Nano-LC FTICR Tandem Mass Spectrometry for Top-Down Proteomics: Routine Baseline Unit Mass Resolution of Whole Cell Lysate Proteins up to 72 kDa

ARTICLE *in* ANALYTICAL CHEMISTRY · MARCH 2012

Impact Factor: 5.64 · DOI: 10.1021/ac202651v · Source: PubMed

10 AUTHORS, INCLUDING:

Jeremiah D Tipton
SCIEX
21 PUBLICATIONS **573** CITATIONS
SEE PROFILE

John C Tran
Genentech
22 PUBLICATIONS **848** CITATIONS
SEE PROFILE

Dorothy R Ahlf Wheatcraft
The Jackson Laboratory
13 PUBLICATIONS **576** CITATIONS
SEE PROFILE

Published in final edited form as:
Anal Chem. 2012 March 6; 84(5): 2111–2117. doi:10.1021/ac202651v.

Nano-LC FT-ICR Tandem Mass Spectrometry for Top-Down Proteomics: Routine Baseline Unit Mass Resolution of Whole Cell Lysate Proteins up to 72 kDa

Jeremiah D. Tipton[†,‡,§,⊥,¶], **John C. Tran**[§], **Adam D. Catherman**[§], **Dorothy R. Ahlf**[§], **Kenneth R. Durbin**[§], **Ji Eun Lee**[§,⊥], **John F. Kellie**[§], **Neil L. Kelleher**[§], **Christopher L. Hendrickson**[†,¶], and **Alan G. Marshall**[*,†,¶]

[†]Ion Cyclotron Resonance Program, National High Magnetic Field Laboratory, Florida State University, 1800 East Paul Dirac Drive, Tallahassee, FL, 32310-4005, U.S.A.

[§]Department of Molecular Biosciences, Northwestern University, 2205 Tech Drive, Evanston, IL 60208-3500

[¶]Department of Chemistry and Biochemistry, Florida State University, 95 Chieftain Way, Tallahassee, FL, 32306-4390

Abstract

Current high-throughput top-down proteomic platforms provide routine identification of proteins less than 25 kDa with 4-D separations. This short communication reports the application of technological developments over the last few years that improve protein identification and characterization for masses greater than 25 kDa. Advances in separation science has allowed increased numbers of proteins to be identified, especially by nano-liquid chromatography (nLC) prior to mass spectrometry (MS) analysis. Further, a goal of high-throughput top-down proteomics is to extend the mass range for routine nLC MS analysis up to 80 kDa because gene sequence analysis predicts that about ~70% of the human proteome is transcribed to be less than 80 kDa. Normally, large proteins greater than 50 kDa are identified and characterized by top-down proteomics through fraction collection and direct infusion at relatively low throughput. Further, other MS based techniques provide top-down protein characterization, however at low resolution for intact mass measurement. Here, we present analysis of standard (up to 78 kDa) and whole cell lysate proteins by Fourier transform ion cyclotron resonance mass spectrometry (nLC ESI FT-ICR MS). The separation platform reduced the complexity of the protein matrix so that at 14.5 Tesla, proteins from whole cell lysate up to 72 kDa are baseline mass resolved on a *nano-LC chromatographic time scale*. Further, the results document routine identification of proteins at improved throughput based on accurate mass measurement (less than 10 ppm mass error) of precursor and fragment ions for proteins up to 50 kDa.

[*]**Corresponding Author** A.G.M. marshall@magnet.fsu.edu. Phone: 850-644-0529 Fax: 850-644-1366.
[‡]Current Address: Florida Center for Drug Discovery and Innovation, University of South Florida, 3702 Spectrum Blvd, Suite 175, Tampa, FL 33612-9444
[⊥]Current Address: Doping Control Center and Center for Theragnosis, Korea Institute of Science and Technology, Seoul, South Korea

ASSOCIATED CONTENT

Supporting Information. Additional information as noted in text. This material is available free of charge via the Internet at http://pubs.acs.org.

Keywords

Fourier transform mass spectrometry; ion cyclotron resonance; FTMS; FT-ICR; top-down proteomics

INTRODUCTION

Top-down proteomics comprises of the direct mass spectrometric (MS) isolation of intact protein(s) followed by fragmentation (MSn) for protein identification and amino acid site location of post-translational modifications.[1–3] Bottom-up proteomics relies on proteolytic digestion of complex protein mixtures, pre-fractionation of peptides, and nano-liquid chromatography (nLC) MS/MS.[4–6] Although the bottom-up approach can achieve extensive proteome coverage, sequence coverage of the identified proteins is typically incomplete, often lacking specific isoform information.[7] Top-down proteomics can potentially provide 100% sequence coverage by accurate precursor mass measurement and efficient dissociation and measurement of resulting fragment ions.[2, 3, 8] The strength of top-down proteomics is the ability to characterize protein truncations, isoforms, combinatorial posttranslational modifications, and even detect errors in annotated genome databases.[7, 9–15]

However, top-down *characterization* of proteins (*above* 25 kDa in mass) is often limited to a targeted approach, i.e. an off-line purified sample is directly infused into the mass spectrometer. Prior to recent developments in protein separation, platforms that automated the fraction collection / direct infusion process improved throughput and number of identified proteins.[14, 16–19] Recent advances in solution-based electrophoresis and nLC separations,[20–22] instrumentation,[23–29] and informatics[28, 30] have enabled an increase in number of identified proteins. Furthermore, an increased number of reports indicate that high-throughput characterization of proteins *below* 25 kDa can be routinely achieved by nano-LC LTQ/FT-ICR MS.[21, 24, 31–34] Recently, high-throughput top-down proteomics with a 4-D separation platform identified over 1000 proteins from whole cell lysate, thus increasing the capacity by 20× for human while also extending the upper molecular weight limit per "discovery" type experiment.[35] Further, the number of proteins was increased by 5× for the yeast proteome.[36] In those studies, there were fragmentation evidence for 2–3× as many species existed for every gene product identification. In the human proteome study, protein fractions above 25 kDa were measured by low-resolution precursor mass measurement, followed by in-source dissociation. This strategy resulted in identification of many protein species above 25 kDa, but without precursor ion measurement. Thus, a strategy for accurate mass measurement of proteins *above* 25 kDa remains a goal for top-down proteomics. The main reason for the need for high resolution measurement is for PTM and isoform characterization: i.e., protein species often co-elute and abundance may be directly measured for a particular system.

Increased throughput for top-down proteomics is also related to the chosen fragmentation method. Top-down protein characterization was originally implemented with electron capture dissociation (ECD).[37–39] With ECD, a protein ion is dissociated into multiple fragment ions of lower abundance. Thus, co-addition (or signal averaging) is typically required to improve the signal-to-noise ratio for *de novo* sequencing of resulting fragments. Recent introduction of electron transfer dissociation (ETD) [40–42] enabled ECD-type fragmentation in a quadrupole ion trap. ETD fragmentation with the LTQ may improve throughput and fragmentation efficiency as advanced instrumentation continues to develop. In contrast to ECD or ETD, "thermal" activation [i.e., collision-induced dissociation (CID), infrared multiphoton dissociation (IRMPD), source-induced dissociation, etc.] heats an ion to a higher Boltzmann temperature and breaks the weakest bond (the peptide C-N

linkage).[43] For example, IRMPD[44–46] dissociates a precursor peptide ion at fewer sites, and thus typically provides less sequence information than ECD.[3, 8, 47] Although CID provides less sequence coverage, throughput and sensitivity are increased because dissociation typically produces fragment ions that have populations required for single measurement events. Thus, the main difference between ECD and thermal activation (with regard to high-throughput analysis with on-line nLC) is rapid fragmentation into fewer modes, thus becoming the method of choice for fragmentation to increase the number of identified proteins by 20×. [35] "Hybrid" instruments generate fragment ions in a quadrupole ion trap through collision-induced dissociation, source-induced dissociation, or high-energy collisional dissociation (HCD) prior to transfer to a high resolution mass analyzer.[14, 17, 24, 27, 48–51] In this report, we apply source-induced dissociation to identify whole cell lysate proteins between 25 and 50 kDa in molecular weight.

The progression of high-throughput top-down proteomics to higher molecular weight is limited by several sources of signal reduction.[3, 52] Even a relatively small protein (e.g., carbonic anhydrase, $C_{1474}H_{2286}N_{412}O_{439}S_{14}$) contains heavy isotopes (^{13}C, ^{15}N, ^{34}S) that generate a broad isotopic distribution for a given charge state. However, the complex isotopic distribution may be "de-isotoped" to provide accurate mass for a large protein.[53] Also, electrospray ionization (ESI),[54, 55] the method of choice for coupling nLC with a mass spectrometer, imparts multiple charges, creating a charge state distribution (each with its own isotopic distribution) for a protein.[56] Moreover, a higher molecular weight protein typically exhibits higher (and more) charge states, further reducing the observed mass spectral signal magnitude.[52] Finally, the mass spectrum of a large electrosprayed protein may be further broadened by substitution of alkali metal ions (Na^+, K^+) for one or more of the protons, and by non-specific binding of solvent or small molecule ligands. In fact, the solvation state of a large protein is suggested to be one of the main limitations for consistent fragmentation by ECD, CID, or nozzle-skimmer dissociation.[57, 58]

Despite the challenges associated with analyzing large polypeptides, we report a simple top-down platform [35] that can provide accurate precursor measurement of proteins (derived from whole cell lysate) up to 72 kDa. Results include rms mass errors of less than 10 ppm for identified proteins. The implementation of the previously reported 4-D separation platform with the addition of FTICR “zoom mapping” precursor detection allow us to resolve intact protein isoforms and posttranslational modifications up to 72 kDa.

EXPERIMENTAL PROCEDURES

M-phase-arrested HeLa Cell Preparation, Solution Isoelectric Focusing, and Gel-Eluted Liquid Fraction Entrapment Electrophoresis

M-phase-arrested HeLa S3 cells were prepared by treating asynchronous cells (3×10^5 cells/mL) with 1 μM colchicine for 18 h. Cells (5×10^8 cell/mL) were harvested, lysed, and split into cytosolic and nuclear fractions as previously described.[21] Isolated protein pellets were acetone-precipitated after lysis and dissolved (3 mg each) in 3.2 mL solution isoelectric focusing (sIEF) buffer (8 M urea, 2 M thiourea, 50 mm DTT, 1% w/v Biolyte 3/10 carrier ampholytes (Bio-Rad Laboratories, Hercules, CA)), and focused by use of a custom designed eightchamber sIEF system as previously described.[59] Acetone-precipitated sIEF fractions were resolved in denaturant buffer and separated by gel-eluted liquid fraction entrapment electrophoresis (GF) as previously described.[20, 60] The GF separation was visualized by silver stained SDS-PAGE (1/15 loading amount of GF fractions corresponding to low μg) to provide sample representative gel images.

Standard Protein Analysis by Automated Pre-fractionation and Direct Infusion

Pre-fractionation of standard proteins was performed with a Triversa Nanomate (Advion Biosciences, Ithaca, NY), followed by automated direct infusion of fractions collected in 96 well plates.[14, 19, 33] Specifically, characterization of transferrin included separation and fraction collection of 5 pmol with a reversed-phase C4 column (GraceVydac, Deerfield, IL): 30 min gradient delivered at 50 μL/min. The Triversa Robot was trained to collect 30 s fractions, followed by direct infusion of the fractions for 10–15 min. Similar experiments were performed with protein samples from GF only or sIEF-GF (fractions visualized to be above 25 kDa). Results from that experiment indicated that sample lifetime deteriorates rapidly due to exposure to air and light, and non-optimal solvents.

Nano-Liquid Chromatography of sIEF/GF Fractions

Nano-LC was performed with a 1-D plus Nano-LC system (Eksigent, Dublin, CA), delivering a gradient (400 nL/min) of 5% B to 80% B in 30 – 50 min (Solvent A: 94.8% H_2O, 5% acetonitrile, 0.2% formic acid, Solvent B: 4.8% H2O, 94% acetonitrile, 0.2% formic acid). Samples were loaded onto a 30 mm × 150 μm pre-column (1 μL/min) and separated over a 100 mm × 75 μm i.d. pico-Frit column with a 15 μm i.d. tip. The columns were packed with PLRP-S (5 μm, 1000 Å) material (New Objective, Woburn, MA).

Fraction Collection, Direct Infusion ESI FT-ICR Mass Spectrometry

Mass spectrometry was performed with a custom-built 14.5 T LTQ FT-ICR mass spectrometer (LTQ-FT, Thermo Fisher Corp., Bremen, Germany).[23, 25] Direct infusion (automatic timed infusion by the Triversa Nanomate) MS and MS/MS analysis were performed by ESI 14.5 T FT-ICR MS (1.8 s transient duration). Two modes of FT-ICR acquisition were used during the 10–15 min automatic ESI timeframe. First, broadband detection was signal averaged for 3 to 5 min. Second, MS/MS by CID activation was performed for 7 to 10 min. Further, the instrument is also equipped with a wired octopole[61], thus increasing the capacity for collection of ions prior to transfer to the ICR cell.[25] For precursor ion measurement, one million charges were accumulated in the LTQ prior to transfer to the wired octopole; 10 cycles were accumulated in the wired octopole prior to transfer to the ICR cell. Further, for CID MS/MS, 0.25 million charges were activated for dissociation in the LTQ prior to 15 cycles and transfer to the ICR cell. Data was compiled and analyzed with custom cRAWler software.[62]

On-line nanoLC MS Thermal Activation with Source-Induced Dissociation

Protein fractions (above 25 kDa) were analyzed by nLC (i.e. analytical column, 75 μm i.d.) "zoom-mapping" (5 × 106 charges with 4 microscans) source-induced dissociation (6×10^6 charges with 6 microscans) Simply, "zoom-mapping" is a sequential "walking window" in which ions spanning an *m/z* range of 80 to 100 are chosen based on the direct infusion experiments (i.e., highest m/z envelope signal magnitude). Based on the direct infusion experiments, the sequential scheme was optimized to be: measurement window 1 is *m/z* 890 to 950, measurement window 2 is *m/z* 940 to 1000, followed by broad mass range source-induced dissociation. The sequence was repeated during the duration of nLC MS period. Data was extracted with custom cRAWler software[62] and searched against differently annotated databases with ProSight PC 2.0.[30] A two-tier search tree included a 200 Da absolute mass search followed by a 50,000 Da (delta mass turned on, fragment ions with a search tolerance of 10 ppm. The cutoff score for accepting a protein as a true identification was an E-value of 1×10^{-4}. The feature map for visualization of the nLC MS data was generated with Proteome Display.[62]

RESULTS AND DISCUSSION

The continued development of high-throughput top-down proteomics increases the number of isoforms and post-translational modifications (PTMs) characterized from a single biological system during a single experiment. A recent example of this improvement was reported for human cell lines: a 20-fold increase in the total number of identified proteins over any prior report.[35] Accurate precursor ion measurement precedes the characterization of increased numbers of isoforms and PTM's: i.e., an intact protein and its phosphorylated form are measured in the same spectrum, differing by 80 Da. If higher molecular weight proteins are accessible with high resolving power and high mass accuracy, multiple posttranslationally modified protein species may be resolved and identified simultaneously. A specific example of the need for high mass accuracy includes proteins that have been deamidated (a mass increase of 1 Da)

The typical experiment for proteins above 25 kDa includes pre-fractionation with reversed-phase chromatography followed by direct infusion mass spectrometry.[14, 19, 33] Initial experiments for whole cell lysate proteins (above 25 kDa in mass) were performed by that method. The constant infusion experiments resulted in the definition of two "zoom-mapping" windows for nLC MS. Transferrin, a commercially available 78 kDa protein (18 disulfide bonds, 1 site for phosphorylation, and 1 site for glycosylation) was fractionated off-line and directly infused. Signal averaging for 5 min at mass resolving power $m/\Delta m_{50\%}$ = 400,000 at *m/z* 800 revealed the full baseline unit mass-resolved isotopic distribution (Figure 1a). Heterogeneous superposition of phosphorylation/glycosylation and oxidation modifications generates unresolved "fine structure" for each resolved peak, in addition to fine structure caused by overlap of heavy isotopes (e.g., ^{13}C vs. ^{15}N). Despite spectral complexity, dissociation of transferrin by LTQ CID and FT-ICR fragment ion measurement resulted in a sequence tag that was identified by ProSight 2.0 (Figure 1b). Transferrin represents a class of proteins currently difficult to analyze by top-down methods—namely, heterogeneous species with multiple sites for glycosylation, phosphorylation, and salt adducts. For characterization of recombinant biologics, the fraction collection/direct infusion strategy is still very useful.

The transition of the direct infusion experiment to nLC FT MS for transferrin (~1.2 pmol loaded on column; ~20 s elution time) is shown in Figure 1c. In contrast to the full direct infusion experiment (10 min for MS and MS/MS data collection), the nLC elution time for less material was only 20 s. On-line nLC MS obviously allows for multiple components to be analyzed, depending on the peak capacity of the chromatography material. Further, non-reduced transferrins were analyzed by nLC. Non-reduced transferrin resulted in the identification of transferrin with a ProSightPC sequence tag score of 3×10^{-10}.

Protein extracts were derived from mitotically arrested HeLa S3 cells by cell lysis, fractionated into nuclear and cytosolic, and separated by isoelectric focusing and GF. Initial experiments for whole cell lysate proteins (above 25 kDa in mass) were performed by the same fraction collection/direct infusion method described above. Although the "targeted" fractionation collection/direct infusion approach for characterization of proteins is typical for proteins above 30 kDa, that method is not considered high throughput. With improved throughput, an increased number of protein forms may be characterized and linked to phenotypes. Also, many of the proteins that were directly infused resulted in broad isotopic distributions due to oxidation, salt, and other solvent adducts. When the same types of samples were analyzed by nLC, broad isotopic distributions due to adducts were typically not observed. For intact protein analysis, samples deteriorate with longer exposure to air and MS compatible solvents. Finally, the constant infusion experiments resulted in the definition

of two "zoom-mapping" (see experimental section for definition) windows for nLC MS.[63, 64]

Fractions that were visualized to fall between 25 and 50 kDa (representative gel image, Figure S1) resulted in 32 confidently identified proteins. In all, 25 fractions greater than 25 kDa were collected from GF or sIEF-GF. Generally, for each sIEF fraction, 5 fractions between 25 and 50 were generated by GF fractionation with a first generation commercial device. The number of proteins that were identified is apparently low; due the dynamic range problem associated with all whole cell lysate proteomic studies. In the 1000 protein report, improved GF separations were achieved with a custom device and the use of low resolution mass spectrometry to detect intact masses greater than 25 kDa. In this study, many of the abundant proteins, such as GAPDH, enolase, and actin were found across multiple fractions: e.g., the "housekeeping" protein, GAPDH, appeared in 8 fractions. Further, highly abundant large proteins (as is also the case for gel based bottom up proteomics) are known to "smear" throughout a gel. The continued improvements to selection and separations will improve top down proteomics in the same way as for bottom-up proteomics. The limit in the number of identified proteins is due to highly abundant proteins in the same molecular weight fractions across sIEF fractions. With regard to reversed-phase nLC, the peak width for proteins less than 25 kDa is typically between 10 and 60 s, whereas the peak width for proteins greater than 35 kDa is typically 1–2 min. Improvements in the separation resins and load/peak capacity have reduced peak widths for proteins up to 25 kDa and recently up to 50 kDa.[35] If the next generation of materials provides improved peak resolution, improved analytics will be observed for the number of proteins above 25 kDa and their protein species and PTM relatives.

In terms of the instrumentation and platform as a whole, the identified proteins from whole cell lysate were baseline unit mass-resolved by on-line nLC (Figures 2b and c) and confidently characterized with mass errors of less than 10 ppm (Table 1). Mass measurements differing by more than 10 ppm could result from *in vivo* modifications, such as deamidation, assignment algorithm error (e.g., by 1 Da shift in fitting a given isotopic distribution), or other adducts (greater than 1 Da mass shift). A feature map presents many baseline resolved proteins above 10:1 S/N ratio, based on signal averaging over 20–40 s of the total ion chromatogram (Figure 2a).[62] Many proteins up to 48 kDa were identified from a single time-domain acquisition. For example, baseline unit mass resolution for keratin 18 (48,272 and 48,216 Da) (Figure 2b) demonstrates the platform's ability to chromatographically separate isoforms with highly similar sequence identity. Source-induced dissociation of those isoforms generated 55 fragments from the N- and C-termini, and a ProSightPC expectation value of 1.61×10^{-64} (i.e., essentially certain identification).[65]

As mentioned earlier in the discussion, bottom-up proteomic analysis of M-phase arrested HeLa S3 cells has been shown to be up-regulated with phosphoproteins.[66] Figure 2c shows identification of phosphoglycerate kinase 1 and a second peak 80 Da higher in mass. The result is based on accurate measurement of precursor and fragment ions. The result is further corroborated by other M-phase arrested HeLa S3 cell studies.[66] The result is significant because many phosphoproteins are known to be between 30 and 50 kDa in mass. Thus, as selection techniques improve along with on-line separations, current instrumentation may be sufficient to characterize protein form complexity.

In this study, the important metric is the accurate precursor mass measurement of larger proteins on a nLC time scale. The duty cycle is typically 33–60 ms for each collection and transfer of 5 to 10 million charges to the ICR. Each data-dependent MS – MS/MS cycle requires ~10 s; the overall duty cycle for the experiment is limited by the FT-ICR data

acquisition period (0.7 s) required to produce a mass resolving power, $m/\Delta m_{50\%}$ = 100,000 at *m/z* 800, to baseline unit mass-resolve proteins up to 60 kDa. For a given mass resolving power, the detection period varies inversely with magnetic field strength.[67] Thus, at 21 Tesla, it should be possible to achieve the same resolving power at an increased duty cycle (shorter transient acquisition period) for nLC MS. Further, at higher magnetic field, an increased number of ions may be analyzed.[68] For proof of concept at 14.5 tesla, we were able to obtain unit mass resolution from a cell-derived 72 kDa protein (Figure 3). At a given magnetic field strength, the sensitivity, mass accuracy, and duty cycle of new ICR MS cells further improve protein identification.[29, 69, 70]

In summary, we have presented several examples of routine (and reliable) identification of proteins up to 50 kDa by 2D solution electrophoresis nLC MS top-down proteomics, with 72 kDa as the current upper limit for baseline unit mass-resolved MS at a nano-LC time scale. Standard proteins up to 78 kDa (~1 pmol loaded on column) were baseline unit mass resolved identified with nLC MS, thereby improving the likelihood of accurate mass detection of proteins derived from cells. Recent encouraging improvements in resolving power for samples analyzed off-line include 0.01 Da mass resolution for bovine serum albumin (66.5 kDa)[69] and 0.25 Da mass accuracy for a monoclonal antibody (148 kDa).[70] The relatively small number of proteins listed in Table 1 is limited mainly by the separation efficiency for large biopolymers (greater than 30 kDa in mass) by the overabundance of "housekeeping" proteins. Moreover, the number of MS-resolved precursor ions is much greater than the number of identified proteins, limited by the current fragmentation efficiency. However, the barriers for analysis of large intact proteins should continue to fall with improvements to separation methods, control of ionization, and instrumentation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We gratefully acknowledge support from the NSF Division of Materials Research through DMR-06-54118, the State of Florida, and NIH GM-067193-08. We also thank Leonid Zamdborg and Paul Thomas for helpful discussions.

REFERENCES

1. Kelleher NL, Lin HY, Valaskovic GA, Aaserud DJ, Fridriksson EK, McLafferty FW. Journal of the American Chemical Society. 1999; 121:806–812.
2. Kelleher NL. Analytical Chemistry. 2004; 76:197A–203A. [PubMed: 14697051]
3. McLafferty FW, Breuker K, Jin M, Han X, Infusini G, Jiang H, Kong X, Begley TP. FEBS Journal. 2007; 274:6256–6268. [PubMed: 18021240]
4. Link AJ, Eng J, Schieltz DM, Carmack E, Mize GJ, Morris DR, Garvik BM, Yates JR. Nat Biotech. 1999; 17:676–682.
5. Washburn MP, Wolters D, Yates JR. Nature Biotechnology. 2001; 19:242–247.
6. Wisniewski JR, Zougman A, Nagaraj N, Mann M. Nat Meth. 2009; 6:359–362.
7. Tipton JD, Tran JC, Catherman AD, Ahlf DR, Durbin KR, Kelleher NL. Journal of Biological Chemistry. 2011; 286:25451–25458. [PubMed: 21632550]
8. Sze SK, Ge Y, Oh H, McLafferty FW. Proceedings of the National Academy of Sciences of the United States of America. 2002; 99:1774–1779. [PubMed: 11842225]
9. Pesavento JJ, Kim Y, Taylor GK, Kelleher NL. Journal of the American Chemical Society. 2004; 126:2286–3387. [PubMed: 14982415]
10. Ge Y, Rybakov IN, Xu Q, Moss RL. PNAS. 2009; 106:12658–12663. [PubMed: 19541641]

11. Siuti N, Kelleher NL. Nature Methods. 2007; 4:817–821. [PubMed: 17901871]
12. Ayaz-Guner S, Zhang J, Li L, Walker JW, Ge Y. Biochemistry. 2009; 48:8161–8170. [PubMed: 19637843]
13. Whitelegge J. TrAC Trends in Analytical Chemistry. 24:576–582.
14. Patrie SM, Ferguson JT, Robinson DE, Whipple D, Rother M, Metcalf WW, Kelleher NL. Molecular and Cellular Proteomics. 2006; 5:14–25. [PubMed: 16236702]
15. Ferguson JT, Wenger CD, Metcalf WW, Kelleher NL. Journal of the American Society for Mass Spectrometry. 2009; 20:1743–1750. [PubMed: 19577935]
16. Johnson JR, Meng F, Forbes AJ, Cargile BJ, Kelleher NL. Electrophoresis. 2002; 23:3217–3223. [PubMed: 12298093]
17. Du Y, Meng F, Patrie SM, Miller LM, Kelleher NL. Journal of Proteome Research. 2004; 3:801–806. [PubMed: 15359734]
18. Patrie SM, Robinson DE, Meng F, Kelleher NL. International Journal of Mass Spectrometry. 2005; 234:175–184.
19. Wenger CD, Boyne MT II, Ferguson JT, Robinson DE, Kelleher NL. Analytical Chemistry. 2008; 80:8055–8063. [PubMed: 18841935]
20. Tran JC, Doucette AA. Analytical Chemistry. 2008; 80:1568–1573. [PubMed: 18229945]
21. Lee JE, Kellie JF, Tran JC, Tipton JD, Catherman AD, Thomas HM, Ahlf DR, Durbin KR, Vellaichamy A, Ntai I, Marshall AG, Kelleher NL. Journal of the American Society for Mass Spectrometry. 2009; 20:2183–2191. [PubMed: 19747844]
22. Vellaichamy A, Tran JC, Catherman AD, Lee LE, Kellie JF, Sweet SMM, Zamdborg L, Thomas PM, D.R. A, Durbin KR, Valaskovic GA, Kelleher NL. Analytical Chemistry. 2010; 82:1234–1244. [PubMed: 20073486]
23. Syka JEP, Marto JA, Bai DL, Horning S, Senko MW, Schwartz JC, Ueberheide B, Garcia B, Busby SA, Muratore T, Shabanowitz J, Hunt DF. J. Proteome Res. 2004; 3:621. [PubMed: 15253445]
24. Parks BA, Jiang L, Thomas PM, Wenger CD, Roth MJ, Boyne MT II, Burke PV, Kwast KE, Kelleher NL. Analytical Chemistry. 2007; 79:7984–7991. [PubMed: 17915963]
25. Schaub TM, Hendrickson CL, Horning S, Quinn JP, Senko MW, Marshall AG. Analytical Chemistry. 2008; 80:3985–3990. [PubMed: 18465882]
26. Makarov A, Denisov E, Kholomeev A, Balschun W, Lange O, Strupat K, Horning S. Analytical Chemistry. 2006; 78:2113–2120. [PubMed: 16579588]
27. Karabacak NM, Li L, Tiwari A, Hayward LJ, Hong P, Easterling ML, Agar JN. Molecular and Cellular Proteomics. 2009; 8:846–856. [PubMed: 19074999]
28. Cobb JS, Easterling ML, Agar JN. Journal of the American Society for Mass Spectrometry. 2010; 21:949–959. [PubMed: 20303285]
29. Tolmachev AV, Robinson EW, Wu S, Paša-Tolic L, Smith RD. International Journal of Mass Spectrometry. 2009; 287:32–38. [PubMed: 20473360]
30. Zamdborg L, LeDuc RD, Glowacz KJ, Kim YB, Viswanathan V, Spaulding IT, Early BP, Bluhm EJ, Babai S, Kelleher NL. Nucleic Acids Research. 2007; 35:W701–W706. [PubMed: 17586823]
31. Lee SW, Berger SJ, Martinović S, Pasa-Tolić L, Anderson GA, Shen Y, Zhao R, Smith RD. PNAS. 2002; 99:5942–5947. [PubMed: 11983894]
32. Zabrouskov V, Senko MW, Du Y, Leduc RD, Kelleher NL. Journal of the American Society for Mass Spectrometry. 2005; 16:2027–2038. [PubMed: 16253516]
33. Roth MJ, Parks BA, Ferguson JT, Boyne MT II, Kelleher NL. Analytical Chemistry. 2008; 80:2857–2866. [PubMed: 18351787]
34. Wynne C, Edwards NJ, Fenselau C. PROTEOMICS. 2010; 10:3631–3643. [PubMed: 20845332]
35. Tran JC, Zamdborg L, Ahlf DR, Lee JE, Catherman AD, Durbin KR, Tipton JD, Vellaichamy A, Kellie JF, Li M, Wu C, Sweet SMM, Early BP, Siuti N, LeDuc RD, Compton PD, Thomas PM, Kelleher NL. Nature. 2011; 480:254–258. [PubMed: 22037311]
36. Kellie JF, Catherman AD, Durbin KR, Tran JC, Tipton JD, Norris JL, Witkowski CE, Thomas PM, Kelleher NL. Analytical Chemistry. 2011

37. Zubarev RA, Kelleher NL, McLafferty FW. Journal of the American Chemical Society. 1998; 120:3265–3266.
38. Horn DM, Ge Y, McLafferty FW. Analytical Chemistry. 2000; 72:4778–4784. [PubMed: 11055690]
39. Cooper HJ, Håkansson K, Marshall AG. Mass Spectrometry Reviews. 2005; 24:201–222. [PubMed: 15389856]
40. Reid GE, McLuckey SA. Journal of Mass Spectrometry. 2002; 37:663–675. [PubMed: 12124999]
41. Coon JJ, Ueberheide B, Syka JEP, Dryhurst DD, Ausio J, Shabanowitz J, Hunt DF. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102:9463–9468. [PubMed: 15983376]
42. Udeshi ND, Compton PD, Shabanowitz J, Hunt DF, Rose KL. Nature Protocols. 2009; 3:1709–1717.
43. McLuckey SA, Goeringer DE. J. Mass Spectrom. 1997; 35:461–474.
44. Woodlin RL, Bomse DS, Beauchamp JL. J. Am. Chem. Soc. 1978; 100:3248–3250.
45. Little DP, Speir JP, Senko MW, O'Connor PB, McLafferty FW. Analytical Chemistry. 1994; 66:2809–2815. [PubMed: 7526742]
46. Li W, Hendrickson CL, Emmett MR, Marshall AG. Analytical Chemistry. 1999; 71:4397–4402. [PubMed: 10517152]
47. Zabrouskov V, Whitelegge JP. Journal of Proteome Research. 2007; 6:2205–2210. [PubMed: 17441748]
48. Chanthamontri C, Jian Liu J, McLuckey SA. International Journal of Mass Spectrometry. 2009; 283:9–16. [PubMed: 20160958]
49. Armirotti A, Benatti U, Damonte G. Rapid Communications in Mass Spectrometry. 2009; 23:661–666. [PubMed: 19184982]
50. Macek B, Waanders LF, Olsen JV, Mann M. Molecular and Cellular Proteomics. 2006; 5:949–958. [PubMed: 16478717]
51. Zhai H, Han X, Breuker K, McLafferty FW. Analytical Chemistry. 2005; 77:5777–5784. [PubMed: 16159106]
52. Compton PD, Zamdborg L, Thomas PM, Kelleher NL. Analytical Chemistry. 2011; 83:6868–6874. [PubMed: 21744800]
53. Horn DM, Zubarev RA, McLafferty FW. Journal of the American Society for Mass Spectrometry. 2000; 11:320–332. [PubMed: 10757168]
54. Fenn J, Mann M, Meng C, Wong S, Whitehouse C. Science. 1989; 246:64–71. [PubMed: 2675315]
55. Loo JA, Udseth HR, Smith RD. Analytical Biochemistry. 1989; 179:404–412. [PubMed: 2774189]
56. Hendrickson CL, Emmett MR. Annual Review of Physical Chemistry. 1999; 50:517–536.
57. Han X, Jin M, Breuker K, McLafferty FW. Science. 2006; 314:109. [PubMed: 17023655]
58. Breuker K, McLafferty FW. Proceedings of the National Academy of Sciences. 2008; 105:18145–18152.
59. Tran JC, Doucette AA. Journal of Proteomic Research. 2008; 7:1761–1765.
60. Tran JC, Doucette AA. Analytical Chemistry. 2009; 81:6201–6209. [PubMed: 19572727]
61. Wilcox BE, Hendrickson CL, Marshall AG. Journal of the American Society for Mass Spectrometry. 2002; 13:1304–1312. [PubMed: 12443021]
62. Durbin KR, Tran JC, Zamdborg L, Sweet SMM, Catherman AD, Lee JE, Li M, Kellie JF, Kelleher NL. Proteomics. 2010; 10:3589–3597. [PubMed: 20848673]
63. Tipton, JD.; Kellie, JF.; Thomas, PM.; D, A.; Hendrickson, CL.; Kelleher, NL.; Marshall, AG. 57th ASMS Conf. on Mass Spectrometry & Allied Topics, Philadelphia, PA, May 31 – June 4; 2009.
64. Tipton, JD.; Lee, JE.; Catherman, AC.; Durbin, KR.; Hendrickson, CL.; Emmett, MR.; Marshall, AG.; Kelleher, NL. 58th Amer. Soc. for Mass Spectrom. Annual Conf. on Mass Spectrometry & Allied Topics, Salt Lake City, UT, May 23–27; 2010.
65. LeDuc RD, Taylor GK, Kim YB, Januszyk TE, Bynum LH, Sola JV, Garavelli JS, Kelleher NL. Nucleic Acids Research. 2004; 32:W340–W345. [PubMed: 15215407]

66. Olsen JV, Vermeulen M, Santamaria A, Kumar C, Miller ML, Jensen LJ, Gnad F, Cox J, Jensen TS, Nigg EA, Brunak S, Mann M. Sci. Signal. 2010; 3 ra3-.

67. Marshall AG, Hendrickson CL, Jackson GS. Mass Spectrom. Rev. 1998; 17:1–35. [PubMed: 9768511]

68. Marshall AG, Hendrickson CL. International Journal of Mass Spectrometry. 2002; 215:59–75.

69. Nikolaev EN, Boldin IA, Jertz R, Baykut GJ. J. Am. Soc. Mass Spectrom. 2011; 22:1125–1133. [PubMed: 21953094]

70. Valeja SG, Kaiser NK, Xian F, Hendrickson CL, Rouse JC, Marshall AG. Analytical Chemistry. 2011; 83:8391–8395. [PubMed: 22011246]

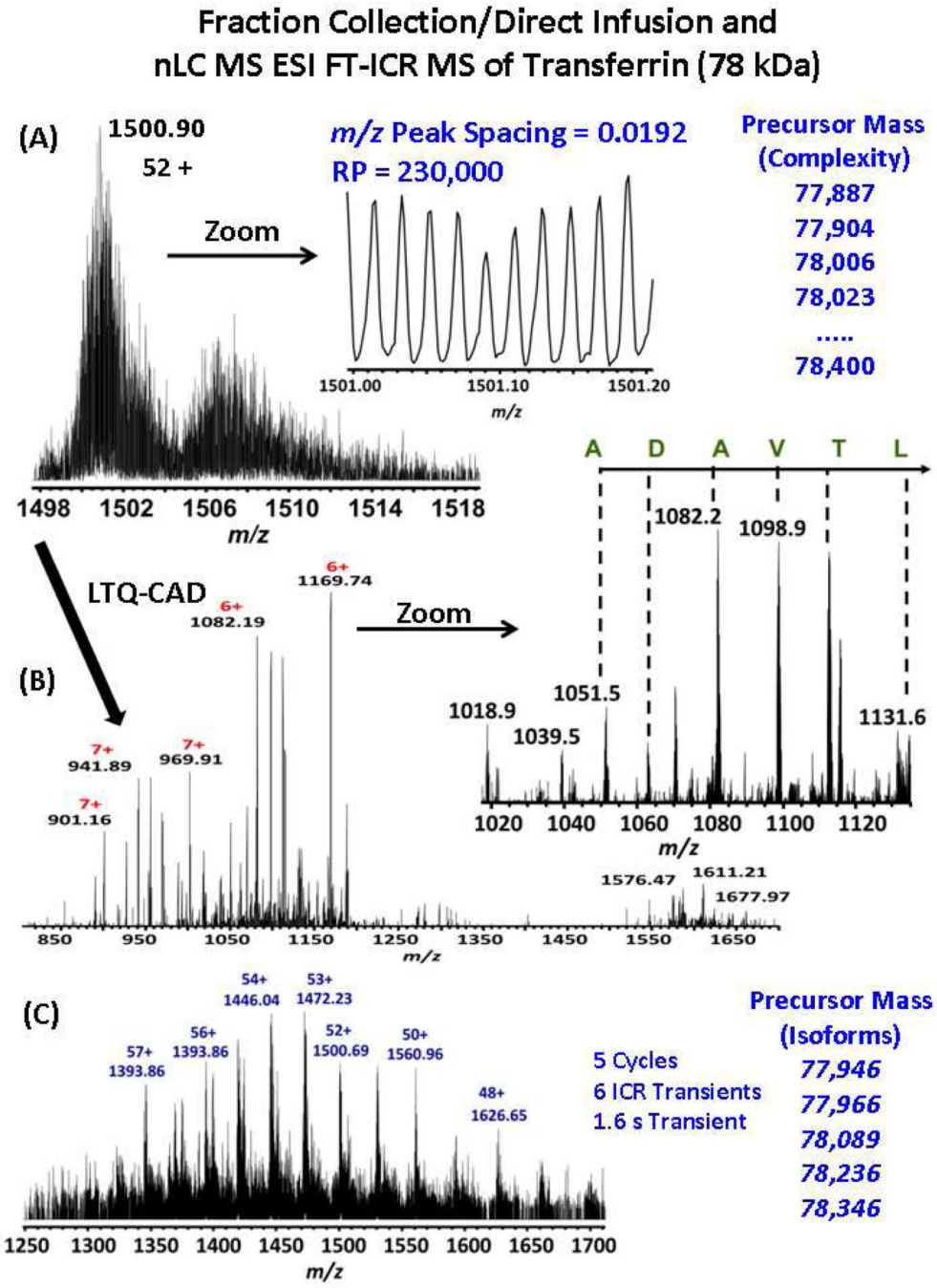

Figure 1.

(A) Baseline unit mass-resolved ESI 14.5 T FT-ICR mass spectrum of bovine transferrin after LC fraction collection and direct infusion. (B) Signal averaged CID product ion mass spectrum, providing a sequence tag for confident identification. (C) Baseline unit mass-resolved nLC ESI FT-ICR precursor ion mass spectrum from ~1.2 pmol of transferrin loaded onto a 75 μm i.d. nano-LC (see text). Each "cycle" represents a repeated LTQ ion accumulation and transfer to an external octopole before injection into the ICR cell.

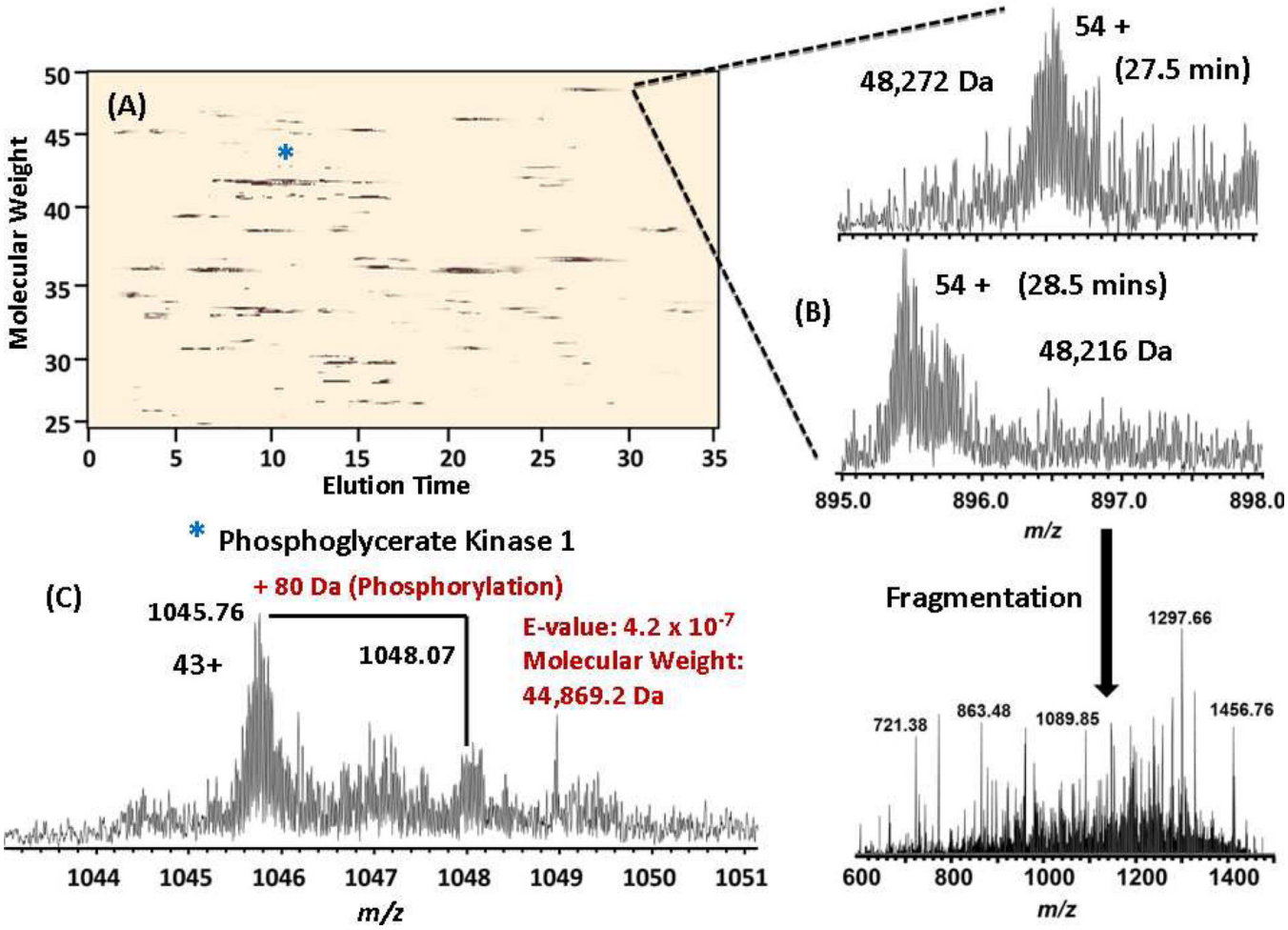

Figure 2.

(A) Selected mass chromatograms for proteins baseline unit mass-resolved by nano-LC ESI 14.5 T FT-ICR MS. (B) Chromatographic separation and baseline unit mass-resolved mass spectrum for keratin 18 (54+ charge state). Proteins identified by activation of the protein in the LTQ are listed in Table 1. (C) Identification of phosphoglycerate kinase 1 and an isoform 80 Da. higher in mass (corresponding to a single phosphorylation). The blue star indicates its location in (A).

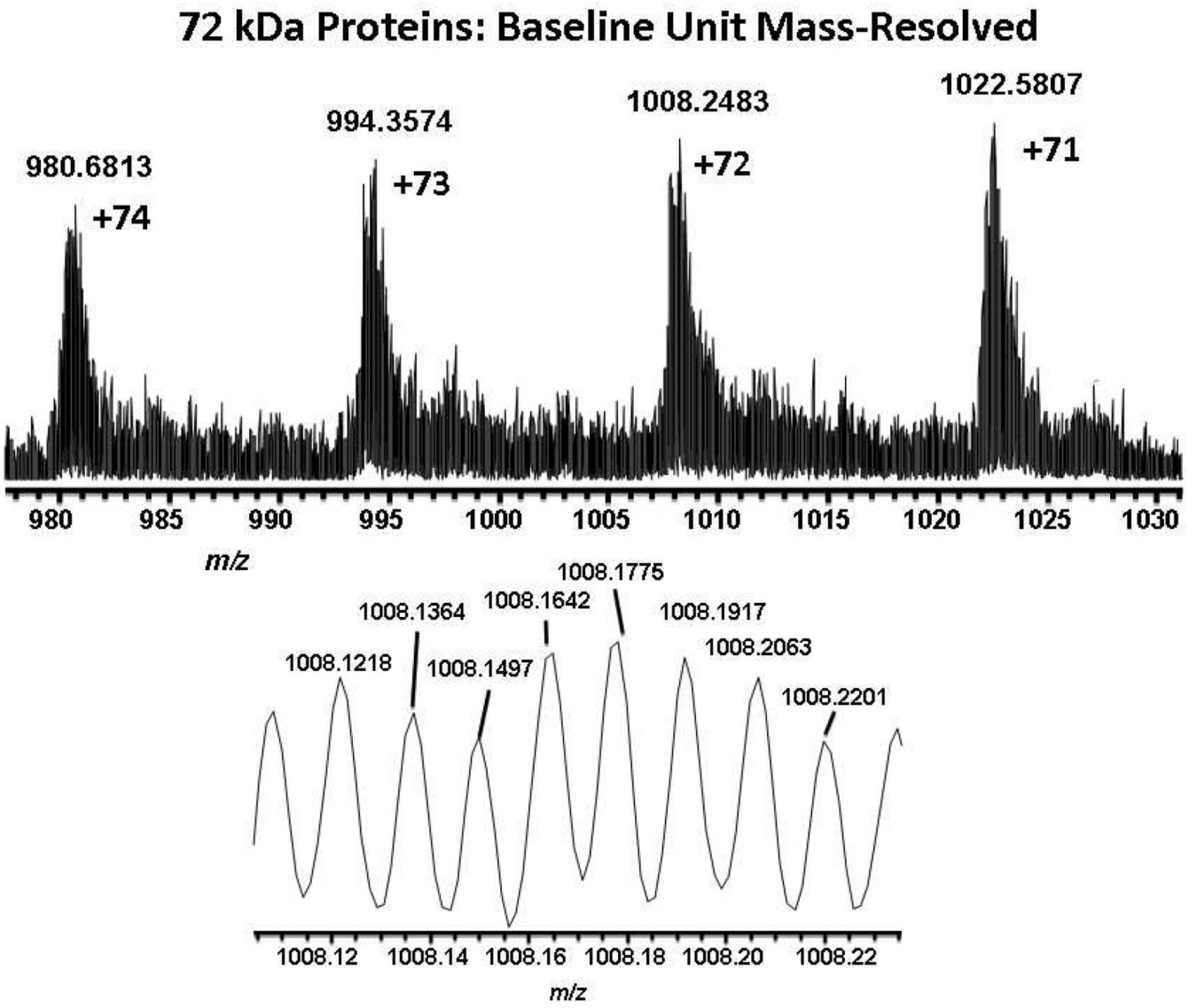

Figure 3.
Cell-derived proteins baseline unit mass-resolved by nLC ion-marching/nozzle-skimmer experiment. Mass resolving power = 400,000 (1.6 s transient) at m/z 400.

Table 1

Results from the analysis of 25 fractions from isoelectric focusing/gel electrophoresis fractions.

Gene	Accession	Fragments	E-value	Calc'd Mass	Exptl Mass	mass	ppm error	Protein Description
K1C17	Q04695	47	1.6E-64	48272.10	48271.02	−1.08	−22.33	Keratin, type I cytoskeletal 17
K1C18	P05783	14	4.4E-17	47939.50	47938.62	−0.88	−18.30	Keratin, type I cytoskeletal 18
ENOA	P06733	10	8.0E-06	47391.40	47391.49	0.09	1.90	Alpha-enolase
PGK1	A8K4W6	8	4.2E-07	44896.20	44896.20	0.01	0.20	Phosphoglycerate kinase
ACTB	Q1KLZ0	27	1.5E-39	41962.89	41962.81	−0.09	−2.14	Actin
ALDOA	P04075	13	1.5E-07	39720.40	39720.40	0.03	0.77	Fructose-bisphosphate aldolase A
ANXA2	P07355	7	8.8E-04	38718.90	38718.90	0.01	0.22	Annexin A2
LDHB	Q5U077	32	8.9E-52	36811.20	36811.10	−0.10	−2.71	L-lactate dehydrogenase
MDHC	P40925	18	2.0E-28	36542.10	36542.11	0.01	0.19	Malate dehydrogenase, cytoplasmic
ALDR	P15121	8	8.7E-05	36140.60	36140.52	−0.08	−2.27	Aldose reductase
ROA2	P22626	32	8.3E-36	36082.90	36082.83	−0.07	−1.94	Heterogeneous nuclear ribonucleoproteins A2/B1
G3P	P04406	42	3.5E-37	36070.40	36070.30	−0.05	−1.40	Glyceraldehyde-3-phosphate dehydrogenase
SFXN1	Q9H9B4	10	5.7E-07	35792.40	35792.51	0.11	3.07	Sideroflexin-1
GBLP	P63244	20	5.4E-37	35421.70	35421.63	−0.07	−2.05	Guanine nucleotide-binding protein subunit beta-2-like 1
RM44	Q9H9J2	16	3.0E-11	34639.10	34639.15	0.05	1.44	39S ribosomal protein L44, mitochondrial
RLA0L	Q8NHW5	11	6.1E-14	34382.70	34570.40	*187.7***	*5429.44*	60S acidic ribosomal protein P0-like
ROA1	P09651	17	3.3E-24	34200.31	34281.11	*80.1**	*5467.89*	Heterogeneous nuclear ribonucleoprotein A1
MTCH2	Q9Y6C9	5	5.1E-05	33847.10	33846.94	−0.16	−4.68	Mitochondrial carrier homolog 2
PHB2	Q99623	9	1.5E-04	33186.90	33186.85	−0.05	−1.51	D-prohibitin
IPYR	Q15181	18	7.1E-25	33006.30	33006.31	0.01	0.36	Inorganic pyrophosphatase
CAZA1	P52907	8	1.5E-05	32984.40	32984.45	0.05	1.52	F-actin-capping protein subunit alpha-1
NPM	P06748	21	4.3E-13	32725.90	32904.86	*178.96***	*5438.66*	Nucleophosmin
DECR	Q16698	10	4.8E-09	32485.80	32486.95	*1.15*##	*35.40*	2,4-dienoyl-CoA reductase
VDAC2	P45880	5	3.1E-05	31970.70	31970.69	−0.01	−0.31	Voltage-dependent anion-selective channel protein 2
PHB	Q6PUJ7	19	6.8E-12	29769.90	29752.83	−17.07	−573.58	Prohibitin
EXOS6	Q5RKV6	10	1.6E-08	28371.71	28371.67	−0.04	−1.41	mRNA transport regulator 3 homolog
HDGF	Q5SZ07	19	2.2E-25	28172.90	28140.37	*−32.53*	*492.58*	Hepatoma-derived growth factor
ETFB	P38117	10	6.8E-08	27965.20	27981.26	*16.06*	*573.96*	Electron transfer flavoprotein subunit beta

Gene	Accession	Fragments	E-value	Calc'd Mass	Exptl Mass	mass	ppm error	Protein Description
HCD2	Q99714	30	1.4E-36	27045.20	27045.26	0.06	2.22	3-hydroxyacyl-CoA dehydrogenase type-2
TPIS	P60174	21	2.6E-20	26806.80	26806.90	−0.01	−0.37	Triosephosphate isomerase
RTN3	O95197	10	1.4E-17	25674.60	25675.82	*1.22*	*47.52*	Neuroendocrine-specific protein-like 2
PRDX6	P30041	10	5.0E-05	25002.20	25002.22	0.02	0.77	Peroxiredoxin-6

** indicates that these mass errors could not be attributed to specific adducts or polymorphism.

corresponds to a mass error of 1 Da that can be attributed to peak extraction errors.

* corresponds to identification of a phosphoprotein pair.