Ultrasensitive proteomics using high-efficiency on-line micro-SPE-nanoLC-nanoESI MS and MS/MS

ARTICLE *in* ANALYTICAL CHEMISTRY · JANUARY 2004

Impact Factor: 5.64 · DOI: 10.1021/ac030096q · Source: PubMed

9 AUTHORS, INCLUDING:

Christophe Masselon
Atomic Energy and Alternative Energies Com…
71 PUBLICATIONS **3,412** CITATIONS
SEE PROFILE

Kim Hixson
Pacific Northwest National Laboratory
41 PUBLICATIONS **2,027** CITATIONS
SEE PROFILE

Gordon A Anderson
Pacific Northwest National Laboratory
125 PUBLICATIONS **5,404** CITATIONS
SEE PROFILE

Richard D Smith
Pacific Northwest National Laboratory
1,131 PUBLICATIONS **45,995** CITATIONS
SEE PROFILE

Anal. Chem. **2004,** *76,* 144−154

Ultrasensitive Proteomics Using High-Efficiency On-Line Micro-SPE-NanoLC-NanoESI MS and MS/MS

Yufeng Shen, Nikola Tolić, Christophe Masselon, Ljiljana Paša-Tolić, David G. Camp, II, Kim K. Hixson, Rui Zhao, Gordon A. Anderson, and Richard D. Smith*

Biological Sciences Division, Pacific Northwest National Laboratory, Richland, Washington 99352

Ultrasensitive nanoscale proteomics approaches for characterizing proteins from complex proteomic samples of <50 ng of total mass are described. Protein identifications from 0.5 pg of whole proteome extracts were enabled by ultrahigh sensitivity (<75 zmol for individual proteins) achieved using high-efficiency (peak capacities of ~10^3) 15-μm-i.d. capillary liquid chromatography separations (i.e., using nanoLC, ~20 nL/min mobile-phase flow rate at the optimal linear velocity of ~0.2 cm/s) coupled on-line with a micro-solid-phase sample extraction and a nanoscale electrospray ionization interface to a 11.4-T Fourier transform ion cyclotron resonance (FTICR) mass spectrometer (MS). Proteome measurement coverage improved as sample size was increased from as little as 0.5 pg of sample. It was found that a 2.5-ng sample provided 14% coverage of all annotated open reading frames for the microorganism *Deinococcus radiodurans,* consistent with previous results for a specific culture condition. The estimated detection dynamic range for detected proteins was 10^5−10^6. An improved accurate mass and LC elution time two-dimensional data analysis methodology, used to both speed and increase the confidence of peptide/protein identifications, enabled identification of 872 proteins/run from a single 3-h nanoLC/FTICR MS analysis. The low-zeptomole-level sensitivity provides a basis for extending proteomics studies to smaller cell populations and potentially to a single mammalian cell. Application with ion trap MS/MS instrumentation allowed protein identification from 50 pg (total mass) of proteomic samples (i.e., ~100 times larger than FTICR MS), corresponding to a sensitivity of ~7 amol for individual proteins. Compared with single-stage FTICR measurements, ion trap MS/MS provided a much lower proteome measurement coverage and dynamic range for a given analysis time and sample quantity.

Present proteomic methods require significant sample quantities (e.g., 10 μg−1 mg), effectively prohibiting many important avenues of research where only smaller samples are available (e.g., studies of cells from small tissue samples, from single cells, or obtained using laser capture microdissection).[1−3] Thus, development of more sensitive, higher throughput, and higher dynamic range proteomic measurements are expected to enable many areas of biological research.

The sensitivity of LC−MS coupling via electrospray ionization (ESI)[4,5] is approximately inversely related to the LC flow rate. Greater sensitivity is achieved with the lower flow rates provided by very narrow diameter LC columns *until* the point where ionization efficiency becomes limited by the number of analyte species available.[6] Fourier transform ion cyclotron resonance (FTICR) MS is currently the most sensitive form of MS,[7] and detection limits of ~30 zmol have been reported for proteins.[8] The sensitivity of capillary separations with FTICR MS and tandem MS (MS/MS) for both peptide and protein analyses has been previously explored.[9−13] For example, capillary electrophoresis-FTICR MS has enabled protein detection at subfemtomole to attomole levels;[9−11] while capillary (50−75-μm i.d.) LC-FTICR MS has demonstrated 0.5-amol peptide detection.[12,13] But even greater sensitivity is desired to enable broader dynamic range proteomic studies, and reducing the LC column inner diameter is the most effective approach for improving sensitivity without trading off separation quality. For example, Haskins et al.[14] used a 25-μm-i.d. capillary LC column with ion trap MS/MS to enable identification of peptides at the ~60-amol level (a detection limit of 4 amol was estimated). However, the reported sensitivity, separation quality, integrated column-ESI emitter (tip) design, and limited

* Corresponding author. E-mail: rds@pnl.gov.

(1) Heppner, G. H.; Miller, F. R. *Int. Rev. Cytol.* **1998**, *177*, 1−56.
(2) Emmert-Buck, M. R.; Bonner, R. F.; Smith, P. D.; Chuaqui, R. F.; Zhuang, Z.; Goldstein, S. R.; Weiss, R. A.; Liotta, L. A. *Science* **1996**, *274*, 998−1001.
(3) Bonner, R. F.; Emmert-Buck, M. R.; Cole, K.; Pohida, T.; Chuaqui, R.; Goldstein, S.; Liotta, L. A. *Science* **1997**, *278*, 1481−1483.
(4) Oosterkamp, A. J.; Gelpí, E.; Abian, J. *J. Mass Spectrom.* **1998**, *33*, 976−983.
(5) Bateman, K. P.; White, R. L.; Thibault, P. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 307−315.
(6) Wahl, J. H.; Goodlett, D. R.; Udseth, H. R. Smith, R. D. *Electrophoresis* **1993**, *14*, 448−457.
(7) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Mass Spectrom. Rev.* **1998**, *17*, 1−35.
(8) Belov, M. E.; Gorshkov, M. V.; Udseth, H. R.; Anderson, G. A.; Smith, R. D. *Anal. Chem.* **2000**, *72*, 2271−2279.
(9) Hofstadler, S. A.; Severs, J. C.; Smith, R. D.; Swanek, F. D.; Ewing, A. G. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 919−922;
(10) Wahl, J. H.; Goodlett, D. R.; Udseth, H. R.; Smith, R. D. *Anal. Chem.* **1992**, *64*, 3194−3196.
(11) Valaskovic, G. A.; Kelleher, N. K.; McLafferty, F. W. *Science* **1996**, *273*, 1199−1202.
(12) Martin, S. E.; Shabanowitz, J.; Hunt, D. F.; Marto, J. A. *Anal. Chem.* **2000**, *72*, 4266−4274.
(13) Quenzer, T. L.; Emmett, M. R.; Hendrickson, C. L.; Kelly, P. H.; Marshall, A. G. *Anal. Chem.* **2001**, *73*, 1721−1725.
(14) Haskins, W. E.; Wang, Z.; Watson, C. J.; Rostand, R. R.; Witowski, S. R.; Powell, D. H.; Kennedy, R. T. *Anal. Chem.* **2001**, *73*, 5005−5014.

10.1021/ac030096q CCC: $27.50
Published on Web 11/25/2003

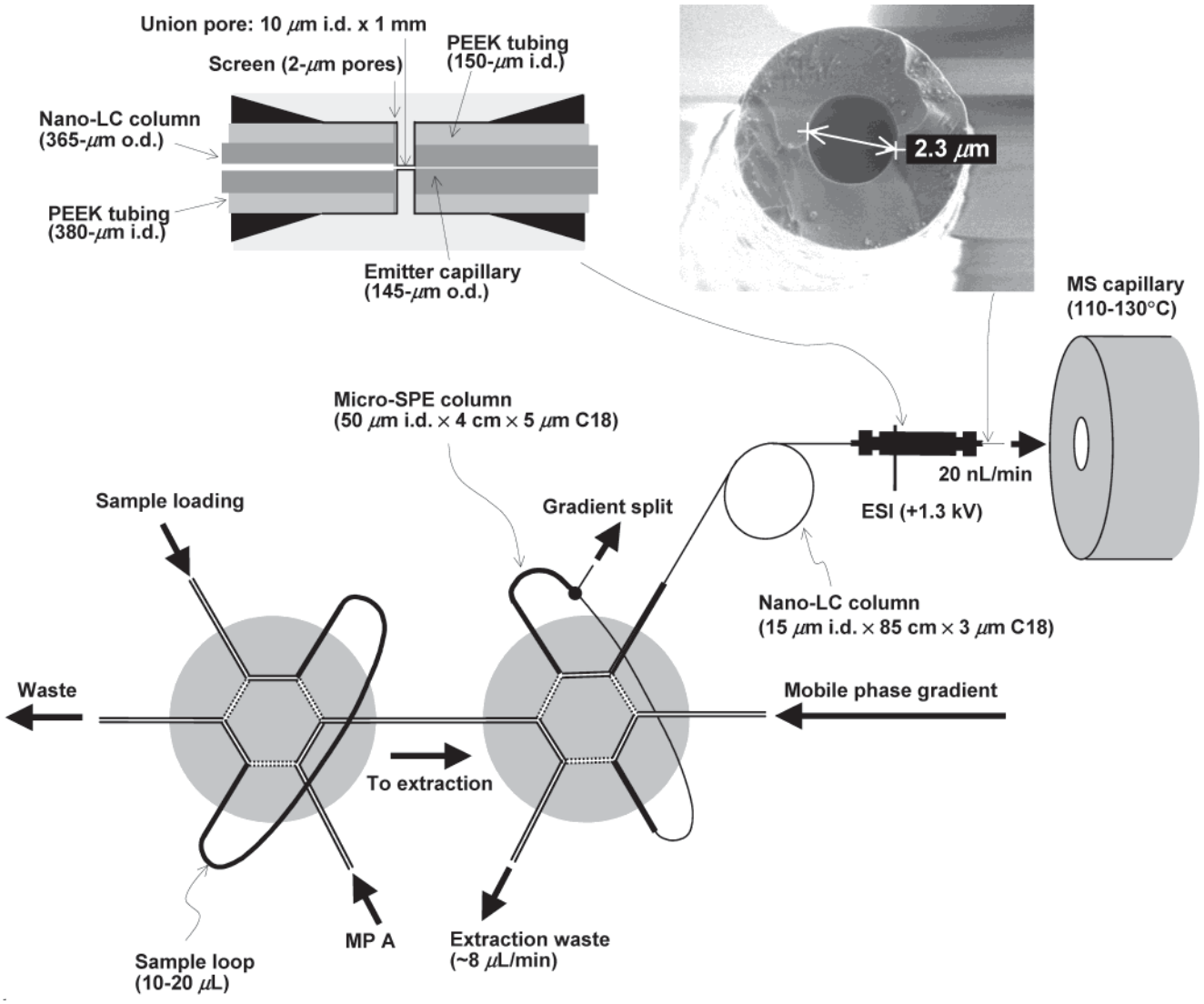

Figure 1. Schematic diagram of high-efficiency microSPE-nanoLC-nanoESI used for MS and MS/MS ultrasensitive proteomic analyses in this study (see text for details).

volume sample-loading capability (~1.8 μL) are generally not well suited for routine measurements of small proteomic samples. In previous studies with capillary LC separations, we achieved high separation efficiencies (peptide peak capacities of ~10^3) using nanoscale LC (nanoLC) with 15-μm-i.d. columns (which provided flow rates of ~20 nL/min at the optimal mobile-phase linear velocity of ~0.2 cm/s) based upon a design using replaceable ESI emitters.[15] The on-line coupling of micro-solid-phase extraction (microSPE) enabled ~400-fold faster sample loading flow rates compared to the LC mobile-phase flow rate (e.g., 8 μm/min sample loading for the 15-μm-i.d. nanoLC), allowing much faster introduction of large sample volumes.[16]

Our previous work also demonstrated that separation efficiency seriously affects the attainable proteome measurement coverage due to the high complexity of typical proteomic samples and the very different relative abundances of their components. For example, the number of detected peptides was increased by ~3-fold when the separation peak capacities were improved from ~10^2 to ~10^3.[17,18] Higher efficiency separations with MS/MS increase the chances of detecting lower abundance peptides, and an improvement from ~550 to ~1000 in separation peak capacity was found to approximately double the number of identified peptides.[16] High-efficiency capillary LC separations at 10 000 psi with MS/MS have enabled us to identify ~1500 peptides in a 6-h run.[16] With moderate-efficiency LC (peak capacities of ~10^2), sample fractionation prior to LC−MS/MS measurements is generally applied to achieve improved proteome coverage, albeit with generally lengthened analysis times and increased sample consumption.[19−21] High-efficiency single-dimension LC provides greater amounts of information from a limited sample size and

(15) Shen, Y.; Zhao, R.; Berger, S. J.; Anderson, G. A.; Rodriguez, N.; Smith, R. D. *Anal. Chem.* **2002**, *74*, 4235−4249.
(16) Shen, Y.; Moore, R. J.; Zhao, R.; Blonder, J.; Auberry, D. L.; Masselon, C.; Paša-Tolić, L.; Hixson, K. K.; Auberry, K. J.; Smith, R. D. *Anal. Chem.* **2003**, *75*, 3596−3605.
(17) 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.; Smith, R. D. *Anal. Chem.* **2001**, *73*, 1766−1775.
(18) 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. *Anal. Chem.* **2001**, *73*, 3011−3021.
(19) Washburn, M. P.; Wolters, D.; Yates, J. R. *Nat. Biotechnol.* **2001**, *19*, 242−247.

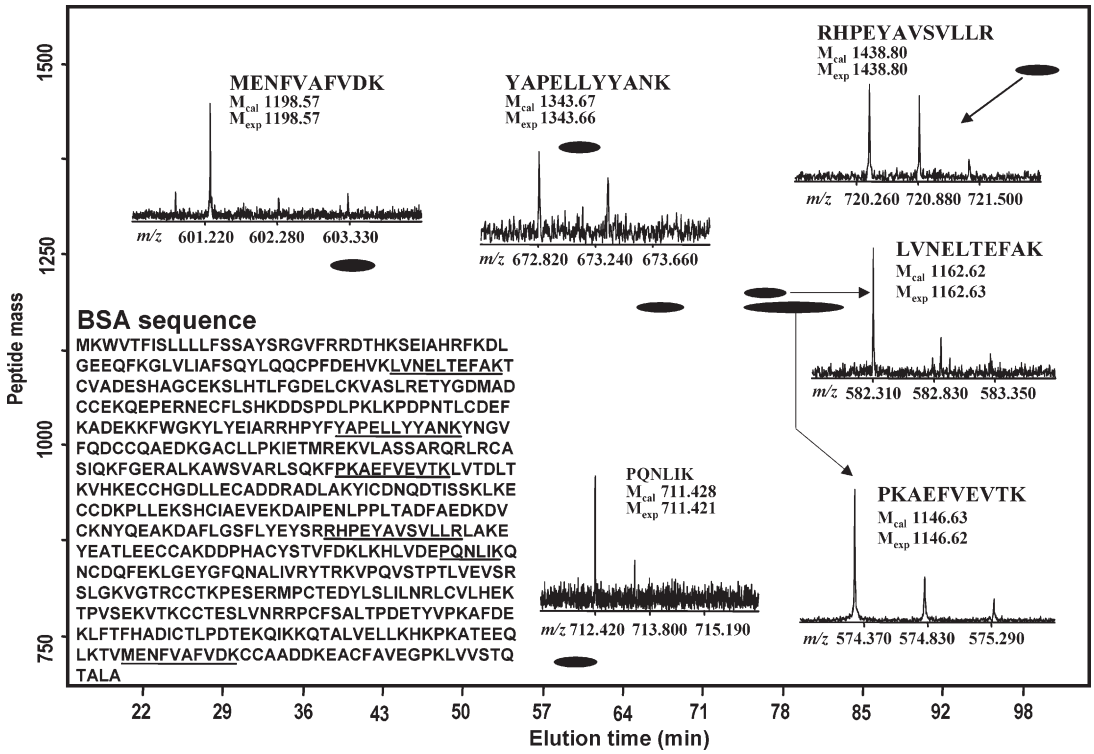

Figure 2. Results for injection of a sample containing 75 zmol of a BSA tryptic digest (10-μL solution) used to evaluate sensitivity for FTICR detection. Peptide assignments used the mass and LC elution time tag methodology described in the text (MMA values used for these assignments are indicated). The location of the peptides according to their mass and relative elution times is indicated by "spots" in the 2D display. The most abundant peptide (PKAEFVEVTK) was detected with S/N of 23, leading to an estimated detection limit of ~10 zmol.

time, while multiple-dimensional separations will generally provide more information (e.g., proteome coverage) if the sample size and analysis time are not constrained. In comparison to commercially available multiple-dimensional LC systems (generally based upon relatively large diameter columns), achieving higher efficiency separations using higher LC pressures has to date required significant expertise to prepare high-efficiency capillary columns, minimize dead volumes, etc. However, these issues are counterbalanced by the attractions of single-dimension LC separations: minimization of specific sample loses from added chromatographic dimensions and improved measurement sensitivity and throughput.

In this report, we describe ultrasensitive nanoscale proteomic analyses based on single-dimension high-efficiency nanoLC combined with both FTICR MS and ion trap MS/MS. The instrumental arrangement reported has demonstrated a ruggedness comparable to conventional capillary LC, but with greatly improved sensitivity. The measurement dynamic range, coverage, and throughput for very small proteomic samples (e.g., pg−ng range) are examined. The instrumental approach is demonstrated in the context of a refinement of the accurate mass and time tag approach for high-throughput proteomics developed at our laboratory.[22−24]

EXPERIMENTAL SECTION

On-Line MicroSPE-NanoLC-NanoESI MS and MS/MS Studies. Figure 1 depicts aspects of the instrumentation developed in this work. A 5-μm C18 particle (300-Å pore size, Phenomenex, Terrence, CA) packed 4 cm × 50 μm i.d. fused-silica capillary (365-μm o.d., Polymicro Technologies, Phoenix, AZ) was used as an on-line microSPE "precolumn" for sample trapping. The sample solution flow rate through the microSPE column was ~8 μL/min at 10 000 psi. Thus, <2 min was required for loading 10-μL samples,[16] allowing multiple introductions and sample washings from an autosampler or syringe to minimize the overall sample losses (the sample remaining in the autosampler or syringe is estimated at <5% after each loading). The loaded sample was then switched on-line to a 3-μm C18 particle (300-Å pore size, Phenomenex) packed 85 cm × 15 μm i.d. fused-silica capillary (365-μm o.d., Polymicro Technologies, prepared as described previously[15]), where the mobile-phase flow rate through the nanoLC separation column is ~20 nL/min (measured using dual-UV detectors, as described elsewhere[15]). The nanoLC operation used two 10 000 psi positive-feedback automatic six-port switching valves (Valco Instruments, Houston, TX), and the packed capillaries for both microSPE and nanoLC were directly connected to the valves using PEEK tubing (380-μm i.d., Upchurch

(20) McDonald, W. H.; Ohi, R.; Miyamoto, D. T.; Mitchison, T. J.; Yates, J. R. *Int. J. Mass Spectrum.* **2002**, *219*, 245−251.
(21) Peng, J.; Elias, J. E.; Thoreen, C. C.; Licklider, L. J.; Gygi, S. P. *J. Proteome Res.* **2003**, *2*, 43−50.
(22) Smith, R. D.; Paša-Tolić, L.; Lipton, M. S.; Jensen, P. K.; Anderson, G. A.; Shen, Y.; Conrads, T. P.; Udseth, H.; Harkewicz, R.; Belov, M. E.; Masselon, C.; Veenstra, T. D. *Electrophoresis* **2001**, *22*, 1652−1668.
(23) Smith, R. D.; Anderson, G. A.; Lipton, M. S.; Paša-Tolić, L.; Shen, Y.; Conrads, T. P.; Veenstra, T. D.; Udseth, H. *Proteomics* **2002**, *2*, 513−523.
(24) Lipton, M. S.; Paša-Tolić, L.; Anderson, G. A.; Auberry, D. L.; Battista, J. R.; Daly, M. J.; Fredrickson, J.; Hixson, K. K.; Kostandarithes, H.; Conrads, T. P.; Masselon, C.; Markille, M. L.; Moore, R. J.; Romine, M. F.; Shen, Y.; Toliç, N.; Udseth, H.; Veenstra, T. D.; Venkateswaran, A.; Wong, K. K.; Zhao, R.; Smith, R. D. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11049−11054.

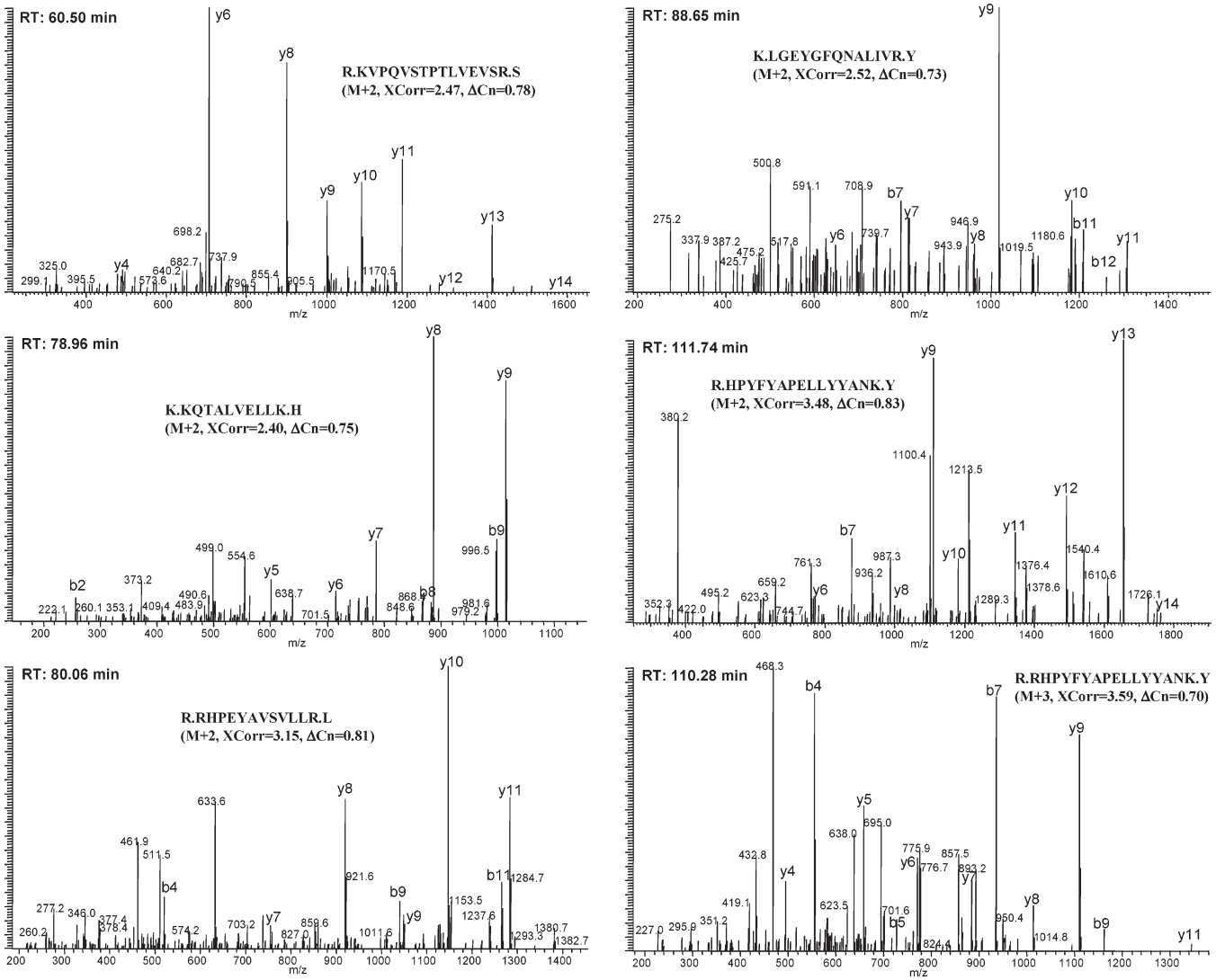

Figure 3. Results from loading 7 amol of the BSA tryptic digest and using ion trap-MS/MS detection. Conditions: same as for Figure 2 except loading 7 amol of the BSA tryptic digest and using MS/MS conditions described in the text. Peptide assignments were made using SEQUEST algorithm and typical assignment rules[17] (the triply charged peptide RHPYFYAPELLYYANK identified with an X_{Corr} of 3.5, close to the assignment acceptance threshold, is also given).

Scientific, Oak Harbor, WA) with a stainless steel mesh screen (2-μm pores, Valco) in the valve adapter to retain the particles. A 10 ± 1 μm i.d. union developed "in-house" was used to connect the 15-μm-i.d nanoLC column outlet to a nanoESI emitter (using another PEEK tube: 150-μm i.d., Upchurch Scientific). The nanoESI emitter had a ~2-μm-i.d orifice drawn from a 10 μm i.d. × 145 μm o.d. fused-silica capillary (Polymicro Technologies). Purified H_2O (Nanopure Infinity ultrapure water system, Barnstead, Newton, WA) was used as the mobile phase A with addition of 0.05% HPLC-grade trifluroacetic acid (TFA, v/v, Aldrich, Milwaukee, WI) and 0.2% HPLC-grade acetic acid (v/v, Aldrich). An aqueous solution of HPLC-grade acetonitrile (ACN) (Aldrich) (ACN/H_2O/TFA, 90:10:0.1, v/v/v) was used as mobile phase B. The mobile phases were pumped at 10 000 psi using two Isco LC pumps (model 100DM with PEEK-coated pressure transducers, Isco, Lincoln, NE) to the gradient at constant pressure (i.e., 10 000 psi). The exponential gradient ranged from 0 to 70% mobile phase B in ~180 min, and a 5-μm C18 packed ~13 cm × 150 μm i.d. capillary was used as a gradient splitter. We have previously demonstrated that the on-line coupled microSPE-nanoLC provides a separation peak capacity of ~1000 for separations of ~3 h.[16]

A previously described 11.4-T FTICR mass spectrometer[25,26] was employed for proteomic analyses based upon accurate mass and elution time information. The ion accumulation time was 4.0 s, with an overall spectrum acquisition time of ~7 s. The ESI emitter orifice was positioned at ~1.5 mm from the heated (105 °C) MS inlet capillary, and 1.3 kV was applied to the connection union between the nanoLC column and nanoESI emitter. MS/MS experiments were performed with a Finnigan ion trap mass spectrometer (model LCQ XP, ThermoQuest Corp., San Jose, CA). The heated capillary temperature and ESI voltage were optimized at 120 °C and 1.8 kV, respectively. The m/z range of 500−2000 was used for a spectrum acquisition cycle of ~9 s, during which

(25) Bruce, J. E.; Anderson, G. A.; Wen, J.; Harkewicz, R.; Smith, R. D. *Anal. Chem.* **1999**, *71*, 2595−2599.
(26) Harkewicz, R.; Belov, M. E.; Anderson, G. A.; Paša-Tolić, L.; Masselon, C. D.; Prior, D. C.; Udseth, H. R.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 144−154.

time the three most abundant ions from each MS analysis were selected for MS/MS using a collision energy setting of 45% with a 2-min dynamic exclusion for data-dependent analyses.

Data Analysis. Data processing was completed using software developed at our laboratory. Briefly, we 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.[27] The output data files were visualized as 2D displays of peptide neutral mass versus LC elution time. Single-stage MS peptide assignments were made by searching the mass versus time 2D database that had been generated from MS/MS sample analyses, using conservative peptide assignment rules[19,28] and the SEQUEST algorithm (ThermoQuest Corp.). The tentatively identified peptides were further refined (i.e., culled) by retaining only those peptides identified in more than one LC−MS/MS run and eliminating peptides having inconsistent LC relative retention times (RRT) in different runs (e.g., where the same peptide was nominally identified in different runs but with significantly different RRT). For *Deinococcus radiodurans*, the result was a set of 9808 tryptic peptides from 1982 ORFs (i.e., covering ~63% of the total of 3116 predicted ORFs), ~7-fold less than the ~70 000[24] identified using looser criteria (X_{Corr} >2). The exact molecular masses of these 9808 peptides were then calculated according to their chemical composition, and the array of peptide molecular masses and corresponding LC RRT values formed the MT-2D database used for subsequent identifications.

To determine $^{14}N/^{15}N$ peptide pairs, we first assembled all measured neutral masses into "unique mass classes" (UMCs), defined as masses within a large mass error (e.g., 25 ppm) and detected in consecutive spectra, so as not to overestimate the number of species detected. Thus, a UMC constitutes the signals from the elution of a specific peptide in terms of its neutral mass (and without the contributions from less abundant isotopes). This set of UMC putative peptides was then searched for pairs of peptides whose mass differences corresponded to unit mass multiples of the $^{14}N - {}^{15}N$ mass difference for the specific peptide and having very similar elution times. The abundance ratios were then calculated for all $^{14}N/^{15}N$ peptide pairs as the ratios of sum of the ^{14}N signal intensities across the peak to the sum ^{15}N signal intensities.

Sample Preparations. A *D. radiodurans* lysate tryptic digest was used to evaluate the sensitivity of proteome analyses. The sample contained equal amounts of cell cultured in ^{14}N- and ^{15}N-enriched media (as used for protein relative abundance quantitation[22,23]) and was prepared as follows: a culture of *D. radiodurans* strain R1 was inoculated with 1 mL of starter culture in TGY media (0.5% tryptone, 0.3% yeast extract, and 0.1% glucose in deionized water). This culture was grown to stationary phase (OD_{600} 0.9) at 30 °C while shaking (225 rpm). A second culture of *D. radiodurans* strain R1 was grown in a defined medium[29] designed for $^{15}N-$ incorporation for all expressed proteins. The defined medium was supplemented with L-glutamic acid (^{15}N, 95−99%), L-methionine (^{15}N, 95−99%), and ammonium sulfate (^{15}N, 98%+) (Cambridge Isotope Laboratories, Andover, MA). The ^{15}N-labeled cells were grown to late logarithmic phase (OD_{600} 0.6) at 30 °C with shaking (225 rpm). Both cell sets were harvested by centrifugation at 4000*g* at 4 °C, washed three times with PBS, aliquoted, and quick frozen for storage at −80 °C. A total of 15.3 and 15.5 mg (wet weight) of ^{14}N- and ^{15}N-labeled cells, respectively, were combined with 200 μL of 50 mM Tris buffer (pH 8.0). Cell lysis used mechanical agitation with a Mini-Beadbeater (Biospec Products, Bartlesville, OK) in three 90-s cycles at 4500 rpm in the presence of 0.1-mm zirconium/silica beads. The cell lysate was recovered by centrifugation at 10000*g* for 2 min. The protein concentration of the cell lysate was determined via BCA protein assay (Pierce, Rockford, IL). A denaturation solution of 150 μL (7 M urea, 2 M thiourea, 5 mM DTT in 40 mM Tris buffer, pH 8.0) was added to the lysate. The protein solution was then incubated at 100 °C for 3 min. After adding 1.5 mL of 40 mM Tris buffer (pH 8.0) and 1.5 μL of 1 M calcium chloride, proteins were digested using sequencing grade modified bovine pancreas trypsin (Promega, Madison, WI) (trypsin/protein, 1:50, w/w) at 37 °C for 3 h. Digested proteins were desalted with a 1-mL SPE C-18 prep-column (Supelco, Bellefonte, PA) followed by evaporation to dryness in a Speed-Vac SC110 (ThermoSavant, Holbrook, NY). The peptides were stored at −80 °C and resuspended in 200 μL of the chromatographic mobile phase A prior to analysis.

A bovine serum albumin (BSA, Sigma, St. Louis, MO) tryptic digest prepared as described above for *D. radiodurans* proteins was also used for sensitivity evaluation.

Table 1. A Portion of the Molecular Masses Detected in a SPE-Nanoscale LC-Nano-FTICR MS Analysis from a 10-ng *D. Radiodurans* Lysate Tryptic Digest Sample[a]

species	mass[b]	RRT	species	mass[b]	RRT
1	1072.459	0.064	14	1072.515	0.183
2	1072.473	0.228	15	1072.517	0.039
3	1072.476	0.090	16	1072.517	0.105
4	1072.494	0.085	17	1072.518	0.064
5	1072.496	0.131	18	1072.519	0.166
6	1072.497	0.227	19	1072.522	0.168
7	1072.500	0.037	20	1072.524	0.126
8	1072.500	0.086	21	1072.524	0.226
9	1072.502	0.157	22	1072.527	0.096
10	1072.506	0.035	23	1072.528	0.155
11	1072.514	0.158	24	1072.529	0.036
12	1072.515	0.096	25	1072.529	0.200
13	1072.515	0.102	26	1072.530	0.026

[a] A 10-ng sample was loaded on the microSPE column with a 10-μL solution and then on-line switched to nanoLC column; the separation was carried out at 10 000 psi with a mobile-phase gradient from A to 75% B in 200 min. Other conditions as described in the Experimental Section. [b] Detected masses are for neutral (zero charge) species after data processing to remove contributions from isotopic peaks (e.g., due to ^{13}C, ^{15}N, etc.).

RESULTS AND DISCUSSION

System Background. The MS and MS/MS backgrounds were initially examined using purified H_2O as the blank. For FTICR MS, data were collected over the m/z range of 500−2000 for a 10-μL blank, using the same LC conditions as used for the proteome samples, and were searched against appropriate peptide (*D. radiodurans* and BSA) databases employed for this work (see discussion below). No "hits" (based upon the peptide identification

(27) Horn, D. M.; Zubarev, R. A.; McLafferty, F. W. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 320−332.
(28) ftp://ftp.tigr.org/pub/data/dradiodurans/GDR.pep; http://www.ncbi.nlm.nih.gov/.
(29) Venkateswaran, A.; McFarlan, S. C.; Ghosal, D.; Minton, K. W.; Vasilenko, A.; Makarova, K.; Wackett, L. P.; Daly, M. J. *Appl. Environ. Microbiol.* **2000**, *66*, 2620−2626.

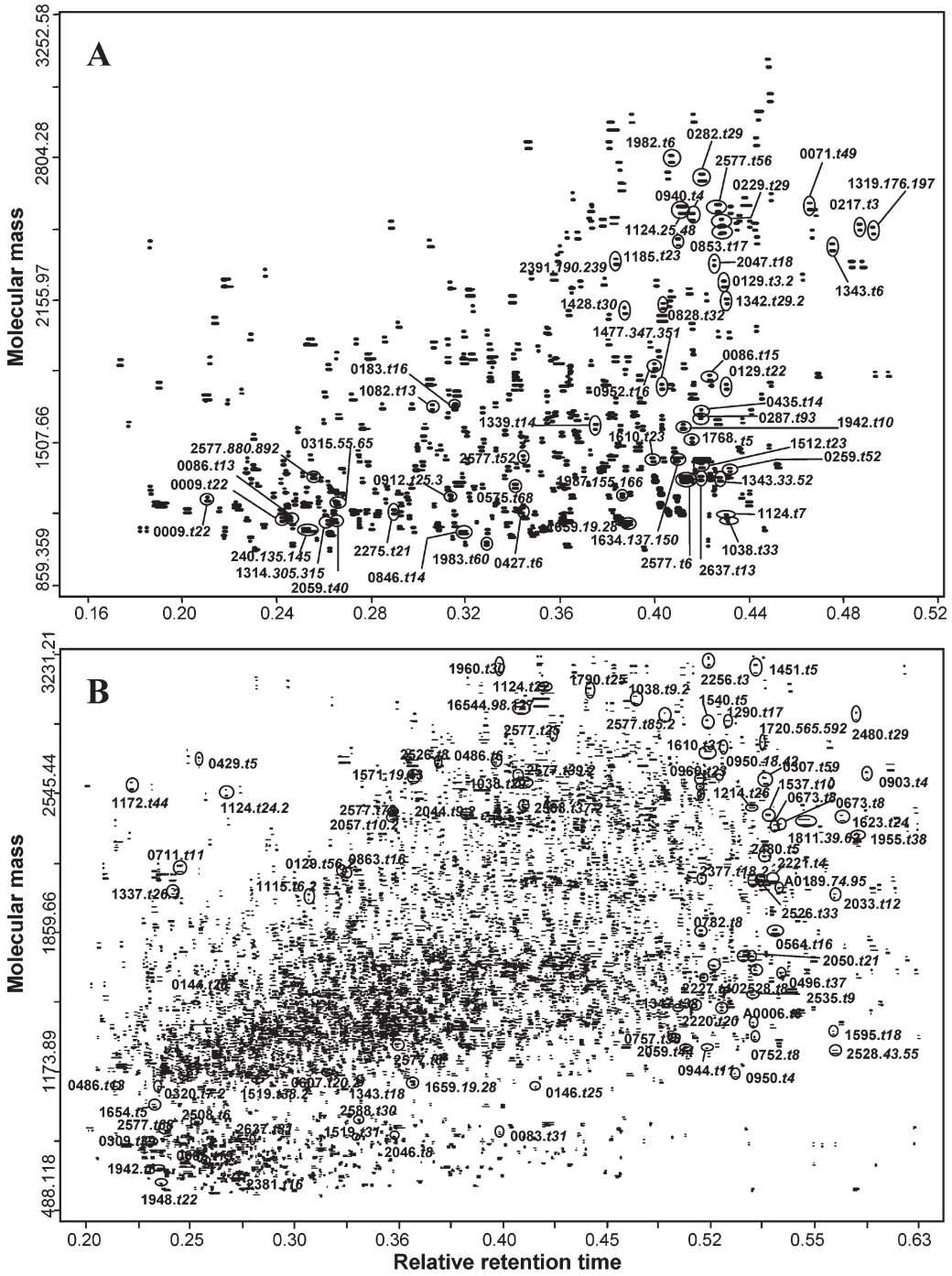

Figure 4. Ultrasensitive proteome measurements using microSPE/nanoLC-FTICR MS. Results from a proteomic analysis from (A) 5 pg and (B) 2.5 ng of a *D. radiodurans* lysate tryptic digest; other conditions same as for Figure 2. In the single 3-h run, 436 proteins were each identified with both ^{14}N- and ^{15}N-labeled peptide pairs from the 2.5-ng sample (B), providing the basis for simultaneous quantification of their relative abundances.[29] The identifications for a few peptides/proteins are indicated (see text).

criteria used) were obtained. The same experiment was performed with the nanoLC-ion trap MS/MS, and the data yielded no acceptable peptide assignments according to the peptide assignment rules.[19] These observations indicated that the instrumentation and data analysis approach used in this work are unlikely to produce spurious peptide/protein identifications due to contaminants or other "noise" contributions.

Sensitivities with NanoLC-FTICR MS and Ion Trap-MS/MS. The sensitivity achievable for protein identification was evaluated using a BSA tryptic digest. A 75-zmol injection in a 10-μL solution identified six BSA tryptic peptides by nanoLC−FTICR MS (Figure 2; based upon <10 ppm MMA and <0.05 TMA). Thus, we conservatively estimate the system sensitivity as 75 zmol for protein identification. Based upon the signal/noise ratio of 23 obtained for the peptide PKAEFVEVTK, we estimate a detection limit of ~10 zmol (~6000 molecules) for this peptide. Note, however, that analyses of such small samples (i.e., ~10 zmol) would likely be impractical in the absence of other peptides (i.e., the "carrier" effect) due to losses to chromatographic and other surfaces. For the nanoLC-ion trap-MS/MS, five doubly, one triply

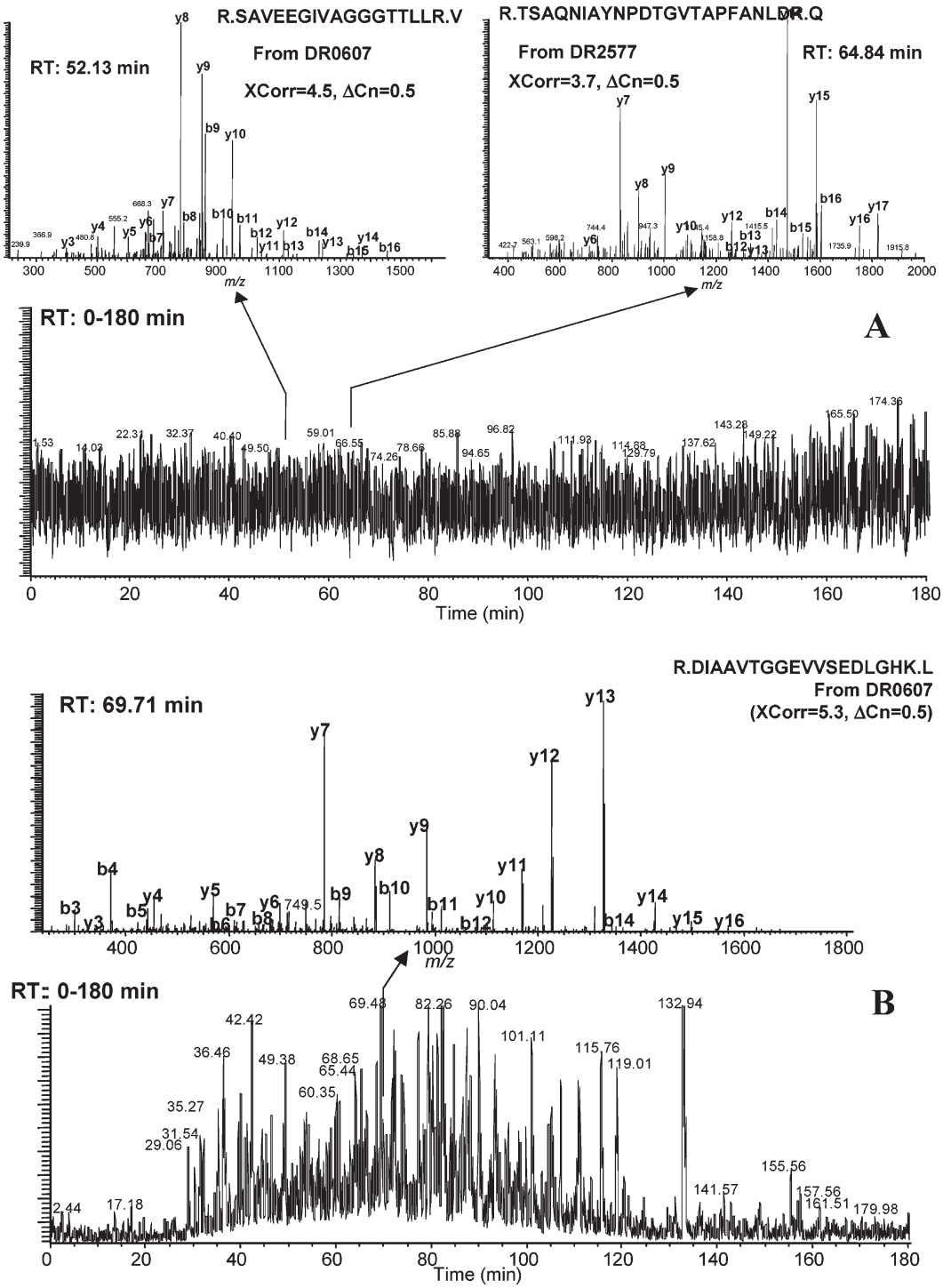

Figure 5. MicroSPE/nanoLC-ion trap-MS/MS from (A) 0.25- and (B) 10-ng injections of a *D. radiodurans* lysate tryptic digest sample. Other conditions same as in Figure 4.

charged tryptic peptides were confidently identified and manually confirmed from a 7-amol BSA tryptic digest sample, as shown in Figure 3. (It should be noted that the peptides identified with greatest sensitivity by MS versus MS/MS are generally different; only one peptide is common with the six peptides from Figure 2.) We thus estimate the corresponding protein identification sensitivity using the nanoLC-ion trap-MS/MS as ~7 amol. Similarly, the estimated signal/noise ratio suggests a peptide detection limit of ~1 amol. The sensitivity provided by the nanoLC-nanoESI-ion trap-MS/MS was ~100-fold less than with FTICR, but still improved by ~50-fold compared with results obtained using a 100-μm-i.d. column.[30]

The sensitivities obtained using nanoLC-MS and MS/MS can provide extremely low concentration detection limits in conjunction with microSPE. For example, analyses of a 250 aM solution with microSPE-nanoLC-FTICR and a 25 fM solution with microSPE-nanoLC-ion trap-MS/MS required only 5 min to load the corresponding 40-μL sample solution.

(30) Guček, M.; Gaspari, M.; Walhagen, K.; Vreeken, R. J.; Verheij, E. R.; van der Greef, J. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1448−454.

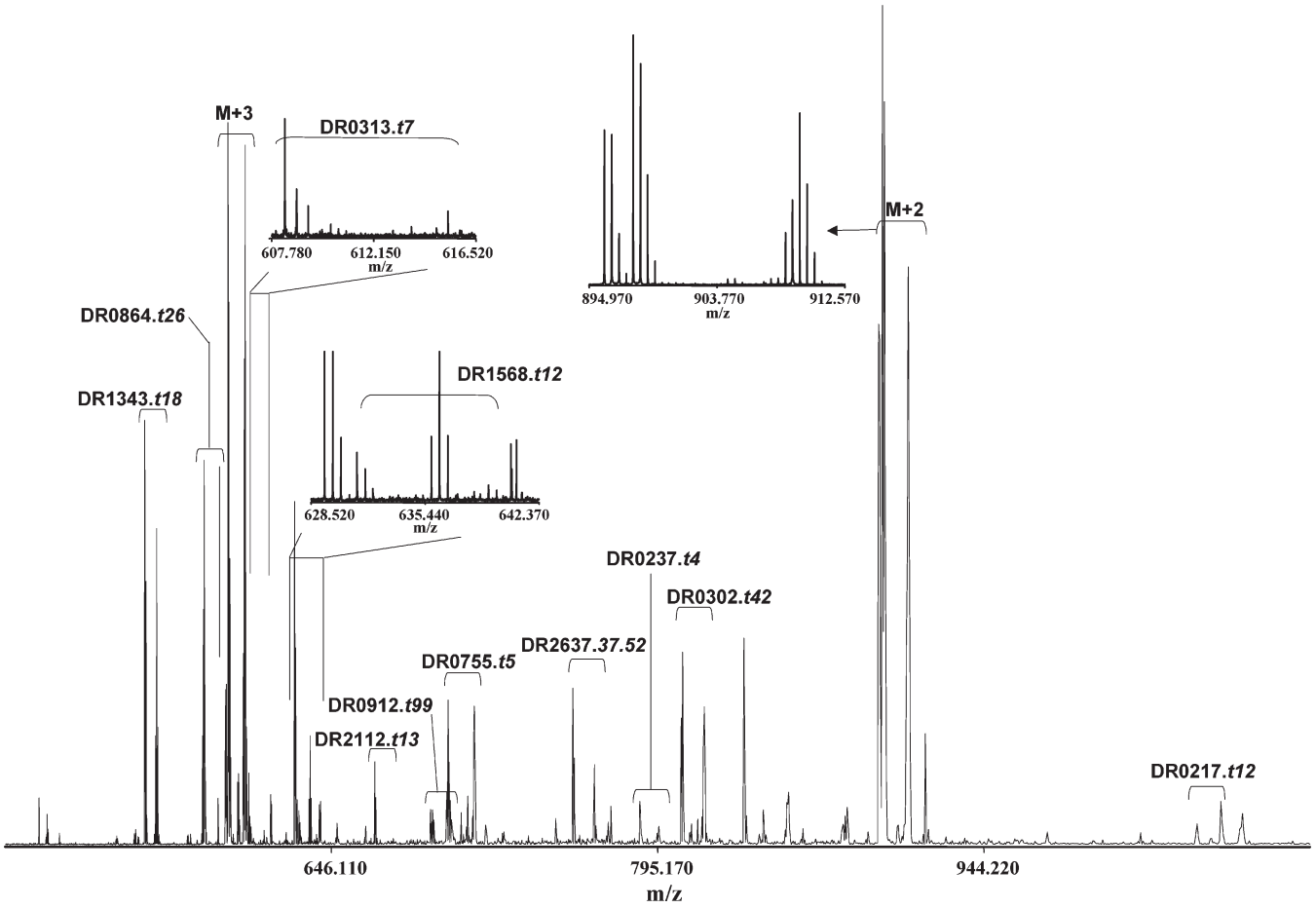

Figure 6. Peptide assignment and protein identification throughput in single spectrum supplied by the LC−MS approach. Conditions same as in Figure 4B (spectrum at RRT = 0.36).

The use of microSPE also facilitates quantitative transfer of small samples. For example, a 10-μL sample introduction could be followed by multiple washings (e.g., 3 times) to minimize losses to the sample vial, syringe, autosampler, connection tubing, and valve adapters and potentially enable transfer of >99% of the sample to the system.

System Ruggedness for Routine Ultrasensitive Application. The on-line coupling of microSPE to the 15 μm i.d. × 85 cm long column nanoLC resulted in an operational ruggedness equivalent to that of a 50-μm-i.d. column. In 12 months of service, encompassing >150 runs with tryptic digests of both microbial and mammalian proteomic samples, the microSPE column was replaced only five times. The 15-μm-i.d. nanoLC column clogged once at its inlet; however, the column was reused by cutting ~2 mm from its inlet end. The 2-μm-i.d. ESI emitter orifice never clogged during operation with high-pressure mobile phases, but did clog once after stopping the mobile-phase flow (the present design makes ESI emitter replacement convenient). The switching valve rotor needed to be replaced every ~2 months.

Single-Stage MS Peptide Assignment Using Molecular Mass and LC Retention Time Information. We used the peptide LC RRT values to improve the confidence in peptide assignments in a variation of our accurate mass and LC elution time measurement approach[18,22−24,31] to compensate for the lower mass measurement accuracy often obtained for very low abundance species.[32] Table 1 lists a subset of species detected from a 10-ng *D. radiodurans* lysate tryptic digest sample in the mass and separation region that was most congested. Species 7, 8, 12−14, 15, 16, 20, 21, 24, and 25 cannot be distinguished based solely upon molecular mass (with 1 ppm MMA). However, LC separation times reflect differences of peptide sequence (and even higher order structure) due to interactions with the chromatographic stationary-phase surface.[33,34] The specificity of peptide measurements in LC−MS is determined by both the MS MMA and the LC RRT measurement accuracy (TMA), and the overall peptide distinguishing capability (R) can be quantified as

$$R = 1/(\mathrm{MMA} \times \mathrm{TMA}) \qquad (1)$$

Equation 1 gives equal weight to both MMA and TMA. With 0.5 ppm overall (2D) specificity of MMA × TMA (i.e., $R = 2 \times 10^6$) resulting from a conservative 10 ppm MMA (allowing use of very low abundance peaks) and 0.05 TMA, the detectable peptides are distinguishable even for the most congested mass−time region of the analysis, as shown in Table 1. The probability of a false positive resulting from detection of a new peptide is only 1 in a

(31) Anderson, G. A.; et al. *Proceedings of the 46th ASMS Conference on Mass Spectrometry and Allied Topics*, 1998; p 1270.
(32) Chen, L.; Cottrell, C.; Marshall, A. G. *Chemom. Intell. Lab.* **1986**, *1*, 51−58.
(33) Meek, J. L. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 1632−1636.
(34) Regnier, F. E. *Science* **1987**, *238*, 319−323.

1000 (based on 10^4 detected peptides). It is noted that the MT-2D *R* value for high-efficiency LC-FTICR MS can be as high as 10^8 based upon 1 ppm MMA[25,35] and 0.01 TMA.[15] It should be noted that this number is larger than the MT-2D resolving power[17,18] since the MMA is up to 10^2-fold higher than the MS resolving power generally used. Additionally, this MT-2D specificity is statistically equivalent to 0.5 ppm MMA for peptide assignments without using TMA, but with somewhat different selectivity (i.e., some peptides can be distinguished by 0.5 ppm MMA, but not by 0.5 ppm MMA × TMA, and vice versa for others). This MT-2D methodology also opens up the use of more moderate MMA measurements, based upon relatively low cost mass spectrometers (e.g., quadrupole time-of-flight MS having MMA of ∼<10 ppm[15,36,37]) for proteomic analyses.

Ultrasensitive Nanoscale Proteomics Using FTICR MS. Figure 4 shows results from analysis of a $^{15}N/^{14}N$-labled *D. radiodurans* lysate tryptic digest where peptide/protein identifications were made using the MT-2D approach described above (i.e., 10 ppm MMA and 0.05 TMA). Figure 4 shows a 2D display of detected peptide where some peptide "spots" have been annotated with the ORF and the peptide identification (either as the tryptic peptide from the N-terminus, e.g., ".t23", or as the start and end amino acids, e.g., ".176.197", after the ORF designation). The sensitivity provided by the nanoLC-FTICR MS led to the identification of 53 *D. radiodurans* ORFs in both their $^{15}N/^{14}N$-labled versions from a 5-pg sample (Figure 4A). From a 0.5-pg sample, we were still able to identify the most abundant proteins (e.g., DR2577, DR0606, and DR0217). These proteins accounted for only ∼1% (w/w) of the total sample (i.e., <10-fg sensitivity). Note that the sample amounts used here are significantly smaller than the protein content of a single mammalian cell (∼50 pg for an "average" eukaryotic cell[38]). In the ∼3-h LC−MS analysis of a 2.5-ng sample, 436 ORFs were identified and simultaneously quantified based upon the relative abundances of the $^{5}N/^{14}N$-labled peptide pairs (Figure 4B).

Ultrasensitive Nanoscale Proteomics Using LC-Ion Trap-MS/MS. The sensitivity of the nanoLC-ion trap-MS/MS approach enabled identification of the most abundant protein (DR2508) from a 50-pg sample (compared to 0.5 pg required for FTICR). While no apparent peaks are evident from the base peak chromatogram for a 0.25-ng sample, the MS/MS spectra quality was sufficient for confident peptide assignments in many cases (Figure 5A). The ∼3-h analysis of a 10-ng sample, the largest sample size used in this work, identified ^{14}N-labeled peptides from 141 ORFs. This number is significantly smaller than that obtained with FTICR, even for the 2.5-ng sample, due to both the lower ion trap-MS/MS sensitivity and the slower data acquisition rate (i.e., one peptide per MS/MS spectrum versus multiple peptides per MS spectrum). Increasing the sample amount accompanied by either extension of the LC gradient elution (e.g., to 5−6 h) or fractionation of the sample prior to LC−MS/MS can increase MS/MS spectrum acquisition but at the cost of decreased sensitivity and throughput. .

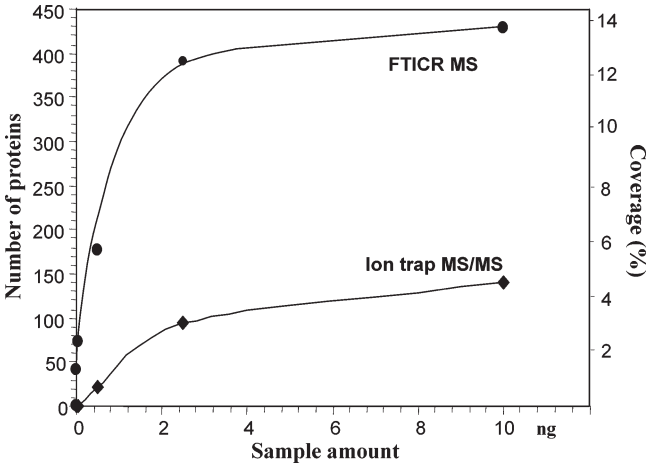

Figure 7. Proteome measurement coverage for various sample amounts obtained using microSPE-nanoLC-nanoESI with FTICR MS and ion trap-MS/MS. For FTICR MS results, proteins were identified using the accurate mass and LC elution time tag approach and the $^{15}N/^{14}N$ paired peptides, as discussed in the text; for ion trap-MS/MS, ^{14}N peptides were assigned by searching against the database using Sequest.

Proteome Measurement Throughput in Ultrasensitive Proteomics Measurements. Proteome measurements using LC−MS offers throughput superior to 2D-gel MS approaches,[39] and single-stage MS provides the greatest throughput since it can detect $>10^2$ peptides in a single spectrum.[25] Figure 6 exemplifies protein identification throughput for a relatively data dense region of a typical analysis. In one spectrum, 22 peptide pairs covering 11 ORFs were identified, and several other stable-isotope-labeled peptide pairs were detected but not identified. The nonuniform distribution of peptides along the LC elution resulted in an overall identification throughput of ∼290 proteins/h (i.e., 872 proteins in a 3-h run) for this sample. For ion trap-MS/MS measurements, the maximum sample size used (i.e., 10 ng) yielded a protein identification throughput of ∼50 proteins/h. This MS/MS throughput potentially could be improved by using multiplexed MS/MS with higher MMA mass spectrometers.[40−42] Algorithms and programs for database searching for confident peptide identification are presently being refined in our laboratory for this and related applications.

Proteome Measurement Dynamic Range and Coverage. Our studies have indicated that 10^5−10^6 dynamic range (between the most and least abundant species detected) can be achieved for a 5-ng *D. radiodurans* sample using microSPE-nanoLC-FTICR MS, which is also consistent with observations that samples in the 2.5−10-ng size range yielded similar protein coverage (see results below). The dynamic range provided by the nanoLC-ion trap-MS/MS was estimated as $\sim 10^4$, consistent with results reported by Yates and co-workers.[43]

(35) Masselon, C.; Tolmachev, A. V.; Anderson, G. A.; Harkewicz, R.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 99−106.
(36) Strittmatter, E. F.; Rodriguez, N.; Smith, R. D. *Anal. Chem.* **2003**, *75*, 460−468.
(37) Strittmatter, E. F.;. Ferguson, P. L.; Tang, K.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 980−991.
(38) Zhang, Z.; Krylov, S.; Arriaga, E. A.; Polakowski, R.; Dovichi, N. J. *Anal. Chem.* **2000**, *72*, 318−322.
(39) Quadroni, M.; James, P. *Electrophoresis* **1999**, *20*, 664−677.
(40) Masselon, C. D.; Anderson, G. A.; Harkewicz, R.; Bruce, J. E.; Paša-Tolić, L.; Smith, R. D. *Anal. Chem.* **2000**, *72*, 1918−1924.
(41) Li, L.; Masselon, C. D.; Anderson, G. A.; Paša-Tolić, L.; Lee, S. W.; Shen. Y.; Zhao, R.; Lipton, M. S.; Conrads, T. P.; Tolic, N.; Udseth, H.; Smith, R. D. *Anal. Chem.* **2001**, *73*, 3312−3322.
(42) Masselon, C. D.; Paša-Tolić, L.; Lee, S.-W.; Li, L.; Anderson, G. A.; Harkewicz, R.; Smith, R. D. *Proteomics* **2003**, *3*, 1279−1286.

Table 2. Protein Identification Overlap for Two Ion Trap-LC-MS/MS Runs[a]

run 1			run 2		
X_{Corr}	peptide references	ORFs	ORFs	peptide references	X_{Corr}
3.3	K.VQDEVEIVGLTDTRK.T	DR0309	DRO309	K.VQDEVEIVGLTDTRK.T	4.8
2.6	K.YEFPGDDLPVVK.G		DR0404	F.NNHAADALALALTHLAHAPMQER.S	2.2
3.9	R.TALQNAASIGALILTTEAIVSDKPEK.A	DR0607	DR0607	K.TNDITGDGTTTATVLGQAIVK.E	5.5
3.2	R.VDKDTTGSANPVYLDGEPIR.I	DR1185		R.SAVEEGIVAGGGTTLLR.V	4.5
2.6	K.LFGEEPSQQIAEDLQK.F	DR2377	DR1185	R.TLGAANTTYSGIAPGSYPVR.V	4.0
3.1	R.FDNVGPTITGSVIR.D	DR2508		R.GGHEVVIIAPGYR.T	2.8
2.9	R.TPWVELGSSDTANTQQK.F		DR1245	K.EVQLDIEENGGQR.F	2.2
4.1	R.DQLGNETASATYELVR.F		DR2508	R.FDNVGPTITGSVIR.D	3.2
3.1	R.AGNATISTTPIVVNAK.I			R.GDVNVFTGNPSLQDR.E	3.3
3.7	R.TSAQNIAYNPDTGVTAPFANLDR.Q	DR2577	DR2577	R.TSAQNIAYNPDTGVTAPFANLDR.Q	3.7
3.1	K.NATTQFSVGNNPVIVTLGQQQK.F			R.DIDGVGGVLNPAVNLDQSR.T	3.0
2.4	R.DIDGVGGVLNPAVNLDQSR.T				
2.4	R.NPAQLPGVIASAALDINNDK.A	DRA0283			

[a] Experimental conditions are the same as for Figure 3 except for loading 0.25 ng of the sample to system; all peptides were detected with +2 charge states.

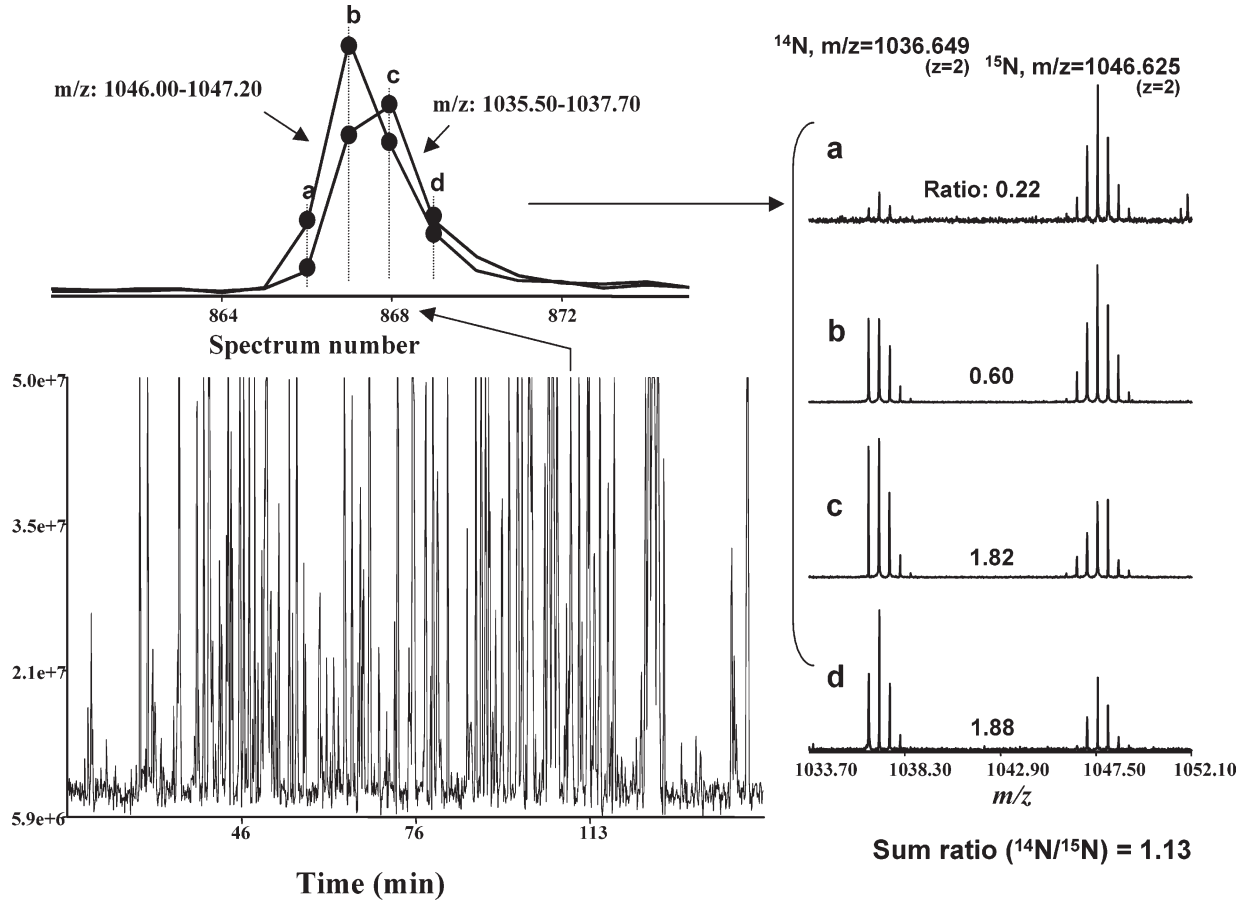

Figure 8. Reversed-phase capillary LC elution profiles for $^{14}N/^{15}N$-labeled peptide pairs detected from injection of a 0.5-ng $^{14}N/^{15}N$-labeled *D. radiodurans* lysate tryptic digest sample. Other conditions same as in Figure 4.

The organism being studied, the details of sample preparation, and the measurement approach all influence proteome coverage. In this work, proteome measurement coverage increased with the amount of *D. radiodurans* sample for both FTICR single-stage MS and ion trap-MS/MS, as shown in Figure 7. The FTICR MS analyses yielded ~14% coverage of all predicted ORFs in a single 3-h nanoLC-FTICR run for 2.5−10 ng of the sample, which was roughly comparable to our previous work with this organism (since a limited fraction of the proteome will generally be expressed under a specific condition);[24] In contrast, the nanoLC-ion trap-MS/MS analyses provided ~4.5% coverage from a 10-ng sample in the same analysis time.

The overlap of proteins identified by FTICR MS and ion trap-MS/MS was examined and found to be >70%. The differences in detected proteins from different nanoLC-ion trap-MS/MS runs were largely due to those proteins identified with only single

(43) Wolters, D. A.; Washburn, M. P.; Yates, J. R., *Anal. Chem.* **2001**, *73*, 5683−5690.

peptides, indicating that these ORFs were most likely to be among the lowest abundance peptides detected. Table 2 lists results from two separate runs with 0.25-ng samples. Both runs identified seven proteins, but two from each run were different, i.e., DR2377 and DRA0283 in run 1 and DR0440 and DR1245 in run 2. These proteins were assigned based upon a single peptide. In contrast, all other proteins detected in multiple runs also gave either multiple peptides or single peptides having high X_{Corr} values (>3.0). For nanoLC-FTICR MS, two runs using a 2.5-ng sample yielded 5% deviations in the number of identified proteins and >70% overlap between the sets of proteins identified. The non-common proteins were similarly also identified from only single peptides and also had low intensities. Unsurprisingly, the lowest level proteins detected (and represented by detecting a single low-intensity peptide peak) showed the greatest run-to-run variability in both the MS and MS/MS approaches.

Proteome Quantitation. Both relative[30,44−46] and absolute[47,48] LC−MS protein quantitation depend on LC elution profile measurements that can be influenced by many factors, including the number of spectra obtained for the eluting peak, details of the ESI source operation, and MS analyzer and detector characteristics, as well as (under some conditions) the nature of the coeluting species and their concentrations. Here the quantification of peptide/protein relative abundances is based upon the relative intensities of stable-isotope-labeled (e.g., $^{14}N/^{15}N$) peptide pairs having essentially identical chemical properties. However, careful examination of peptide LC elution profiles shows that there are small differences in chromatographic retentions of the paired peptides, and improvements in the quality of separations aggravate this problem. Figure 8 shows the difference in elution profiles (over four spectra) of a $^{14}N/^{15}N$-labeled peptide pair. Typically the ^{15}N-labeled peptide elutes earlier, but the difference varied from peptide to peptide. Summing the total intensities obtained from all four spectra yields a light/heavy relative abundance ratio of 1.13, while the spectra individually give ratios that vary from 0.22 to 1.88. Such shifts were also observed for solid-phase isotope-coded affinity tag labeled peptide pairs (data not shown), consistent with the report of Zhang and Regnier.[49] The single-stage MS data acquisition approach allows continuous and simultaneous data acquisition for more accurate quantitation of multiple peptides eluting from the LC column, while the MS/MS approach suffers from the discontinuous nature of detection due to shifting between MS and MS/MS analysis modes.

CONCLUSIONS

We have described the basis for ultrasensitive nanoscale proteomics using on-line high-efficiency microSPE-(15-μm-i.d. column) nanoLC nanoESI-MS (2-μm-i.d emitter) and MS/MS. Using FTICR MS, we were able to detect six BSA tryptic peptides when a 75-zmol BSA tryptic digest sample was loaded with 10 μL of solution, leading to an estimated peptide mass detection limit of 10 zmol (i.e., ∼6000 molecules) and a peptide concentration detection limit of ∼250 aM. This sensitivity applied to a much more complex proteomic mixture enabled identification of high-abundance proteins from as little as 0.5 pg of total sample which, given appropriate sample processing methods, would potentially allow proteome measurements covering a $\sim 10^3$ range of relative peptide abundances from a single mammalian cell. Combined with the mass and elution time data analysis method, the nanoLC-FTICR MS provided a protein identification throughput of ∼870 proteins/run (∼290 proteins/h). The ∼3-h proteome measurements provided coverage of 14% of the predicted proteins from 2.5 to 10 ng of tryptic digest from *D. radiodurans* sample. Using the same approach with an ion trap-MS/MS analysis, the sensitivity was reduced to ∼7 amol for confident protein identification, and 50 pg of the sample was required to identify the most abundant proteins (∼100-fold greater than needed with FTICR). Precise quantitation using stable isotope labeling is also effective for studies with these small samples. The reduced sensitivity and increased sample size requirement with ion trap-MS/MS analyses led to reduced proteome measurement throughput, coverage, and dynamic range. Finally, we note that the incorporation of a microSPE with very small i.d. LC columns allowed robust operation and effective manipulation of much smaller samples than previously feasible and facilitated achieving extremely low concentration detection limits.

ACKNOWLEDGMENT

We thank the U.S. Department of Energy's Office of Biological and Environmental Research and the National Cancer Institute (Grant CA 86340) and the NIH National Center for Research Resources (RR 18522) for supporting portions of this research. We also thank Stanley Sterns of Valco Instruments for providing injection valves for 10 000 psi operation. Pacific Northwest National Laboratory is operated by the Battelle Memorial Institute for the U.S. Department of Energy through Contract DE-ACO6-76RLO 1830.

Received for review March 11, 2003. Accepted October 8, 2003.

AC030096Q

(44) Oda, Y.; Huang, K.; Cross, F. R.; Cowburn, D.; Chait, B. T. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6591−6596.
(45) Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. *Nat. Biotechnol.* **1999**, *17*, 994−999.
(46) Mirgorodskaya, O. A.; Kozmin, Y. P.; Titov, M. I.; Körner, R.; Sönksen, C. P.; Roepstorff, P. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1226−1232.
(47) Gerber, S. A. Rush, J.; Stemmann, O.; Steen, H.; Kirschner, M. W.; Gygi, S. P. *50th ASMS Conference on Mass Spectrometry and Allied Topics*, 2002; p 1538.
(48) Wang, W.; Shaler, T. A.; Norton, S. M.; Hill, L. R.; Lin, H.; Zhou, H.; Roy, S. M.; Hastings, C. A.; Becker, C. H. *50th ASMS Conference on Mass Spectrometry and Allied Topics*, 1620, 2002; 1620.
(49) Zhang, R.; Regnier, F. E. *J. Proteome Res.* **2002**, *1*, 139−147.