

Increased Proteome Coverage for Quantitative Peptide Abundance Measurements Based upon High Performance Separations and DREAMS FTICR Mass Spectrometry

Ljiljana Paša-Tolić, Richard Harkewicz, Gordon A. Anderson, Nikola Tolić, Yufeng Shen, Rui Zhao, Brian Thrall, Christophe Masselon, and Richard D. Smith

Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington, USA

A primary challenge in proteome measurements is to be able to detect, identify, and quantify the extremely complex mixtures of proteins. The relative abundances of interest span at least six orders of magnitude for mammalian proteomes, and this constitutes an intractable challenge for high throughput proteome studies. We have recently described a new approach, Dynamic Range Enhancement Applied to Mass Spectrometry (DREAMS), which is based upon the selective ejection of the most abundant species to expand the dynamic range of Fourier transform ion cyclotron resonance (FTICR) measurements. The basis of our approach is on-the-fly data-dependent selective ejection of highly abundant species, followed by prolonged accumulation of remaining low-abundance species in a quadrupole external to the FTICR ion trap. Here we report the initial implementation of this approach with high efficiency capillary reverse phase LC separations and high magnetic field electrospray ionization FTICR mass spectrometry for obtaining enhanced coverage in quantitative measurements for mammalian proteomes. We describe the analysis of a sample derived from a tryptic digest of proteins from mouse B16 cells cultured in both natural isotopic abundance and ^{15}N -labeled media. The FTICR mass spectrometric analysis allows the assignment of peptide pairs (corresponding to the two distinctive versions of each peptide), and thus provides the basis for quantitative measurements when one of the two proteomes in the mixture is perturbed or altered in some fashion. We show that implementation of the DREAMS approach allows assignment of approximately 80% more peptide pairs, thus providing quantitative information for approximately 18,000 peptide pairs in a single analysis. (J Am Soc Mass Spectrom 2002, 13, 954–963) © 2002 American Society for Mass Spectrometry

Probably the greatest challenges presently facing analytical chemistry are related to the development of high throughput methods for the comprehensive study of biological systems. In the era of systems biology, computational and high throughput experimental biological approaches are combined to provide global snapshots of entire genomes and proteomes under tissue- and disease-specific conditions. While methods to simultaneously assess the abundances of thousands of expressed genes at the transcriptome (i.e., mRNA) level now exist [1–3], in many cases it is difficult to predict the resulting proteins since post-transcriptional processes play a major role in determining protein abundances and modification states.

For these and other reasons protein abundances often show poor correlation with mRNA levels [4]. Thus, considerable attention is now focused on the proteome, the complement of proteins expressed by a particular cell, organism or tissue at a given time and under a specific set of conditions [5–7].

The study of proteomes is presently largely based upon two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) methods capable of visualizing hundreds to thousands of protein “spots”, generally with 1 to 100 femtomole detection limits, and augmented by subsequent use of matrix-assisted laser desorption/ionization (MALDI) [8] or electrospray ionization (ESI) [9] with mass spectrometry (MS) for protein identification or characterization. These conventional MS approaches rely on the analysis of enzymatically processed polypeptides from an isolated “spot” for a protein (or a small subset of proteins) followed by either peptide mapping or tandem MS (i.e.,

Published online June 25, 2002

Address reprint requests to Dr. R. D. Smith, Environmental and Molecular Sciences Laboratory, PacificNorthwest National Laboratory, MSIN: K8-98, P.O. Box 999, Richland, WA 99352, USA. E-mail: rd_smith@pnl.gov

MS/MS). However, current level of performance obtained by this methodology falls far short of what is needed to elevate proteomics to throughput levels approaching those of mRNA expression measurements. Many important regulatory proteins are expressed at such low levels (e.g., <1000 copies per cell) that their detection is precluded unless 2D-PAGE is preceded by extensive fractionation of large quantities of protein and the processing of a large numbers of narrow-range gels [10]. Additionally, proteome coverage provided by 2D-PAGE is problematic for proteins having extreme isoelectric points or molecular masses, as well as for membrane proteins (which in total account for roughly half of all proteins). Even with the use of narrow-range immobilized pH gradients [10, 11], 2D-PAGE/MS approach remains lacking in proteome coverage, sensitivity, dynamic range, throughput, and the precision needed to discern small (but often biologically important) changes in protein abundances. Consequently, there are considerable efforts being devoted to the development of new proteomics technologies that attempt to bypass 2D-PAGE, by the use of alternative separation methodologies, and combined with some form of MS, often after protein digestion using specific proteases (e.g., trypsin). For example, Washburn et al. have recently reported on the use of 2-D capillary liquid chromatography (LC)-tandem MS strategy, in which peptides were separated using cation exchange LC followed by reversed-phase (RP) LC coupled online to the mass spectrometer, and demonstrated the ability to identify 1484 yeast proteins [12].

We have recently described a global proteomics strategy that provides improvements in sensitivity, dynamic range, comprehensiveness, and throughput based upon the use of polypeptide “accurate mass tags” (AMTs), i.e., peptides (or modified peptides) with the molecular masses measured with high enough mass measurement accuracy (MMA) such that their masses are effectively unique among all of the possible peptides predicted from an annotated genome [13–16]. The two-stage strategy initially described exploits high resolution capillary LC separations combined with Fourier transform ion cyclotron resonance (FTICR) instrumentation [17], to first validate polypeptide AMTs using accurate mass measurements combined with MS/MS measurements. The AMTs then provide the basis for second stage high throughput studies without the need for routine MS/MS measurements. Key potential attractions of the approach include the feasibility of completely automated high confidence protein identification, more comprehensive proteome coverage, and the capability for exploiting stable-isotope labeling methods for high precision abundance measurements. Another attraction of this approach is the potential for realizing a significant increase in throughput for subsequent quantitative measurements of changes in gene expression (following assault), based upon predetermined AMTs with the use of stable isotope labeling (i.e.,

during cell culture using labeled media or Cys-peptide specific labeling methods) [18–20].

We have also recently described the basis of a new approach for dynamic range expansion in FTICR mass spectrometric measurements, which we have termed Dynamic Range Enhancement Applied to MS (DREAMS), and provided an initial evaluation of its utility [21, 22]. This approach shares some conceptual similarities to an approach we have described earlier based completely upon FTICR [23], but uses data-directed ejection of the most abundant ions prior to ion accumulation in a 2-D ion trap located external to the FTICR trap. Due to the elimination of the major ions in this way, which would otherwise result in the rapid filling of the external 2-D quadrupole ion trap (and consequently also the ICR trap), one can then accumulate ions from the ion source for significantly extended times, and achieve greater sensitivity and extended dynamic range. A key advantage of this approach, in comparison with that described earlier [23], is its speed and compatibility with on-line separations and higher magnetic field FTICR instrumentation.

In this work we have initially evaluated the DREAMS FTICR method in conjunction with high efficiency capillary reverse phase LC separations and high magnetic field electrospray ionization-FTICR mass spectrometry as a basis for obtaining enhanced coverage for mammalian proteomes and more comprehensive quantitative measurements. We build upon our earlier studies and describe the analysis of a complex tryptic digest of mouse proteins from B16 cells cultured in both natural isotopic abundance, i.e., ^{14}N - and ^{15}N -labeled media. In previous work we applied stable-isotope labeling for quantitative proteomics and demonstrated the ability to make precise measurements of differences in the abundances of many proteins using capillary isoelectric focusing coupled with FTICR for intact *E. coli* proteins [20]. We have also previously described the combined use of ^{15}N -metabolic labeling and a commercially available Cys-affinity tag [i.e., (+)-biotinyl-iodoacetamidyl-3,6-dioxaoctanediamine; iodoacetyl-PEO-biotin] to isolate and quantitate Cys-containing polypeptides as an alternative to the ICAT (i.e., isotope-coded affinity tag) [18] quantitative proteomics strategy [24]. In addition to being an effective method to determine relative Cys-polypeptide abundances, this approach provides added constraints for protein identification (e.g., the mass difference for a pair of the Cys-peptides corresponds to the number of nitrogen atoms in that particular peptide) and avoids introducing significant differences in the chromatographic elution times for the two versions of labeled peptides, and thus improves quantitation [25]. Herein, we demonstrate the DREAMS approach in the context of analyses of *all* tryptic peptides using ^{14}N - and ^{15}N -labeling of mouse B16 melanoma cells, and show that the number of detected peptide pairs suitable for use in quantitative analyses is significantly increased due to the enhanced dynamic range of the measurements.

Experimental

Sample Preparation

Murine B16 melanoma cells (ATCC #30-2002) were routinely maintained in Dulbecco's Modified Eagles Media supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 100 u/ml penicillin and streptomycin in a humidified atmosphere of 5% CO₂ in air at 37 °C. The cells were plated on 6-well culture plates at a concentration of 10⁴ cells/well and following attachment the medium was changed to Celtone M Natural or Celtone M-N (>98% ¹⁵N) (Martek Biosciences Corporation catalog no. 53998), both of which were supplemented with 10% heat-inactivated fetal bovine serum and 100 U/ml penicillin and streptomycin. The cells were fed with fresh medium daily and allowed to grow to an estimated 85% confluence (4 days). Daily analysis of the number of viable cells by trypan blue exclusion and hemacytometer counts demonstrated that there were no significant differences in proliferation rates or viability between the ¹⁵N-labeled and natural mediums (data not shown). For protein extraction, the cells were washed twice with phosphate-buffered saline and trypsinized, and equal numbers of cells lysed by resuspending in buffer containing 50 mM Tris-HCl (pH 7.4), 100 mM sodium chloride, 50 mM sodium fluoride, and 1% Triton X-100. To evaluate the efficiency of the labeling and achievable precision of quantitation from the normal and ¹⁵N-labeled cell cultures, identical aliquots of the soluble protein extracts obtained from the B16 cells were mixed. The sample was desalted into 100 mM NH₄HCO₃, 5 mM EDTA, pH 8.4, and digested with trypsin (1:50 enzyme to protein ratio) overnight at 37 °C.

High Efficiency Capillary Reverse Phase LC

Packed capillary LC separations were carried out at constant pressure as previously described [26]. The polypeptide solution was injected into a capillary column (150 μm i.d., 360 μm o.d., 50 cm long or 120 cm long), packed in-house [27] with 5 μm diameter Jupiter C₁₈ resin (300 Å pore size, Jupiter, Phenomenex, Terrence, CA). A solvent gradient was used to elute the polypeptides with solvent A, 0.2% acetic acid, 0.1% TFA, water (vol %) and solvent B, 80% acetonitrile, 20% water, 0.1% TFA (vol %), mixed in a steel mixer (~2.8 mL) with a magnetic stirrer before entering the separation capillary. When operated at pressure <8000 psi (i.e., 50 cm column), the mobile phase switching and sample introduction were performed using a 4-port switching valve and 6-port sample injector (Valco, Houston, TX), with a sample loop of 10 μL. When operated at pressure >8000 psi (i.e., 120 cm column), passive feedback switching valves (Valco) were used for the mobile phase switching and sample introduction with a sample loop of 50 μL. The peptides were eluted using a linear gradient of 0 to 80% solvent B over ~120

min (50 cm column) or ~15 h (120 cm column). Solvents were delivered to the capillary column (at a pressure of 5000 psi or 8500 psi for the 50 and 120 cm long columns, respectively) using two Isco model 100 DM pumps (ISCO, Lincoln, NE) controlled by a series D controller resulting in a capillary flow rate of ~1 μL/min.

FTICR Mass Spectrometry

For FTICR analyses, the capillary LC was coupled on-line via ESI with an 11.4 tesla FTICR mass spectrometer designed and constructed in our laboratory [22]. The output of the capillary LC was coupled to an atmospheric ESI source comprising a stainless steel heated capillary inlet to the mass spectrometer that incorporates an electrodynamic ion funnel assembly to improve ion transmission through the interface [28–30] followed by a collisional focusing quadrupole. Following ion accumulation in an external 2-D quadrupole, the trapped ions were ejected into a quadrupole ion guide and then dynamically trapped in a cylindrical open-ended capacitively coupled FTICR cell, operating at a pressure of ~10⁻⁹ torr. An Odyssey data station (Finnigan Corp., San Jose, CA) was used to generate experimental control scripts and to acquire data from the FTICR trap detector preamplifier.

Capillary LC DREAMS FTICR

To implement selection based on RF-only resonant dipolar excitation, two opposite rods of the selection quadrupole (located just before an accumulation quadrupole) were coupled to the secondary coil of an auxiliary 1:1 transformer. The center point of the transformer secondary coil was driven by the main RF-drive at an amplitude of ~300 V_{p-p} and frequency of ~980 kHz. The resonant dipolar excitation voltages are applied to the primary coil of the transformer. These auxiliary or "tickle" frequency voltages were synthesized using our ICR-2LS software program and generated using a plug-in PC DAC board (National Instruments, Austin, TX). An ancillary PC running the ICR-2LS was used to data-dependently generate the tickle frequencies necessary for ion ejection from the selection quadrupole, as described previously [21, 22].

The capillary LC DREAMS FTICR experiments consisted of two spectrum acquisition sequences conducted repetitively during the separation: a standard acquisition followed by a second acquisition during which the most abundant species were ejected in the 2-D quadrupole during (a prolonged) accumulation event. For the first "normal" spectral acquisition we used an extended ion accumulation period and ion trapping conditions that provide largely unbiased accumulation of ions under conditions of very large ion currents that applied during elution of significant peaks [31, 32]. This approach allowed the use of extended ion accumulation periods that would otherwise have resulted in significant "overflowing" of the FTICR trap or other undesired

artifacts. (While we cannot dismiss the possibility that the accumulation conditions resulted in some bias across the m/z range, it is unlikely that any such contribution would affect the measurement of the relative abundances of the closely spaced stable isotope labeled peptide pairs from which the quantitative information is obtained here.)

For the DREAMS spectra a 12-bit ADC (National Instruments, Austin, TX), triggered by the Odyssey data station, was used to acquire time-domain signals which were converted to m/z spectra by the ICR-2LS software program, on the basis of a series of user predefined processing steps. The DREAMS tickle frequency calibration was determined as described previously [21, 22]. The waveform was then downloaded to a National Instruments DAQ 5411 arbitrary waveform generator (AWG). A superposition of resonant (dipolar) sine waveforms was applied to the selection quadrupole for ejection of up to 10 most abundant peaks simultaneously before accumulation (actually small m/z regions were ejected due to an effective resolution of 30 to 50, which also served to eject related isotopic peaks for the species selected as well as other closely spaced peaks). This allowed the accumulation quadrupole to be filled for an extended period with the remaining lower abundance species. Since the fraction of the total ion current ejected varied from spectrum to spectrum, this approach also results in an increase in dynamic range for the DREAMS spectra that varies from spectrum to spectrum.

Data Analysis

Data processing used the ICR-2LS and LaV2DG software packages developed at our laboratory. Briefly, ICR-2LS converted the raw data into m/z spectra that were subsequently transformed to generate a list of neutral masses using an implementation of the THRASH algorithm originally developed by Horn et al. [33]. The output data file was visualized as a 2-D display of polypeptide neutral mass versus spectrum number (see, for example, Figure 3). The software developed for this purpose also allowed assignment of the peptide pairs (e.g., $^{14}\text{N}/^{15}\text{N}$ peptide pairs) based on user-defined criteria. Since we did not actually characterize each detected peptide from the tryptic digest of the mouse B16 cell proteins for this initial evaluation of the DREAMS method, we used as a practical figure of merit the number of distinguishable peptide pairs that were detected as a measure of proteome coverage. To determine the number of $^{14}\text{N}/^{15}\text{N}$ peptide pairs detected in this work, we first assembled all measured neutral masses into “unique mass classes” (UMCs), defined as masses within a large mass tolerance (e.g., 25 ppm) and detected in consecutive spectra so as not to overestimate the number of species detected. This set of UMC putative peptide masses was then searched for pairs of peptides whose mass differences corresponded to unit mass multiples of the $^{14}\text{N}/^{15}\text{N}$ mass difference

within a conservative mass error tolerance (e.g., 0.02 μm), and for which there was a substantial overlap of elution times (since the ^{14}N - and ^{15}N -labeled peptides largely coelute [24]). The abundance ratios (ARs; the ratio of the average of the ^{14}N signal intensities across the peak to the average ^{15}N signal intensities of each peptide) were then calculated for all $^{14}\text{N}/^{15}\text{N}$ peptide pairs, the number of peptide pairs determined for both the normal spectra and for the DREAMS spectra. The set of pairs common to both spectra sets also determined (i.e., peptide pairs corresponding to lower abundances species not ejected from the DREAMS spectra). To obtain the small number of highly confident unambiguous $^{14}\text{N}/^{15}\text{N}$ pairs, we then excluded all peptide pairs on the tails of the AR distribution and discarded any possible peptide pairs that had a peptide assigned to more than one peptide pair.

Results and Discussion

The large variation of protein relative abundances having potential biological significance in mammalian systems (>6 orders of magnitude) presents a major challenge for proteomics. While FTICR mass spectrometry has demonstrated a capability for ultrasensitive characterization of biopolymers (e.g., achieving subattomole detection limits) [34, 35], the maximum dynamic range for a single mass spectrum (i.e., without the use of spectrum averaging or summation) is typically constrained to about 10^3 . An important factor conventionally limiting achievable FTICR sensitivity and dynamic range is the maximum charge capacity of either the external ion accumulation device or the FTICR mass analyzer itself. Prolonged ion accumulation would be helpful during the LC elution of low-abundance components (i.e., during the “valleys” in chromatograms), potentially allowing measurable signals to be obtained for otherwise undetectable species, and increasing the effective overall dynamic range of proteome measurements. Unfortunately, “overfilling” of external multipole ion traps by high-abundance species often results in a biased accumulation process in which parts of the m/z range are selectively retained or lost [31] and/or extensive ion activation and dissociation occurs well before sufficient populations of low level species can be accumulated [36, 37]. We have recently reported on approaches for minimizing or eliminating such overfilling of the 2-D quadrupole accumulation trap [32], and make use of such an approach in this work. However, even when artifacts associated with trap overfilling are avoided, and allowing use of greatly prolonged accumulation times, the ion capacity of an FTICR trap beyond which unacceptable space charge effects occur (e.g., peak coalescence) will still limit the dynamic range achievable in a single spectrum to $\sim 10^3$ for the case of the analysis of complex peptide mixtures [26]. Thus, the species present at low relative abundances will be effectively “masked” by the presence of coeluting highly abundant species. Under such conditions the

overall dynamic range achievable will depend on factors that include: (1) the quality of the preceding sample processing and LC separations steps and (2) the extent to which the maximum capacity of the FTICR trap can be effectively utilized in generating each spectrum. As we have shown previously, when combined with high-resolution LC separations, the overall effective dynamic range of FTICR based proteome measurements is approximately 10^4 to 10^5 [26]. While the overall dynamic range can certainly be increased further by the use of additional “up-front” separation dimensions, in principle to any desired extent, doing so necessarily results in selective sample losses, increased sample size requirements, and greatly decreased overall throughput since many more samples will have to be analyzed to achieve broad proteome coverage.

The DREAMS FTICR approach [21, 22] allows the use of longer accumulation times for species present at low abundance by the elimination of the most highly abundant species during the ion accumulation process. This provides an overall improvement in sensitivity for any lower abundance species not separated from high abundance species (e.g., by chromatography), and thus an extension of the overall dynamic range of proteome measurements. In earlier work we have demonstrated methods for FTICR dynamic range expansion based upon ion manipulations in the FTICR cell based upon the selective ejection of the most abundant species by excitation of their magnetron motion in the FTICR cell [23]. Unfortunately, this approach (1) requires transient introduction of a relatively high gas pressure in the FTICR cell that must be removed before analysis, and (2) takes longer to apply as magnetic field increases due to the longer times for ion ejection by magnetron excitation [38], and is thus not readily compatible with on-line separations. In this work the ion ejection step is preformed in an external 2-D quadrupole, such that FTICR performance is unaffected. Our previous DREAMS reports [21, 22] illustrated basic principles of the technique using yeast global tryptic digest and similar data analyses as described here (for the mouse B16 cells tryptic digest) to estimate the increase in the number of detected species. Herein we describe the improved peptide abundance measurements using $^{14}\text{N}/^{15}\text{N}$ peptide pairs for quantitative measurements, and illustrate significant increase in the number of peptides for which precise ARs can be determined.

In this initial evaluation of the DREAMS approach for quantitative proteomics, we have used both high performance capillary LC separations and a high magnetic field FTICR mass spectrometer, and evaluated the improvement in mammalian quantitative proteome measurements achievable with this approach. For the latter purpose we prepared a global tryptic digest of ^{14}N - and ^{15}N -labeled peptides from soluble proteins extracted from mouse B16 cells and mixed in approximately 1:1 ratio. In the analysis every second mass spectrum resulted from the accumulation of ions that remained after the dipolar ejection of the most abun-

dant ions (i.e., the DREAMS spectrum). Thus, it was possible to reconstruct two different ion chromatograms for this experiment, corresponding to the “normal” (i.e., broadband mode acquisition with a 100 ms ion accumulation time) and the DREAMS (i.e., data dependent RF-only selective acquisition with a 300 ms to 1 s accumulation time) spectra sets.

Figure 1 illustrates dynamic range enhancement capability using DREAMS by comparing the two mass spectra for one point in a capillary LC separation obtained using a 50 cm long column and 120 min reverse phase gradient. The data allow two chromatograms to be reconstructed from the data corresponding to the normal and DREAMS spectra (Figure 1, left top and bottom). As shown by the comparison of the full mass spectra, the normal spectrum (Figure 1, middle top) is dominated by a number of major peptide ions, most prominently three pairs of ^{14}N - ^{15}N -labeled peptides in the $500 < m/z < 700$ range. The information from this spectrum was used “on-the-fly” to apply dipolar RF excitation in the 2-D quadrupole at the secular frequencies corresponding to the m/z of these major ions during ion accumulation for the next spectrum. As a result the species in this m/z region were effectively ejected prior to accumulation and the FTICR spectrum is now dominated by a much different set of species (Figure 1, middle bottom). As noted in the Experimental section, the use of nominally “unbiased” ion accumulation conditions in the 2-D quadrupole results in a much greater gain in sensitivity for low level components in the DREAMS spectrum than would be expected from the difference in ion accumulation times alone. As a single example selected from many similar cases, the inset right shows that the signal-to-noise ratio (S/N) for a peptide pair at $m/z \sim 1000$ is greatly improved from a level at which no effective identification could be obtained, to a level at which a very precise relative abundance ratio (AR) can be determined for the peptide pair, with the gain in the S/N in this case of ~ 50 .

In this initial implementation, the overall speed of the DREAMS process is limited by both software for data processing, data transfer times, the generation of the DREAMS rf waveforms, and the ion cooling times used to improve FTICR spectrum quality. Thus, the total time required for acquisition of both the normal and DREAMS spectra pair is ~ 20 s. In order to provide a basis for a more quantitative evaluation of the additional information acquired with the DREAMS analysis we used longer high performance reverse phase gradient separations where the minimum peak width is about one minute. Figure 2 shows chromatograms reconstructed from the FTICR spectra acquired during a single high pressure capillary LC separation of 1:1 mixture of ^{14}N - and ^{15}N -labeled mouse B16 tryptic peptides using 120 cm long column and a nearly 15 h gradient, with the data being acquired for ~ 10 h (started before the first peaks were expected to elute). The average LC peak widths in this case spanned from

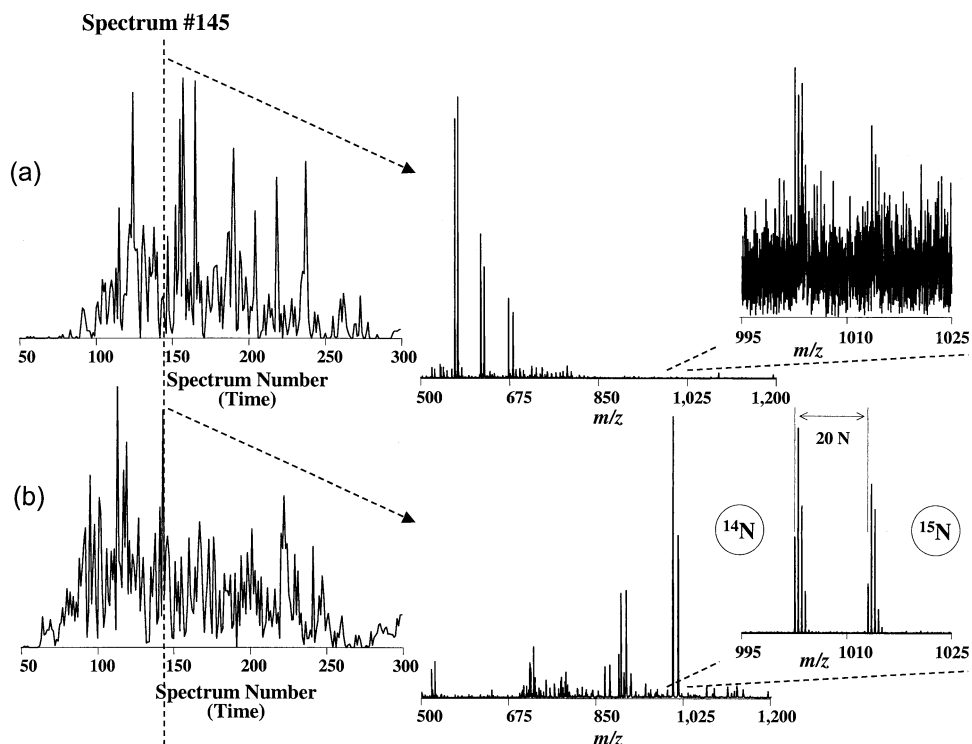


Figure 1. Capillary LC-FTICR results showing total ion chromatogram (TIC) and portions of the mass spectra acquired during the normal spectrum acquisition process and the alternating DREAMS spectrum acquisition process obtained using a 50 cm long column and 120 min reverse phase gradient. (a) TIC reconstructed from the FTICR spectra 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 and representative spectrum obtained using broadband mode acquisition and a 100 ms accumulation time. (b) corresponding TIC and representative spectrum obtained using RF-only selective acquisition and 300 ms accumulation time. Resonant frequencies for RF-only dipolar excitation were identified during a broadband ion acquisition (a) and up to five species having relative abundance $>10\%$ were then data-dependently ejected during a selective acquisition that immediately followed (b). Total amount of sample injected ~ 5 micrograms.

three to more than ten spectra for both the normal and DREAMS chromatograms. It is evident from the quite different chromatographic profiles that many additional species were detected with much greater signal intensities in the DREAMS spectra, particularly at longer retention times. However, the chromatograms only provide a qualitative view of the added information from the DREAMS approach, and one that is obviously limited to the more abundant species.

To further evaluate the utility of the DREAMS approach for providing additional information on proteomes we use as a figure of merit the number of peptide pairs that could be confidently assigned to both the normal and DREAMS spectra sets for the separation shown in Figure 2. Briefly, we first assembled a list of all the neutral masses that could be distinguished during the separations. Here each distinguishable isotopic cluster gives one putative peptide mass that is used. Any given mass that appeared at two distinctive elution times was counted only once, thus leading to a small potential underestimation of the number of polypeptides actually detected. This list was then searched for pairs whose mass difference corresponds to a multiple

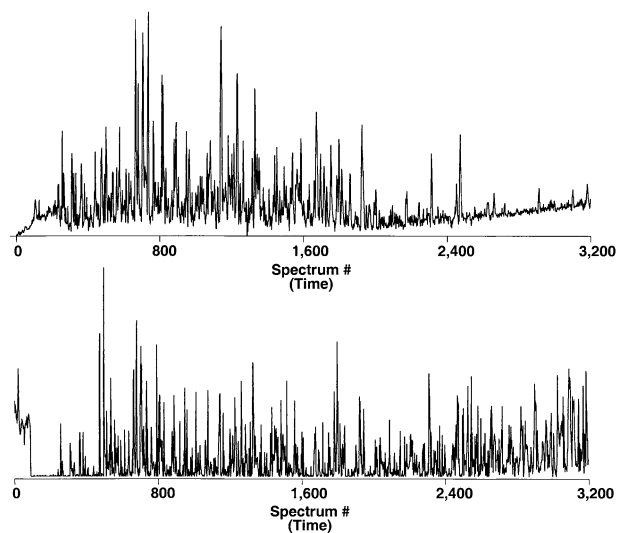


Figure 2. Total ion chromatograms (TICs) reconstructed from FTICR mass spectra acquired using broadband mode acquisition with 100 ms accumulation time (top) and with RF-only selective (i.e., DREAMS) acquisition and 1 s accumulation time (bottom) during a capillary LC separation of soluble proteins that were digested with trypsin from a 1:1 mixture of ^{14}N - and ^{15}N -labeled mouse B16 cells.

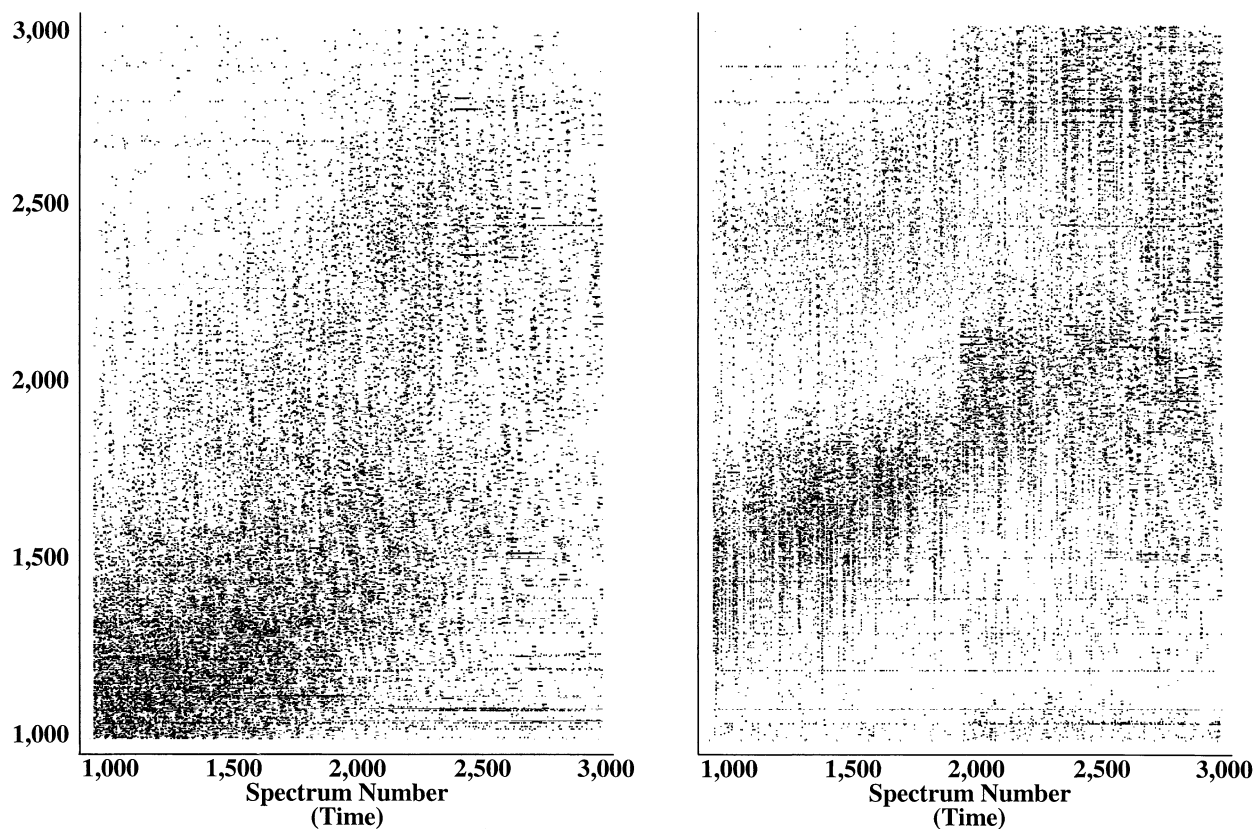


Figure 3. 2-D display showing the molecular masses and spectrum number for peptide pairs detected from a 1:1 mixture of ^{14}N - and ^{15}N -labeled peptides from mouse B16 cells (see Figure 2). The left panel shows peptide pairs detected in the normal spectrum set and the right panel shows peptide pairs 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.

of the $^{14}\text{N}/^{15}\text{N}$ mass difference and for which there was also an overlap in the elution time. From this smaller set we also discarded any possible peptide pairs that had a peptide assigned to more than one peptide pair (i.e., an ambiguous pair assignment). Finally, we also discarded a set of possible peptide pairs with relative abundances outside of the range of 0.4 to 1.6 (since an approximately 1:1 ratio was expected). These criteria resulted in a conservative determination of the number of peptide pairs that were detected in the normal and DREAMS spectrum sets. This number represented a lower bound for the number of cases where quantitative comparisons of the relative intensities of such pairs could be derived, and thus the cases where stable-isotope labeling methodology could potentially be applied to obtain quantitative measurements for relative protein abundances.

Figure 3 shows a 2-D display that includes spots for *only* the peptide pairs measured from the capillary LC-FTICR analysis of Figure 2. This figure is a result of substantial post-processing of the capillary LC-FTICR analysis to convert the mass spectral information into molecular masses and to identify the subset of $^{14}\text{N}/^{15}\text{N}$ -labeled peptide pairs as described above. The analysis of the normal spectra resulted in detection of a total of

9896 $^{14}\text{N}/^{15}\text{N}$ -labeled peptide pairs (Figure 3, left). The average abundance ratio (AR) for the peptide pairs was 1.05, only a slight deviation for the nominal value of 1.0 expected. The standard deviation for the peptide pairs was 0.28. The second set of DREAMS mass spectra revealed 8856 $^{14}\text{N}/^{15}\text{N}$ -labeled peptide pairs (Figure 3, right), of which 7917 were “new” peptide pairs not detected in the normal spectra, and 939 that were lower level peptide pairs also observed in the normal spectra. The average AR for the peptides detected in the DREAMS portion of the analysis was 1.015, and the standard deviation was 0.31. It should be noted that the assigned peptide pairs accounted for less than 25% of the total number of species detected in the analysis that presumably contains contaminants, other mixture components and pairs of peptides excluded by the criteria set above. Finally, the combined total of 17,813 unambiguous and unique peptide pairs gave an average AR of 1.035 and a standard deviation of 0.29.

While different approaches for the analysis of this data, primarily the use of different peptide pair selection criteria, will lead to slightly different results, the key conclusion is that the number of detected peptides is greatly increased by the use of the DREAMS meth-

odology. Indeed, the number of peptides for which quantitative information could be obtained from the relative abundances of peptide pairs was increased by 80% using the criteria described above. However, it is not possible from the present analysis to readily determine the increase in the number of proteins detected due to the fact that available genomic data for mouse, and the reliability of the suite of predicted proteins, is still problematic and significantly incomplete.

In order to initially address the question if the increase in the number of observed peptides translates into markedly better proteome coverage, we have analyzed a mixture of identical aliquots of a natural isotopic abundance and ^{15}N -labeled version of *D. radiodurans* cells. DREAMS LC-FTICR analysis of this mixture resulted in detection of 9559 peptide pairs in the “normal” set of spectra and 7926 peptide pairs in the DREAMS set of spectra. We then used *D. radiodurans* AMT database, assembled in our initial demonstration of an AMT approach for high throughput quantitative proteomics [15, 16], to uniquely identify 2244 peptides, corresponding to 965 ORFs, in the normal set of spectra and 2259 peptides, corresponding to 1000 ORFs, in the DREAMS set of spectra, for a total of 3264 peptides and 1244 ORFs. 1022 (45%) unique peptides, corresponding to 279 (30%) ORFs, identified in the DREAMS dataset were “new” (i.e., not identified in the normal dataset). Thus, both the number of identified peptides and the number of proteins for which quantitative information could be obtained from the relative abundances of peptide pairs was increased by >30% in this case. The overall proteome coverage can be compared with recent work by Washburn et al. [39] using their multidimensional protein identification procedure (MudPIT) [12], which reported an average of 869 yeast proteins (13.8% of the predicted larger set of yeast proteins) observed in analyses of 1:1, 5:1, and 10:1 mixtures of *S. cerevisiae* grown in ^{14}N - and ^{15}N -enriched minimal media. In this work, we observed 1244 *D. radiodurans* proteins (39.9% of the predicted *D. radiodurans* proteome) as peptide pairs and with good precision ARs in a single DREAMS LC-FTICR analysis using our AMT-based approach. A detailed statistical analysis of such peptide abundance measurements achievable using AMT-based proteomics is in progress and will be the subject of a future report.

Additional efforts are planned towards improvements of the resolution of the ion ejection step (currently 30 to 50) and to also implement the DREAMS analysis approach as a routine part of our accurate mass tag [13–16] strategy for quantitative proteome measurements. In particular, the effective integration of this step with the use of a variable ion accumulation time for automated gain control is in progress. This step will be important for enabling the use of absolute peak intensities, and also eliminate small errors associated with AR measurements due to possible variations in sensitivity during peak elution and the small elution time offsets associated with the different stable isotope-

labeled version of each peptide. Since this capability would more effectively exploit the optimum trap capacity for each spectrum, we anticipate that this combination of capabilities would further increase the dynamic range in proteome measurements. Finally, the speed of the methodology needs to be increased. At present we utilize a 5 s ion cooling step to enhance the quality of the resulting data. However, when this is done in conjunction with DREAMS analysis (or multiplexed-MS/MS) [40, 41], the result is either the need for an extended analysis time or effectively decreased chromatographic resolution (and significantly decreased proteome coverage). We have recently shown the potential to decrease the ion cooling time by up to an order of magnitude using an adiabatic cooling scheme that involves the ramping of the FTICR ion trapping well voltages [42]. The combination of this step with decreased data transfer and processing times promises to result in substantially reduced overall analysis times.

Conclusions

Global proteomic measurements for mammalian systems pose an enormous challenge because of factors that include both the complexity and dynamic range of the constituents. In this work we have shown that the DREAMS approach applied in conjunction with high resolution capillary LC separations and high performance FTICR results in a substantial increase in the number of peptides detected from global protein enzymatic digests. The observed 80% increase in the number of stable-isotope labeled peptide pairs is attributed to the detection of lower level peptides that would otherwise be missed in our present global proteome analysis approach. The additional peptides detected from the DREAMS analysis represent high quality measurements from which useful quantitative information can be extracted, and the present results provide a strong validation of the stable-isotope labeling approach for quantitation. It should be noted that the DREAMS FTICR approach is broadly applicable to other similar types of analytical problems (e.g., complex fuel or environmental mixtures). Additionally, the DREAMS approach would also offer significant gains when coupled with conventional ion trap instrumentation in place of FTICR analyzer used in this work.

As noted above, additional efforts are needed to improve the FTICR spectrum acquisition speed, increase the resolution of the DREAMS ion ejection process, and to implement active control of the ion accumulation times based upon signal intensity information. These efforts hold the promise of considerable further gains, and to even allow *multiple* DREAMS ion ejections steps of progressively lower level species for measurements of even lower abundance proteins, as we have demonstrated in earlier work based upon ejection from the FTICR trap [23]. The implementation of the DREAMS approach in conjunction with multiplexed-MS/MS [38, 39] analyses will

serve to increase the effectiveness and throughput for the de novo characterization of low level peptides and to provide initial identifications in support of the use of AMTs for subsequent high throughput proteome measurements.

Finally, we believe that the DREAMS approach will become an important part of future FTICR and ion trap instrumentation, and serve to substantially increase the sensitivity and dynamic range for the studies of complex biological systems.

Acknowledgments

The authors are grateful to Drs. Mikhail Belov and to Mikhail Gorshkov for helpful discussions, and Dr. Thomas C. Conrads for processing of the mouse B16 cell proteome sample used in this work. Portions of this research were supported by the office of Biological and Environmental Research, U.S. Department of Energy, and the National Cancer Institute under grants CA81654 and CA86340. Pacific Northwest National Laboratory is operated by Battelle Memorial Institute for the U.S. Department of Energy under contract DE-AC06-76RLO 1830.

References

- Velculescu, V. E.; Zhang, L.; Zhou, W.; Vogelstein, J.; Basrai, M. A.; Bassett, D. E. J.; Hieter, P.; Vogelstein, B.; Kinzler, K. W. Characterization of the Yeast Transcriptome. *Cell* **1997**, *88*, 243–251.
- Zhang, L.; Zhou, W.; Vogelstein, B.; Velculescu, V. E.; Kern, S. E.; Hruban, R. H.; Hamilton, S. R.; Kinzler, K. W. Gene Expression Profiles in Normal and Cancer Cells. *Science* **1997**, *276*, 1268–1272.
- Schena, M.; Shalon, D.; Davis, R. W.; Brown, P. O. Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray. *Science* **1995**, *270*, 467–70.
- Gygi, S. P.; Rochon, Y.; Franza, B. R.; Aebersold, R. Correlation Between Protein and mRNA Abundance in Yeast. *Mol. Cell Biol.* **1999**, *19*, 1720–1730.
- Wilkins, M. R.; Sanchez, J.-C.; Gooley, A. A.; Appel, R. D.; Humphery-Smith, I.; Hochstrasser, D. F.; Williams, K. L. Progress with Proteome Projects: Why All of Proteins Expressed by a Genome Should Be Identified and How to Do it. *Biotechnol. Gene Eng. Rev.* **1995**, *13*, 19–50.
- Wilkins, M. R.; Pasquali, C.; Appel, R. D.; Ou, K.; Golaz, O.; Sanchez, J. C.; Yan, J. X.; Gooley, A. A.; Hughes, G.; Humphery-Smith, I.; Williams, K. L.; Hochstrasser, D. F. From Proteins to Proteomes: Large Scale Protein Identification by Two-Dimensional Electrophoresis and Amino Acid Analysis. *Bio/Technol.* **1996**, *14*, 61–65.
- Wilkins, M. R.; Williams, K. L.; Appel, R. D.; Hochstrasser, D. F., Eds. *Proteome Research: New Frontiers in Functional Genomics*. Springer-Verlag: Berlin, Heidelberg, 1997.
- Hillenkamp, F.; Karas, M.; Beavis, R. C.; Chait, B. T. Matrix-Assisted Laser Desorption Ionization Mass-Spectrometry of Biopolymers. *Anal. Chem.* **1991**, *63*, 1193–1202.
- Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science* **1989**, *246*, 64–71.
- Gygi, S. P.; Corthals, G. L.; Zhang, Y.; Rochon, Y.; Aebersold, R. Evaluation of Two-Dimensional Gel Electrophoresis-Based Proteome Analysis Technology. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9390–9395.
- Westbrook, J. A.; Wait, R.; Welson, S. Y.; Dunn, M. J. Zooming-in on the Proteome: Very Narrow-Range Immobilized pH Gradients Reveal More Protein Species and Isoforms. *Electrophoresis* **2001**, *22*, 2865–2871.
- Washburn, M. P.; Wolters, D.; Yates, J. R., III. Large-Scale Analysis of the Yeast Proteome by Multidimensional Protein Identification Technology. *Nat. Biotechnol.* **2001**, *19*, 242–247.
- Conrads, T. P.; Anderson, G. A.; Veenstra, T. D.; Paša-Tolić, L.; Smith, R. D. Utility of Accurate Mass Tags for Proteome-Wide Protein Identification. *Anal. Chem.* **2000**, *72*, 3349–3354.
- Smith, R. D.; Paša-Tolić, L.; Lipton, M. S.; Jensen, P. K.; Anderson, G. A.; Shen, Y.; Conrads, T. P.; Udseth, H. R.; Harkewicz, R.; Belov, M. E.; Masselon, C.; Veenstra, T. D. Rapid Quantitative Measurements of Proteomes by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Electrophoresis* **2001**, *22*, 1652–1668.
- Smith, R. D.; Anderson, G. A.; Lipton, M. S.; Masselon, C.; Paša-Tolić, L.; Shen, Y.; Udseth, H. R. The Use of Accurate Mass Tags for High-Throughput Microbial Proteomics. *OMICS* **2002**, *6*, 61–90.
- Smith, R. D.; Anderson, G. A.; Lipton, M. S.; Paša-Tolić, L.; Shen, Y.; Conrads, T. P.; Veenstra, T. D.; Udseth, H. R. An Accurate Mass Tag Strategy for Quantitative and High Throughput Proteome Measurements. *Proteomics* **2002**, in press.
- Marshall, A. G. Milestones in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry Technique Development. *Int. J. Mass Spectrom. Ion Processes* **2000**, *200*, 331–356.
- Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. Quantitative Analysis of Complex Protein Mixtures Using Isotope-Coded Affinity Tags. *Nat. Biotechnol.* **1999**, *17*, 994–999.
- Oda, Y.; Huang, K.; Cross, F. R.; Cowburn, D.; Chait, B. T. Accurate Quantitation of Protein Expression and Site-Specific Phosphorylation. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6591–6596.
- Paša-Tolić, L.; Jensen, P. K.; Anderson, G. A.; Lipton, M. S.; Peden, K. K.; Martinović, S.; Tolić, N.; Bruce, J. E.; Smith, R. D. High Throughput Proteome-Wide Precision Measurements of Protein Expression using Mass Spectrometry. *J. Am. Chem. Soc.* **1999**, *121*, 7949–7950.
- Belov, M. E.; Anderson, G. A.; Angell, N. H.; Shen, Y.; Tolić, N.; Udseth, H. R.; Smith, R. D. 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.* **2001**, *73*, 5052–5060.
- Harkewicz, R.; Belov, M. E.; Anderson, G. A.; Paša-Tolić, L.; Masselon, C. D.; Prior, D. C.; Udseth, H. R.; Smith, R. D. ESI-FTICR Mass Spectrometry Employing Data-Dependent External Ion Selection and Accumulation. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 144–54.
- Bruce, J. E.; Anderson, G. A.; Smith, R. D. “Colored” Noise Waveforms and Quadrupole Excitation for the Dynamic Range Expansion of Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.* **1996**, *68*, 534–541.
- Conrads, T. P.; Alving, K.; Veenstra, T. D.; Belov, M. E.; Anderson, G. A.; Anderson, D. J.; Lipton, M. S.; Paša-Tolić, L.; Udseth, H. R.; Chrisler, W. B.; Thrall, B. D.; Smith, R. D. Quantitative Analysis of Bacterial and Mammalian Proteomes Using a Combination of Cysteine Affinity Tags and ¹⁵N-Metabolic Labeling. *Anal. Chem.* **2001**, *73*, 2132–2139.
- Zhang, R.; Sioma, C. S.; Wang, S.; Regnier, F. E. Fractionation of Isotopically Labeled Peptides in Quantitative Proteomics. *Anal. Chem.* **2001**, *73*, 5142–5149.
- Shen, Y.; Tolić, N.; Zhao, R.; Paša-Tolić, L.; Li, L.; Berger, S. J.; Harkewicz, R.; Anderson, G. A.; Belov, M. E.; Smith, R. D. High-Throughput Proteomics Using High Efficiency Multiple-Capillary Liquid Chromatography with On-Line High Perfor-

- mance ESI FTICR Mass Spectrometry. *Anal. Chem.* **2001**, *73*, 3011–3021.
27. Shen, Y.; Zhao, R.; Belov, M. E.; Conrads, T. P.; Anderson, G. A.; Tang, K.; Paša-Tolić, L.; Veenstra, T. D.; Lipton, M. S.; Udseth, H. R.; Smith, R. D. Packed Capillary Reverse Phase Liquid Chromatography/Electrospray Ionization-Fourier Transform Ion Cyclotron Resonance Mass Spectrometry for Proteomics. *Anal. Chem.* **2001**, *73*, 1766–1775.
 28. Shaffer, S. A.; Prior, D. C.; Anderson, G. A.; Udseth, H. R.; Smith, R. D. An Ion Funnel Interface for Improved Ion Focusing and Sensitivity Using Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **1998**, *70*, 4111–4119.
 29. Kim, T.; Tolmachev, V.; Harkewicz, R.; Prior, D. C.; Anderson, G. A.; Udseth, H. R.; Smith, R. D.; Bailey, T. H.; Rakov, S.; Futrell, J. H. Design and Implementation of a New Electrodynamic Ion Funnel. *Anal. Chem.* **2000**, *72*, 2247–2255.
 30. Belov, M. E.; Gorshkov, M. V.; Udseth, H. R.; Anderson, G. A.; Smith, R. D. Initial Implementation of an Electrodynamic Ion Funnel with Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.* **2000**, *72*, 2271–2279.
 31. Belov, M. E.; Nikolaev, E. N.; Harkewicz, R.; Masselon, C.; Alving, K.; Smith, R. D. Ion Discrimination During Ion Accumulation in a Quadrupole Interface External to a Fourier Transform Ion Cyclotron Resonance Mass Spectrometer. *Int. J. Mass Spectrom. Ion Processes* **2001**, *208*, 205–225.
 32. Belov, M. E.; Gorshkov, M. V.; Alving, K.; Smith, R. D. Optimal Pressure Conditions for Unbiased External Ion Accumulation in a Two-Dimensional Radio-Frequency Quadrupole for Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1988–1996.
 33. Horn, D. M.; Zubarev, R. A.; McLafferty, F. W. Automated Reduction and Interpretation of High Resolution Electrospray Mass Spectra of Large Molecules. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 320–332.
 34. Valaskovic, G. A.; Kelleher, N. L.; McLafferty, F. W. Attomole Protein Characterization by Capillary Electrophoresis Mass Spectrometry. *Science* **1996**, *273*, 1199–1202.
 35. Belov, M. E.; Gorshkov, M. V.; Udseth, H. R.; Anderson, G. A.; Smith, R. D. Zeptomole-Sensitivity Electrospray Ionization-Fourier Transform Ion Cyclotron Resonance. *Anal. Chem.* **2000**, *72*, 2271–2279.
 36. Sannes-Lowery, K.; Griffey, R. H.; Kruppa, G. H.; Speir, J. P.; Hofstadler, S. A. Multipole Storage Assisted Dissociation, a Novel In-Source Dissociation Technique for Electrospray Ionization Generated Ions. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1957–1961.
 37. Tolmachev, A. V.; Udseth, H. R.; Smith, R. D. Radial Stratification of Ions as a Function of Mass to Charge Ratio in Collisional Cooling Radio Frequency Multipoles Used as Ion Guides or Ion Traps. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1907–1913.
 38. Guan, S. H.; Kim, H. S.; Marshall, A. G.; Wahl, M. C.; Wood, T. D.; Xiang, X. Z. Shrink-Wrapping an Ion Cloud for High-Performance Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Chem. Rev.* **1994**, *94*, 2161–2182.
 39. Washburn, M. P.; Ulaszek, R.; Deciu, C.; Schieltz, D. M.; Yates, J. R. III. Analysis of Quantitative Proteomics Data Generated via Multidimensional Protein Identification Technology. *Anal. Chem.* **2002**, in press.
 40. Li, L.; Masselon, C.; Anderson, G. A.; Paša-Tolić, L.; Lee, S.-W.; Shen, Y.; Zhao, R.; Lipton, M. S.; Conrads, T. P.; Tolić, N.; Smith, R. D. High-Throughput Peptide Identification from Protein Digests Using Data-Dependent Multiplexed Tandem FTICR Mass Spectrometry Coupled with Capillary Liquid Chromatography. *Anal. Chem.* **2001**, *73*, 3312–3322.
 41. Masselon, C.; Anderson, G. A.; Harkewicz, R.; Bruce, J. E.; Paša-Tolić, L.; Smith, R. D. Accurate Mass Multiplexed Tandem Mass Spectrometry for High-Throughput Polypeptide Identification from Mixtures. *Anal. Chem.* **2000**, *72*, 1918–1924.
 42. Gorshkov, M. V.; Masselon, C.; Anderson, G. A.; Udseth, H. R.; Harkewicz, R.; Smith, R. D. A Dynamic Ion Cooling Technique for FTICR Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 1169–1173.