1900.

Laskin J, Denisov EV, Shukla AK, Barlow SE, Futrell JH. 2002. Surface-induced dissociation in a Fourier transform ion cyclotron resonance mass spectrometer: Instrument design and evaluation. Anal Chem 74:3255–3261.

Lee S-W, Kim HS, Beauchamp JL. 1998. Salt bridge chemistry applied to gas-phase peptide sequencing: Selective fragmentation of sodiated gas-phase peptide ions adjacent to aspartic acid residues. J Am Chem Soc 120:3188–3195.

Lee S-W, Berger SJ, Martinović S, Paša-Tolić L, Anderson GA, Shen Y, Zhao R, Smith RD. 2002. Direct mass spectrometric analysis of intact proteins of the yeast large ribosomal subunit using capillary LC-FTICR. Proc Natl Acad Sci USA 99:5942–5947.

Leymarie N, Berg EA, McComb ME, O'Connor PB, Grogan J, Oppenheim FG, Costello CE. 2002. Tandem mass spectrometry for structural characterization of proline-rich proteins: Application to salivary PRP-3. Anal Chem 74:4124–4132.

Li J, Taraska JA, Counterman AE, Clemmer DE. 1999a. Influence of solvent composition and capillary temperature on the conformations of electrosprayed ions: Unfolding of compact ubiquitin conformers from pseudo-native and denatured solutions. Int J Mass Spectrom 185–187:37–47.

Li W, Hendrickson CL, Emmett MR, Marshall AG. 1999b. Identification of intact proteins in mixtures by alternated capillary liquid chromatography electrospray ionization and LC ESI infrared multiphoton dissociation Fourier transform ion cyclotron resonance mass spectrometry. Anal Chem 71:4397–4402.

Li L, Masselon CD, Anderson GA, Paša-Tolić L, Lee S-W, Shen Y, Zhao R, Lipton MS, Conrads TP, Tolić N, Smith RD. 2001. High-throughput peptide identification from protein digests using data-dependent multiplexed tandem FTICR mass spectrometry coupled with capillary liquid chromatography. Anal Chem 73:3312–3322.

Lipton MS, Paša-Tolić L, Anderson GA, Anderson DJ, Auberry DL, Battista JR, Daly MJ, Fredrickson J, Hixson KK, Kostandarithes H, Masselon C, Markillie LM, Moore RJ, Romine MF, Shen Y, Stritmatter E, Tolić N, Udseth HR, Venkateswaran A, Wong K-K, Zhao R, Smith RD. 2002. Global analysis of the *Deinococcus radiodurans* proteome by using accurate mass tags. Proc Natl Acad Sci USA 99:11049–11054.

Little DP, Speir JP, Senko MW, O'Connor PB, McLafferty FW. 1994. Infrared multiphoton dissociation of large multiply charged ions for biomolecule sequencing. Anal Chem 66:2809–2815.

Loo JA. 1997. Studying noncovalent protein complexes by electrospray ionization mass spectrometry. Mass Spectrom Rev 16:1–23.

Marshall AG. 1996. Ion cyclotron and nuclear magnetic resonance spectroscopies: Magnetic partners for elucidation of molecular structure and reactivity. Acc Chem Res 9:307–316.

Marshall AG. 2000. Milestones in Fourier transform ion cyclotron resonance mass spectrometry technique developments. Int J Mass Spectrom 200:331–356.

Marshall AG, Hendrickson CL, Jackson GS. 1998. Fourier transform ion cyclotron resonance mass spectrometry: A primer. Mass Spectrom Rev 17:1–35.

Marshall AG, Guan S. 1996. Advantages of high magnetic field for FT-ICR mass spectrometry. Rapid Commun Mass Spectrom 10:1819–1823.

Marshall AG, Hendrickson CL. 2001. Charge reduction lowers mass resolving power for isotopically resolved electrospray ionization Fourier transform ion cyclotron resonance mass spectra. Rapid Commun Mass Spectrom 15:232–235.

Martin SE, Shabanowitz J, Hunt DF, Marto JA. 2000. Subfemtomole MS and MS/MS peptide sequence analysis using nano-HPLC micro-ESI Fourier transform ion cyclotron resonance mass spectrometry. Anal Chem 72:4266–4274.

Martinović S, Berger SJ, Paša-Tolić L, Smith RD. 2000. Separation and detection of intact noncovalent protein complexes from mixtures by on-line capillary isoelectric focusing-mass spectrometry. Anal Chem 72:5356–5360.

Martinović S, Veenstra TD, Anderson GA, Paša-Tolić L, Smith RD. 2002. Selective incorporation of isotopically labeled amino acids for identification of intact proteins on a proteome-wide level. J Mass Spectrom 37:99–107.

Masselon C, Anderson GA, Harkewicz R, Bruce JE, Paša-Tolić L, Smith RD. 2000. Accurate mass multiplexed tandem mass spectrometry for high-throughput polypeptide identification from mixtures. Anal Chem 72:1918–1924.

Masselon C, Tolmachev AV, Anderson GA, Harkewicz R, Smith RD. 2002. Mass measurement error caused by local frequency perturbations in FTICR mass spectrometry. J Am Soc Mass Spectrom 13:99–106.

McCormack AL, Schieltz DM, Goode B, Yang S, Barnes G, Drubin D, Yates JR. 1997. Direct analysis and identification of proteins in mixtures by LC/MS/MS and database searching at the low-femtomole level. Anal Chem 69:767–776.

McLafferty FW. 1994. High-resolution tandem FT mass spectrometry above 10 kDa. Acc Chem Res 94:2161–2182.

McLafferty FW, Guan Z, Haupts U, Wood TD, Kelleher NL. 1998. Gaseous conformational structures of cytochrome c. J Am Chem Soc 120:4732–4740.

Meng F, Cargile BJ, Miller LM, Forbes AJ, Johnson JR, Kelleher NL. 2001. Informatics and multiplexing of intact protein identification in bacteria and the archaea. Nat Biotech 19:952–957.

Meng F, Cargile BJ, Patrie SM, Johnson JR, McLoughlin SM, Kelleher NL. 2002. Processing complex mixtures of intact proteins by direct analysis by mass spectrometry. Anal Chem 74:2923–2929.

Mirgorodskaya E, O'Connor PB, Costello CE. 2002. A general method for precalculation of parameters for sustained off-resonance irradiation/collision-induced dissociation. J Am Soc Mass Spectrom 13:318–324.

Nousiainen M, Vainiotalo P, Feng X, Derrick PJ. 2001. Calmodulin–RS20-Ca$_4$ complex in the gas phase: Electrospray ionization and Fourier transform ion cyclotron resonance. Eur J Mass Spectrom 7:393–398.

Paša-Tolić L, Bruce JE, Lei QP, Anderson GA, Smith RD. 1998. In-trap cleanup of proteins from electrospray ionization using soft sustained off-resonance irradiation with Fourier transform ion cyclotron resonance mass spectrometry. Anal Chem 70:405–408.

Paša-Tolić L, Harkewicz R, Anderson GA, Tolić N, Shen Y, Zhao R, Thrall B, Masselon CD, Smith RD. 2002a. Increased proteome coverage for quantitative peptide abundance measurements based upon high performance separations and DREAMS FTICR mass spectrometry. J Am Soc Mass Spectrom 13:954–963.

Paša-Tolić L, Lipton MS, Masselon CD, Anderson GA, Shen Y, Tolić N, Smith RD. 2002b. Gene expression profiling using advanced mass spectrometric approaches. J Mass Spectrom 37:1185–1198.

Palmblad M, Tsybin YO, Ramström M, Bergquist J, Håkansson P. 2002a. Liquid chromatography and electron-capture dissociation in Fourier transform ion cyclotron resonance mass spectrometry. Rapid Commun Mass Spectrom 16:988–992.

Palmblad M, Ramström M, Markides KE, Håkansson P, Bergquist J. 2002b. Prediction of chromatographic retention and protein identification in liquid chromatography/mass spectrometry. Anal Chem 74:5826–5830.

Perkins DN, Pappin DJC, Creasy DM, Cottrell JS. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20:3551–3567.

Peters EC, Brock A, Horn DM, Phung QT, Ericson C, Salomon AR, Ficarro SB, Brill LM. 2002. An automated LC-MALDI FT-ICR MS platform for high-throughput proteomics. LC GC Europe 7:2–7.

Petritis K, Kangas LJ, Ferguson PL, Anderson GA, Paša-Tolić L, Lipton MS, Auberry KJ, Strittmatter EF, Shen Y, Zhao R. 2003. Use of artificial neural networks for the accurate prediction of peptide liquid chromatography elution times in proteome analyses. Anal Chem 75:1039–1048.

Price WD, Schnier PD, Williams ER. 1996. Tandem mass spectrometry of large biomolecular ions by blackbody infrared radiative dissociation. Anal Chem 68:859–866.

Quenzer TL, Emmett MR, Hendrickson CL, Kelly PH, Marshall AG. 2001. High sensitivity Fourier transform ion cyclotron resonance mass spectrometry for biological analysis with nano-LC and microelectrospray ionization. Anal Chem 73:1721–1725.

Reid GE, McLuckey SA. 2002. Top down protein characterization via tandem mass spectrometry. J Mass Spectrom 37:663–675.

Rodgers RP, Blumer EN, Hendrickson CL, Marshall AG. 2000. Stable isotope incorporation triples the upper mass limit for determination of elemental composition by accurate mass measurement. J Am Soc Mass Spectrom 11:835–840.

Sannes-Lowry KA, Hofstadler SA. 2000. Characterization of multipol storage assisted dissociation: Implications for electrospray ionization mass spectrometry characterization of biomolecules. J Am Soc Mass Spectrom 11:1–9.

Shen Y, Tolić N, Zhao R, Paša-Tolić L, Li L, Berger SJ, Harkewicz R, Anderson GA, Belov ME, D. Smith RD. 2001. High-throughput proteomics using high-efficiency multiple-capillary liquid chromatography with on-line high-performance ESI FTICR mass spectrometry. Anal Chem 73:3011–3021.

Shen Y, Tolić N, Masselon CD, Paša-Tolić L, Camp DG II, Lipton MS, Anderson GA, Smith RD. 2004. Nanoscale proteomics. Anal and Bioanal Chem 378:1037–1045.

Shi SD-H, Hendrickson CL, Marshall AG. 1998. Counting individual sulfur atoms in a protein by ultrahigh-resolution Fourier transform ion cyclotron resonance mass spectrometry: Experimental resolution of isotopic fine structure in proteins. Proc Natl Acad Sci USA 95:11532–11537.

Shi SD-H, Hemling ME, Carr SA, Horn DM, Lindh I, McLafferty FW. 2001. Phosphopeptide/phosphoprotein mapping by electron capture dissociation mass spectrometry. Anal Chem 73:19–22.

Smith RD, Shen Y, Tang K. 2004. Ultra-sensitive and quantitative analyses from combined separations-mass spectrometry for the characterization of proteomes. Acc Chem Res web release 02/25/04.

Stensballe A, NØrregaard Jensen O, Olsen JV, Haselmann KF, Zubarev RA. 2000. Electron capture dissociation of singly and multiply phosphorylated peptides. Rapid Commun Mass Spectrom 14:1793–1800.

Strittmatter EF, Ferguson PL, Tang K, Smith RD. 2003. Proteome analysis using accurate mass and elution time peptide tags with capillary LC time-of-flight mass spectrometry. J Am Soc Mass Spectrom 14:980–991.

Suckau D, Shi Y, Beu SC, Senko MW, Quinn JP, Wampler FM, McLafferty FW. 1993. Coexisting stable conformations of gaseous protein ions. Proc Natl Acad Sci USA 90:790–793.

Sze SK, Ge Y, Oh HB, McLafferty FW. 2002. Top-down mass spectrometry of a 29-kDa protein for characterization of any posttranslational modification to within one residue. Proc Natl Acad Sci USA 99:1774–1779.

Sze SK, Ge Y, Oh HB, McLafferty FW. 2003. Plasma electron capture dissociation for the characterization of large protein by top down mass spectrometry. Anal Chem 75:1599–1603.

Taylor OK, Amster IJ. 2003. Space charge effects on mass accuracy for multiply charged ions in ESI–FTICR. Int J Mass Spectrom 222:351–361.

Tsybin YO, Håkansson P, Budnik BA, Haselmann KF, Kjeldsen F, Gorshkov M, Zubarev RA. 2001. Improved low-energy electron injection systems for high rate electron capture dissociation in Fourier transform ion cyclotron resonance mass spectrometry. Rapid Commun Mass Spectrom 15:1849–1854.

Tsybin YO, Håkansson P, Wetterhall M, Markides KE, Bergquist J. 2002. Capillary electrophoresis and electron capture dissociation Fourier transform ion cyclotron resonance mass spectrometry for peptide mixture and protein digest analysis. Eur J Mass Spectrom 8:389–395.

Uchiki T, Hettich R, Gupta V, Dealwis C. 2002. Characterization of monomeric and dimeric forms of recombinant Sml1p-histag protein by electrospray mass spectrometry. Anal Biochem 301:35–48.

Valaskovic GA, Kelleher NL, McLafferty FW. 1996. Attomole protein characterization by capillary electrophoresis-mass spectrometry. Science 273:1199–1292.

Veenstra TD, Martinović S, Anderson GA, Paša-Tolić L, Smith RD. 2000. Proteome analysis using selective incorporation of isotopically labeled amino acids. J Am Soc Mass Spectrom 11:78–82.

VerBerkmoes NC, Bundy JL, Hauser L, Asano KG, Razumovskaya J, Larimer F, Hettich RL, Stephenson JL Jr. 2002. Integrating top-down and bottom-up mass spectrometric approaches for proteomic analysis of *Shewanella oneidensis*. J Proteome Res 1:239–252.

Wetterhall M, Palmblad M, Håkansson P, Markides KE, Bergquist J. 2002. Rapid analysis of tryptically digested cerebrospinal fluid using capillary electrophoresis-electrospray ionization-Fourier transform ion cyclotron resonance-mass spectrometry. J Prot Res 1:361–366.

Wigger M, Eyler JR, Benner SA, Li W, Marshall AG. 2002. Fourier transform-ion cyclotron resonance mass spectrometric resolution, identification, and screening of non-covalent complexes of Hck Src homology 2 domain receptor and ligands from a 324-member peptide combinatorial library. J Am Soc Mass Spectrom 13:1162–1169.

Wu QY, Vanorden S, Cheng XH, Bakhtiar R, Smith RD. 1995. Characterization of cytochrome c variants with high-resolution FTICR mass spectrometry: Correlation of fragmentation and structure. Anal Chem 67:2498–2509.

Xi J, Ge Y, Kinsland C, McLafferty FW, Begley TP. 2001. Biosynthesis of the thiazole moiety of thiamin in *Escherichia coli*: Identification of an acyldisulfide-linked protein–protein conjugate that is functionally analogous to the ubiquitin/E1 complex. Proc Natl Acad Sci USA 98:8513–8518.

Yamada N, Suzuki E, Hirayama K. 2002. Identification of the interface of a large protein–protein complex using H/D exchange and Fourier transform ion cyclotron resonance mass spectrometry. Rapid Commun Mass Spectrom 16:293–299.

Yao X, Freas A, Ramirez J, Demirev PA, Fenselau C. 2001. Proteolytic ^{18}O labeling for comparative proteomics: Model studies with two serotypes of adenovirus. Anal Chem 73:2836–2842.

Yates JR, McCormack AL, Eng J. 1996. Mining genomes with MS. Anal Chem 68:534A–540A.

Yu L-R, Johnson MD, Conrads TP, Smith RD, Morrison RS, Veenstra TD. 2002. Proteome analysis of camptothecin-treated cortical neurons using isotope-coded affinity tags. Electrophoresis 23:1591–1598.

Zhang Z, Marshall AG. 1998. A universal algorithm for fast and automated charge state deconvolution of electrospray mass-to-charge ratio spectra. J Am Soc Mass Spectrom 9:225–233.

Zubarev RA, Kelleher NL, McLafferty FW. 1998. Electron capture dissociation of multiply charge protein cations: A nonergodic process. J Am Chem Soc 120:3265–3266.

Bogdan Bogdanov has been a Postdoctoral Research Fellow in Dr. Richard D. Smith's group since 2001. He has been working on the analysis of proteome samples by reversed phase liquid chromatography coupled to electrospray ionization Fourier transform ion cyclotron resonance tandem mass spectrometry, as well as new tandem mass spectrometry methods for proteomics analysis. Dr. Bogdan received his M.Sc. degree in Chemistry from the University of Amsterdam, where he worked in Prof. N. M. M. Nibbering's laboratory, studying the unimolecular dissociation of organic ions, and his Ph.D. from the University of Waterloo in 2001 in Prof. T. B. McMahon's laboratory, investigating the thermochemistry and the structure of negative cluster ions in the gas phase by pulsed-ionization high pressure mass spectrometry and quantum chemistry calculations.

Richard D. Smith currently holds the positions of Chief Scientist and Battelle Fellow at the Pacific Northwest National Laboratory (PNNL). His research is primarily focused on the development of advanced analytical methods and instrumentation, with particular emphasis on high resolution separations and mass spectrometry, and their applications in biological research. His current research involves the development and application of new methods for quantitatively probing the entire array of proteins expressed by a cell, tissue, or organism, i.e., its "proteome." Dr. Smith is the author or co-author of more than 500 publications, has been awarded 21 patents and 7 R&D 100 Awards, and has presented more than 250 invited or plenary lectures at national and international scientific meetings. He received his B.Sc. in Chemistry at the University of Massachusetts at Lowell in 1971 and his Ph.D. in Physical Chemistry at the University of Utah in 1975 in Prof. J. H. Futrell's laboratory. Following a Postdoctoral Fellowship at the Naval Research Laboratory in Washington, DC, he moved to PNNL in 1977.