Published in final edited form as:
Mol Biosyst. 2010 September ; 6(9): 1532–1539. doi:10.1039/c000896f.

The emerging process of Top Down mass spectrometry for protein analysis: biomarkers, protein-therapeutics, and achieving high throughput†

John F. Kellie, **John C. Tran**, **Ji Eun Lee**, **Dorothy R. Ahlf**, **Haylee M. Thomas**, **Ioanna Ntai**, **Adam D. Catherman**, **Kenneth R. Durbin**, **Leonid Zamdborg**, **Adaikkalam Vellaichamy**, **Paul M. Thomas**, and **Neil L. Kelleher**
Technology Development Team, Center for Top Down Proteomics, University of Illinois at Urbana-Champaign, USA

Neil L. Kelleher: kelleher@scs.uiuc.edu

Abstract

Top Down mass spectrometry (MS) has emerged as an alternative to common Bottom Up strategies for protein analysis. In the Top Down approach, intact proteins are fragmented directly in the mass spectrometer to achieve both protein identification and characterization, even capturing information on combinatorial post-translational modifications. Just in the past two years, Top Down MS has seen incremental advances in instrumentation and dedicated software, and has also experienced a major boost from refined separations of whole proteins in complex mixtures that have both high recovery and reproducibility. Combined with steadily advancing commercial MS instrumentation and data processing, a high-throughput workflow covering intact proteins and polypeptides up to 70 kDa is directly visible in the near future.

Introduction

Top Down mass spectrometry differs from the more traditional Bottom Up strategy of protein analysis by starting with intact mass measurement (Fig. 1). If one is interested in just detection of a protein's presence—or its simple identification from a database—then Bottom Up tends to be easier, with a wide variety of well-honed tools resulting from the lion's share of efforts from vendors and academic laboratories. However, if one is interested in characterization of an entire protein's primary structure (including amino acid sequence and modifications), a targeted strategy using Top Down is more efficient in many cases. Today, the hardware and software combinations are more available than ever to acquire Top Down data. Top Down proteomics—where hundreds of intact proteins are fragmented directly in a mass spectrometer (no protease)—is becoming more feasible every year. Prior to 2007 only a handful of Top Down studies reported greater than 25 protein identifications; however, in recent years a dozen or so studies have been reported, some observing a few hundreds or even thousands of proteins.[1–5]

Currently, many labs are using aspects of Top Down MS for measurement of endogenous proteins like histones and biomarkers under 30 kDa or analysis of protein-based therapeutics undergoing the process of drug development.[6,7] For example, Amgen and Amylin have used

†This article is part of the 2010 *Molecular BioSystems* 'Emerging Investigators' issue: highlighting the work of outstanding young scientists at the chemical- and systems-biology interfaces.

Correspondence to: Neil L. Kelleher, kelleher@scs.uiuc.edu.

Top Down for characterization of recombinant antibodies and endogenous secretory peptides, respectively.[8–10] In a targeted mode (Fig. 1B), Top Down analysis of single proteins containing multiple post-translational modifications (PTMs) has shown clear advantages. Top Down concepts are taking hold for efficient characterization of protein pharmacophores where routine characterization of deamidations, synthetic modifications on Cys or Lys residues, and simple glycosylation patterns are prevalent.[11,12]As generic protein-based drugs emerge and drop in cost, the characterization of large molecules with the analytical rigor now routine for small molecules could be a stronghold for Top Down concepts and ultimately replace the concept of "biosimilarity" as the *de facto* metric of product quality and safety.[13] With such expanding efforts to produce high quality biopharmaceuticals, formulators and process engineers often change conditions or production hosts from prokaryotic to higher organisms such as yeast, plant, insect or mammalian cells that harbor the machinery for complex splice forms and PTMs. So, due to the benefits of Top Down MS, we envision a steady shift from classical peptide mapping *via* Bottom Up MS to the use of 20–100 kDa peptides and intact proteins for precise isoform and PTM analysis of such biopharmaceuticals with ~100% sequence coverage. Currently, measurements of intact antibodies are a key part of a new characterization workflow incorporating elements of both Top Down and Bottom Up.[14] While it is unclear if direct MS/MS of intact antibodies will replace the approach of using PGNaseF and protein digestion, direct MS/MS analysis of heavy and light chains is currently feasible.[10]

One example of the deep insight provided by Top Down is in the area of histone biology.[15] Beyond such boutique applications that tend to be low-throughput, we introduced a high-throughput version of Top Down that could soon become a reasonable option for many labs wanting to identify and characterize hundreds of proteins per day.[3] For a systems approach, Top Down analysis can facilitate understanding of protein-level complexity by dissecting isoform relative abundances and profiling endogenous arrays of PTMs without use of isotopic labeling for quantification.[6,16] Such measurements in Top Down are possible with LC-MS/MS because multiple protein forms can elute simultaneously allowing precursor and fragment ion relative ratios to be measured. With these motivations, we outline here some very recent developments for processing intact proteins in ways amenable for large scale work. Selected options for each of the three pillars of proteome analysis are highlighted here to give the reader a sense for the most recent developments in this growing sub-field of protein analysis.

Separation science for intact proteins

Given the enormous complexity of proteomic mixtures ranging from sub-organellar complexes to whole-cell proteomes, MS alone is not sufficient in adequately characterizing a proteome. Effective separations are critical for decreasing sample complexity and increasing the dynamic range of detection. While some front end separations for MS-based proteomics are performed at the intact protein level, the eventual method of analysis—Top Down or Bottom Up—plays a key role in determining what front end separations are to be used. For example, the most dominant separation platform for intact proteins is two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) due to its unrivalled peak capacity; however, extraction of intact proteins from gels results in low recovery. Bottom Up proteomics after 2D-PAGE is possible, thanks to enzymes (such as trypsin) which literally cut proteins from the solid phase. Agilent's OFFGEL isoelectric focusing system is capable of separating intact proteins, but is most commonly used for peptide separations due to the relatively low sample recovery of intact proteins from the device.[17–19] Therefore, intact protein separations prior to Top Down MS are generally conducted in a fashion that produces fractionated proteins in solution. Separations are based on a wide range of protein-intrinsic parameters, which include charge, size, and hydrophobicity.

Solution-phase separations based on protein charge have historically been dominated by ion exchange chromatography partly because of widespread familiarity and relatively high loading capacity. While popular, this method does not yield high resolution separation nor does it separate proteins predictably according to isoelectric point. Despite this, combining ion exchange chromatography with reversed-phase liquid chromatography and Top Down mass spectrometry, we have shown that 133 protein forms could be identified from human white blood cells.[20] Alternatively, chromatofocusing (a derivative of ion exchange chromatography) and solution isoelectric focusing, which both utilize pH gradients, are capable of separating proteins with high isoelectric point correlation and are promising as alternatives to salt gradient ion exchange chromatography for Top Down proteomic separations.[21]

A custom-designed solution isoelectric focusing (sIEF) device featuring effective focusing for protein separation shows promise for Top Down.[22] In contrast to other devices based on carrier ampholytes such as the Rotofor or Free Flow Electrophoresis, this method enables near-quantitative recovery of precipitated proteins in addition to rapid focusing. The eight fractions from sIEF have recently been coupled to a multiplex tube gel electrophoresis separation device termed gel-eluted liquid fraction entrapment electrophoresis (GELFrEE), which is in our view the most promising new development for robust Top Down. GELFrEE shows highly predictable molecular weight separations at high resolution and with >90% recovery.[23,24]

The impact of coupling sIEF to the GELFrEE device creates a highly orthogonal separation, with proteins available in the solution phase at high yield. Akin to 2D gels, this separation combines IEF with the ability to separate proteins according to molecular weight. Using GELFrEE alone, Tran and Doucette were able to achieve rapid partitioning of a proteome into about 20 well-resolved, discrete mass range fractions of complex proteomic mixtures in as fast as one hour, over mass ranges from below 10 kDa to 250 kDa. A recent advancement in Top Down has been the connection of this size-based separation approach with micro-capillary reversed-phase liquid chromatography coupled to tandem mass spectrometry (μRPLC-MS/MS), demonstrating high-throughput analysis of intact proteins. This new workflow is highlighted in Fig. 2. This GELFrEE-LC-MS/MS platform routinely identifies 20–60 proteins in a one-hour LC-MS/MS run.[3,25] Example data are shown in Fig. 3 using ~100–200 μg sample amounts. Also, direct LC-MS analysis of protein standards (ranging from 8 to 45 kDa) shows clear detection of amounts as low as 0.3 pmol.[25]

The GELFrEE device also can be applied to Bottom Up proteomics, peptide analysis, and large protein purification. A recent GELFrEE manuscript described in detail a method to separate intact proteins and then digest proteins in each fraction after separation.[24] Many endogenous peptides are of interest for biomarker characterization, such as highly processed neuropeptides and protein hormones in complex biological fluids.[9,26,27] We predict that the GELFrEE-μRPLC-MS/MS approach will find wide application in these areas. For >50 kDa proteins, shifting from typical pore sizes (15% T) to gels with high pore sizes (~5% T) affords fast elution of proteins as a means of purification. Such large proteins can then be detected by Western blotting and subjected to Top Down and/or Bottom Up MS.

Mass spectrometry hardware for Top Down proteomics

There are four basic components for a typical mass spectrometer: sample inlet, ion source, mass analyzer, and detector. Usually, the mass analyzer defines the instrument type. Ion Trap, Time-of-Flight (TOF), Orbitrap, and Fourier Transform Ion Cyclotron Resonance (FTICR) mass analyzers are commonly used in Top Down MS, and these mass analyzers vary in one or more fundamental performance metrics (Table 1).

Ion traps are commonly used for Bottom Up proteomic analyses, but have been used for Top Down as well.[28,29] High sensitivity afforded by ion traps means that a large number of protein fragment ions can be detected, but the instruments have lower mass accuracy and unit resolution. For Top Down use, an ion trap as a sole mass analyzer is becoming more feasible.[30] However, the low mass accuracy of the instrument results in lower confidence identifications than identifications from instruments with high mass accuracy.[28] The true utility for ion traps in Top Down proteomics lies in a number of hybrid instruments (LTQ-Orbitrap, LTQ-FT).

Recently, a TOF analyzer has been used for Top Down analysis of intact proteins between 10 and 30 kDa.[31] Deterding and co-workers showed that a Top Down strategy using a q-TOF was complementary to Top Down FTICR MS and Bottom Up for characterizing oxidatively-damaged myoglobin.[32] Also a hybrid quadrupole-TOF MS capable of unit resolution of fragment ions was used for Top Down MS/MS on proteins approaching 50 kDa.[8]

FTICR and Orbitrap mass analyzers both rely on Fourier transformation of image currents recorded at the detector. FTICR is based on excitation of ions under the influence of a high magnetic field,[33] and the Orbitrap analyzer uses electrostatic trapping between a cylindrical center and outer barrel-like electrode.[34] Much of the early Top Down work was achieved using FTICR MS and associated tightly with formalization of the Top Down strategy of protein characterization in 1999.[35] Meng *et al.* demonstrated large scale Top Down protein identifications from *Saccharomyces cerevisiae*,[36] and Pesavento *et al.* observed histone H4 modification during the cell cycle using only Top Down.[37] These are just a few examples of the success of Top Down using custom instrumentation; however, if Top Down is to appeal to a wide range of researchers as a mainstream method for proteomics, commercial instrumentation is necessary. Recently, high-throughput Top Down proteomics has been demonstrated for the low molecular weight proteome.[3] This was achieved using specially-built Thermo LTQ-FTICR instruments. Also, a hybrid Top Down/Bottom Up approach has shown promise using a Bruker q-FTICR MS.[38] Thermo extended the LTQ-FTICR hybridization concept to the recently-released Orbitrap,[39] creating the LTQ-Orbitrap. Top Down was first achieved on this instrument by Macek and co-workers in 2006,[40] and many Top Down studies have followed including SILAC quantitation, lipidomics, and antibody analysis.[10,41,42] We believe the use of ion traps, TOFs, and (benchtop) Orbitraps, along with the classic FTICR MS approach, will all have major roles to play in writing the next chapter of Top Down (Fig. 2, lower right).

Methods of ion fragmentation

Collision-induced dissociation (CID) is often implemented in lower resolution ion traps or quadrupoles with fragmentation products able to be analyzed at high resolution in FT mass analyzers. Electron capture dissociation (ECD) occurs in the superconducting magnet portion of FTICR mass analyzers and has proven to be a robust method for extensive characterization of whole proteins in direct infusion analysis.[43] More recently, ETD implements an electron transfer reaction in an ion trap, with fragmentation readout in an Orbitrap mass analyzer (Table 1).[44] Newer modes of ETD introduce the reagents from the ion source of the mass spectrometer.[45,46]

Indeed, many labs are engaging in Top Down, both developing new methods and demonstrating high throughput. Liu and Schey recently displayed Top Down identification using CID fragmentation on a MALDI TOF-TOF instrument.[31] CID has also proved effective for Top Down using a q-TOF.[47] Incorporating front end separations, Ouvry-Patat and co-workers have displayed utility of a free-flow electrophoresis coupled to FTICR

MS.[48] New methods will enable more users to engage in Top Down, especially for laboratories that cannot justify the cost of a high-field FTICR system. ETD has also been utilized for Top Down in high-throughput mode, where strong anion exchange prior to LC-MS/MS analysis identified 322 protein forms corresponding to 174 unique protein species in about two days.[1] The use of electron-based MS/MS continues to mature, with hybrid MS/MS approaches (*e.g.*, ECD with activation) likely to play a major role for the fine mapping of modifications and sequence changes; however, for high-throughput use, CID will maintain a central role in Top Down for the next few years.

Computational proteomics for Top Down MS and MS/MS

Bioinformatics for MS-based proteomics is expanding rapidly. As Top Down progresses, analysis software is becoming more readily available, a sign that researchers acknowledge the need for tailored Top Down software. Three current Top Down search engines are MascotTD (Big Mascot), PIITA, and ProSight. Each uses different methodologies but recognize a key element: Top Down requires special attention in data processing that is different from Bottom Up.

Mascot is widely used for Bottom Up analyses and has been thoroughly described.[49,50] A restriction for Top Down analyses is the lack of Mascot support over 16 kDa, but in the recently-released MascotTD, support is extended to 110 kDa.[51] This new mode requires assignment of both precursor and fragment masses; however, if no intact precursor is determined, a “default” intact mass can be applied to permit database searching over a very wide mass range. MascotTD is able to identify multiple forms of a single protein, including PTMs and sequence variants.

Another engine for Top Down data processing is known as the Precursor Ion Independent Top-down Algorithm (PIITA).[4] PIITA uses fragmentation data to identify a protein from a predicted gene. Once a gene match is made, the intact mass is used to map observed shifts from the gene-predicted mass. In this model, a “relaxed” search has no *a priori* expectation of either PTM mass or total PTMs present. PIITA has been used to identify 154 unique proteins from *Salmonella typhimurium*.

The first and most flexible search engine for Top Down MS is ProSight. There are three primary search modes that can be used within the ProSight software suite for protein identification: absolute mass, biomarker, and sequence tag (Fig. 4).[52] Absolute mass searches involve matching the observed mass to a theoretical intact mass stored in a database (within a user-specified tolerance) and then comparing observed fragment masses to those calculated from possible forms. Biomarker searching involves matching an observed mass to the theoretical masses of possible subsequences in the database and then comparing calculated fragments of those subsequences to observed fragments. Combinations of PTMs derived from known sites of modification are stored in the database, enabling simultaneous identification and characterization of intact proteins. Sequence tag searching identifies a series of fragment ions which match amino acid additions and attempts to match the tag to sequences in the database.

Proteome visualization software for Top Down proteomics

For comparative proteomics studies, there exists a need to “visualize” the proteome. However, unless 2D gels are used, an ultimate visualization solution is neither simple nor straight-forward. For example, if more than one front end separation is involved, the researcher must choose which is displayed. Methods highlighted here display protein mass as a function of some front end separation metric. The Lubman group began displaying detected protein masses as a function of isoelectric focusing fraction (p*I*) in 2002.[53] They

extended this methodology to chromatofocusing in 2006 and displayed differences in proteins detected using LC-TOF MS between estrogen-treated and untreated premalignant, estrogen-responsive human breast cells.[54]Also, Wu and co-workers implemented a hybrid Top Down/Bottom Up strategy and plotted detected protein mass as a function of LC retention time.[55]

Using our newly developed LC-MS visualization technologies coupled to exact-mass protein identification, it is possible to monitor differences in protein and PTM (and sequence variant) expression between two different cell states. Mass spectral intensity for each protein form is correlated to a map of molecular weight as a function of LC elution time. In Fig. 5, LC-MS visualizations for injections of GELFrEE fraction from asynchronous and M-Phase arrested HeLa are shown correlated to MS detection. Relative changes in a single protein can be observed between two samples or cell states. Detection of individual proteins builds each map for different cell states and Top Down offers a novel impact on biological studies with a “bird’s eye view” of complex PTM dynamics on the forms of proteins present, providing insight into disparity of protein expression between two samples.

Summary and outlook

Top Down mass spectrometry has progressed substantially in the last few years and now provides answers for intact protein *characterization* that Bottom Up cannot—particularly when multiple mass shifts due to sequence differences and modifications are present over a stretch of sequence >20 residues. Targeting abundant proteins in cells, endogenous biomarkers, and protein pharmaceuticals has shifted from possibility to reality on plug-and-play commercial instruments.

Refined and even automated front end sample separations/preparations will further facilitate the use of Top Down. Varied and tailored separation schemes will almost certainly emerge to improve the analysis of very large proteins (approaching 100 kDa) and membrane proteins. This coupled with instrumentation progress such as increased sensitivity, resolution, and scan speed will gradually increase the high-throughput molecular weight limit of Top Down. Until the point when an entire proteome can be analyzed quickly the lack of throughput will ultimately deter some users; however, hybrid Top Down/Bottom Up or Middle Down (digestion of the proteome into large peptides) strategies can fill the high mass “void” and be used to obtain useful information from the entire proteome.

In the next decade, optimized front end separations, sophisticated yet affordable bench top instrumentation, and refined data analysis/visualization will solidify Top Down as both practical above 50 kDa and viable in a high-throughput environment. Within a few years, Top Down will emerge as a method of choice for those wishing to definitively interrogate proteins at the biochemical level in the context of protein therapeutics, targeted questions in biomedical/clinical research, and even discovery-type systems biology. When robust Bottom Up protein analysis became high-throughput, it was adopted by many laboratories—no matter the level of difficulty. Alternative strategies now must be compared in terms of cost-effectiveness, accessibility to “non-experts”, and ability to produce high-value and definitive data. Using several of these metrics, Top Down mass spectrometry can now compete with Bottom Up in some contexts. Through achieving implementation of Top Down on a truly proteomic scale, more people in diverse laboratories will start to believe that a new world of protein analysis lies on the horizon.

Acknowledgments

The authors thank the Packard Foundation, the Roy J. Carver Charitable Trust, the University of Illinois, the National Institutes of Health (GM 067193), and the National Science Foundation (DMS 0800631) which have

supported the development of Top Down mass spectrometry and Top Down proteomics. Additionally, we thank Michael Senko, Fred McLafferty, Alan Marshall and the many former members of the Kelleher Research Team that have all contributed to the development of tandem mass spectrometry above 10 kDa.

References

1. Bunger MK, Cargile BJ, Ngunjiri A, Bundy JL, Stephenson JL. Anal. Chem. 2008; 80:1459–1467. [PubMed: 18229893]
2. Collier TS, Hawkridge AM, Georgianna DR, Payne GA, Muddiman DC. Anal. Chem. 2008; 80:4994–5001. [PubMed: 18512951]
3. Lee JE, Kellie JF, Tran JC, Tipton JD, Catherman AD, Thomas HM, Ahlf DR, Durbin KR, Vellaichamy A, Ntai I, Marshall AG, Kelleher NL. J. Am. Soc. Mass Spectrom. 2009; 20:2183–2191. [PubMed: 19747844]
4. Tsai YS, Scherl A, Shaw JL, MacKay CL, Shaffer SA, Langridge-Smith PRR, Goodlett DR. J. Am. Soc. Mass Spectrom. 2009; 20:2154–2166. [PubMed: 19773183]
5. Zhou F, Hanson TE, Johnston MV. Anal. Chem. 2007; 79:7145–7153. [PubMed: 17711353]
6. Pesavento JJ, Mizzen CA, Kelleher NL. Anal. Chem. 2006; 78:4271–4280. [PubMed: 16808433]
7. Ayaz-Guner S, Zhang J, Li L, Walker JW, Ge Y. Biochemistry. 2009; 48:8161–8170. [PubMed: 19637843]
8. Ren D, Pipes GD, Hambly D, Bondarenko PV, Treuheit MJ, Gadgil HS. Anal. Biochem. 2009; 384:42–48. [PubMed: 18834850]
9. Taylor SW, Andon NL, Bilakovics JM, Lowe C, Hanley MR, Pittner R, Ghosh SS. J. Proteome Res. 2006; 5:1776–1784. [PubMed: 16823986]
10. Zhang ZQ, Shah B. Anal. Chem. 2007; 79:5723–5729. [PubMed: 17591752]
11. Zhang ZQ, Pan H, Chen XY. Mass Spectrom. Rev. 2009; 28:147–176. [PubMed: 18720354]
12. Yu XC, Borisov OV, Alvarez M, Michels DA, Wang YJ, Ling V. Anal. Chem. 2009; 81:9282–9290. [PubMed: 19852494]
13. Kresse GB. Eur. J. Pharm. Biopharm. 2009; 72:479–486. [PubMed: 19286455]
14. Pristatsky P, Cohen SL, Krantz D, Acevedo J, Ionescu R, Vlasak J. Anal. Chem. 2009; 81:6148–6155. [PubMed: 19591437]
15. Siuti N, Kelleher NL. Nat. Methods. 2007; 4:817–821. [PubMed: 17901871]
16. Zabrouskov V, Ge Y, Schwartz J, Walker JW. Mol. Cell. Proteomics. 2008; 7:1838–1849. [PubMed: 18445579]
17. de Godoy LMF, Olsen JV, Cox J, Nielsen ML, Hubner NC, Frohlich F, Walther TC, Mann M. Nature. 2008; 455:1251–1254. [PubMed: 18820680]
18. Heller M, Michel PE, Morier P, Crettaz D, Wenz C, Tissot JD, Reymond F, Rossier JS. Electrophoresis. 2005; 26:1174–1188. [PubMed: 15704250]
19. Michel PE, Reymond F, Arnaud IL, Josserand J, Girault HH, Rossier JS. Electrophoresis. 2003; 24:3–11. [PubMed: 12652567]
20. Roth MJ, Parks BA, Ferguson JT, Boyne MT, Kelleher NL. Anal. Chem. 2008; 80:2857–2866. [PubMed: 18351787]
21. Lubman DM, Kachman MT, Wang HX, Gong SY, Yan F, Hamler RL, O'Neil KA, Zhu K, Buchanan NS, Barder TJ. J. Chromatogr. B. 2002; 782:183–196.
22. Tran JC, Doucette AA. J. Proteome Res. 2008; 7:1761–1766. [PubMed: 18284188]
23. Tran JC, Doucette AA. Anal. Chem. 2008; 80:1568–1573. [PubMed: 18229945]
24. Tran JC, Doucette AA. Anal. Chem. 2009; 81:6201–6209. [PubMed: 19572727]
25. Vellaichamy A, Tran JC, Catherman AD, Lee JE, Kellie JF, Sweet SMM, Zamdborg L, Thomas PM, Ahlf DR, Durbin KR, Valaskovic GA, Kelleher NL. Anal. Chem. 2010; 82:1234–1244. [PubMed: 20073486]
26. Kultima K, Nilsson A, Scholz B, Rossbach UL, Falth M, Andren PE. Mol. Cell. Proteomics. 2009; 8:2285–2295. [PubMed: 19596695]
27. Lee JE, Atkins N Jr, Hatcher NG, Zamdborg L, Gillette MU, Sweedler JV, Kelleher NL. Mol. Cell. Proteomics. 2010; 9:289–297.

28. Amunugama R, Hogan JM, Newton KA, McLuckey SA. Anal. Chem. 2004; 76:720–727. [PubMed: 14750868]
29. Chi A, Bai DL, Geer LY, Shabanowitz J, Hunt DF. Int. J. Mass Spectrom. 2007; 259:197–203. [PubMed: 17364019]
30. Second TP, Blethrow JD, Schwartz JC, Merrihew GE, MacCoss MJ, Swaney DL, Russell JD, Coon JJ, Zabrouskov V. Anal. Chem. 2009; 81:7757–7765. [PubMed: 19689114]
31. Liu ZL, Schey KL. J. Am. Soc. Mass Spectrom. 2008; 19:231–238. [PubMed: 17693096]
32. Deterding LJ, Lhattacharjee S, Ramirez DC, Mason RP, Tomer KB. Anal. Chem. 2007; 79:6236–6248. [PubMed: 17637042]
33. Marshall AG, Hendrickson CL, Jackson GS. Mass Spectrom. Rev. 1998; 17:1–35. [PubMed: 9768511]
34. Hu QZ, Noll RJ, Li HY, Makarov A, Hardman M, Cooks RG. J. Mass Spectrom. 2005; 40:430–443. [PubMed: 15838939]
35. Kelleher NL, Lin HY, Valaskovic GA, Aaserud DJ, Fridriksson EK, McLafferty FW. J. Am. Chem. Soc. 1999; 121:806–812.
36. Meng FY, Du Y, Miller LM, Patrie SM, Robinson DE, Kelleher NL. Anal. Chem. 2004; 76:2852–2858. [PubMed: 15144197]
37. Pesavento JJ, Mizzen CA, Kelleher NL. Biophys. J. 2004; 86:421A.
38. Running WE, Ravipaty S, Karty JA, Reilly JP. J. Proteome Res. 2007; 6:337–347. [PubMed: 17203977]
39. Makarov A, Denisov E, Kholomeev A, Baischun W, Lange O, Strupat K, Horning S. Anal. Chem. 2006; 78:2113–2120. [PubMed: 16579588]
40. Macek B, Waanders LF, Olsen JV, Mann M. Mol. Cell. Proteom. 2006; 5:949–958.
41. Waanders LF, Hanke S, Mann M. J. Am. Soc. Mass Spectrom. 2007; 18:2058–2064. [PubMed: 17920290]
42. Graessler J, Schwudke D, Schwarz PEH, Herzog R, Shevchenko A, Bornstein SR. PLoS One. 2009; 4:e6261. [PubMed: 19603071]
43. Zubarev RA, Kelleher NL, McLafferty FW. J. Am. Chem. Soc. 1998; 120:3265–3266.
44. Syka JEP, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Proc. Natl. Acad. Sci. U. S. A. 2004; 101:9528–9533. [PubMed: 15210983]
45. Huang TY, Emory JF, O'Hair RAJ, McLuckey SA. Anal. Chem. 2006; 78:7387–7391. [PubMed: 17073403]
46. Compton PD, Shabanowitz J, Hunt DF. manuscript in preparation.
47. Armirotti A, Benatti U, Damonte G. Rapid Commun. Mass Spectrom. 2009; 23:661–666. [PubMed: 19184982]
48. Ouvry-Patat SA, Torres MP, Quek HH, Gelfand CA, O'Mullan P, Nissum M, Schroeder GK, Han J, Elliott M, Dryhurst D, Ausio J, Wolfenden R, Borchers CH. Proteomics. 2008; 8:2798–2808. [PubMed: 18655049]
49. Perkins DN, Pappin DJC, Creasy DM, Cottrell JS. Electrophoresis. 1999; 20:3551–3567. [PubMed: 10612281]
50. Kapp EA, Schutz F, Connolly LM, Chakel JA, Meza JE, Miller CA, Fenyo D, Eng JK, Adkins JN, Omenn GS, Simpson RJ. Proteomics. 2005; 5:3475–3490. [PubMed: 16047398]
51. Karabacak NM, Li L, Tiwari A, Hayward LJ, Hong PY, Easterling ML, Agar JN. Mol. Cell. Proteom. 2009; 8:846–856.
52. Zamdborg L, LeDuc RD, Glowacz KJ, Kim YB, Viswanathan V, Spaulding IT, Early BP, Bluhm EJ, Babai S, Kelleher NL. Nucleic Acids Res. 2007; 35:W701–W706. [PubMed: 17586823]
53. Wang HX, Kachman MT, Schwartz DR, Cho KR, Lubman DM. Electrophoresis. 2002; 23:3168–3181. [PubMed: 12298089]
54. Zhao J, Zhu K, Lubman DM, Miller FR, Shekhar MPV, Gerard B, Barder TJ. Proteomics. 2006; 6:3847–3861. [PubMed: 16767785]
55. Wu S, Lourette NM, Tolic N, Zhao R, Robinson EW, Tolmachev AV, Smith RD, Pasa-Tolic L. J. Proteome Res. 2009; 8:1347–1357. [PubMed: 19206473]

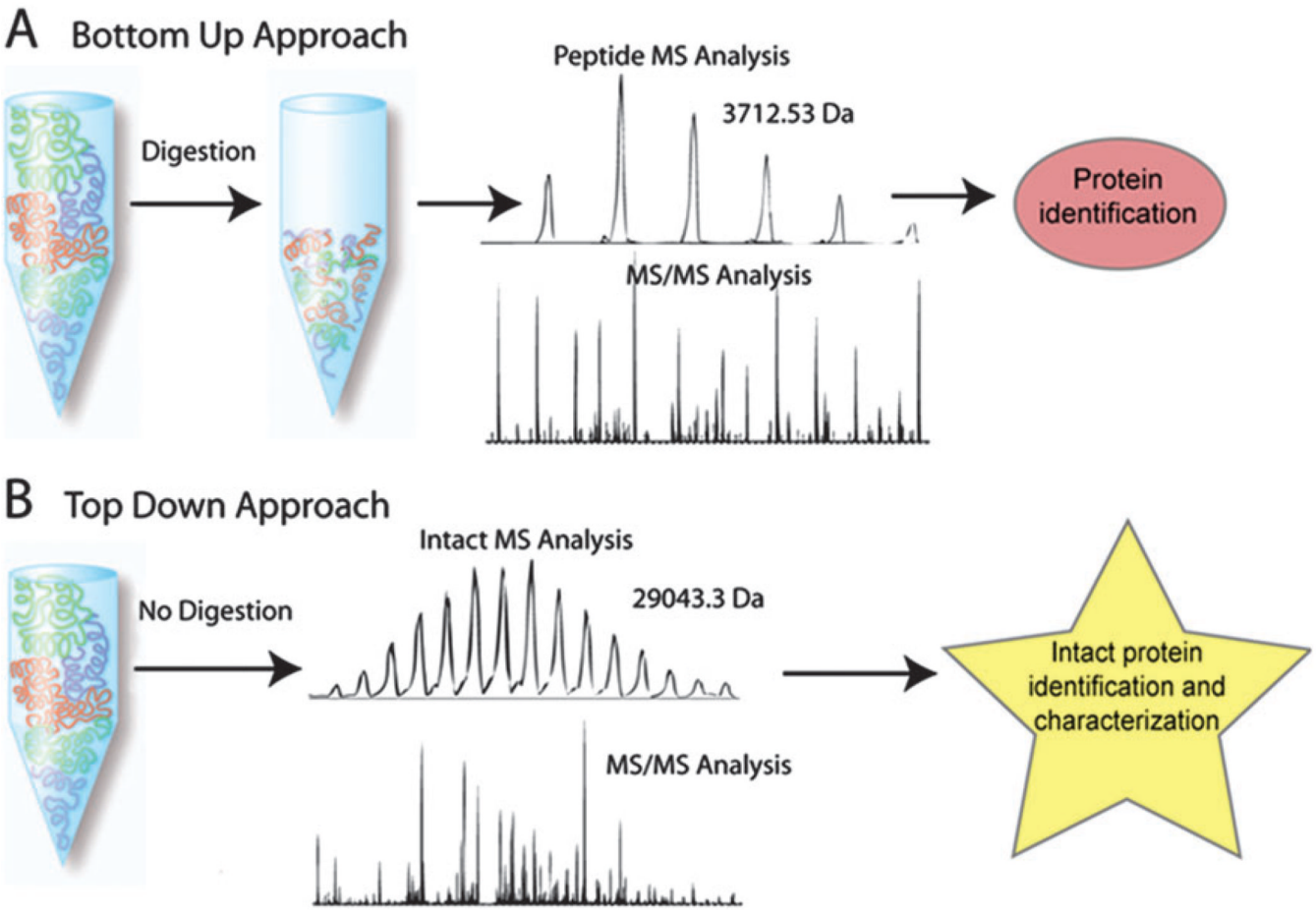


Fig. 1.

Schematic of the Bottom Up and Top Down approaches to protein identification. In the Bottom Up approach (A), enzymatic digestion is utilized to cleave intact proteins into peptides. Peptides are analyzed through tandem mass spectrometry, and protein identification can occur through identifying peptides matched *in silico* to a protein database by a search algorithm. Although low sequence coverages are typical, very high proteome coverages are common. Alternatively, in the Top Down approach (B), intact proteins are directly analyzed in the mass spectrometer (no enzymatic digestion). The resulting precursor and fragment masses are then matched to candidate sequences from a protein database. This enables the potential for full localization and characterization of post-translational modifications.

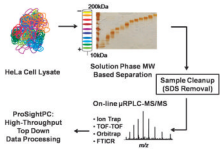


Fig. 2.

Example work flow for Top Down proteomics. Total protein content from HeLa-S3 nuclei or cytosol is quantified and loaded onto a GELFrEE column. The GELFrEE device separates the protein samples according to molecular weight. Proteins of increasing molecular weight elute into solution-phase fractions, which can be visualized on a slab gel (top right). The solution-phase fractions are cleaned up to remove SDS before injection onto a µRPLC column for MS/MS. LC-MS/MS files are processed with ProSightPC 2.0, a software suite tailored for Top Down analysis in a high-throughput setting.

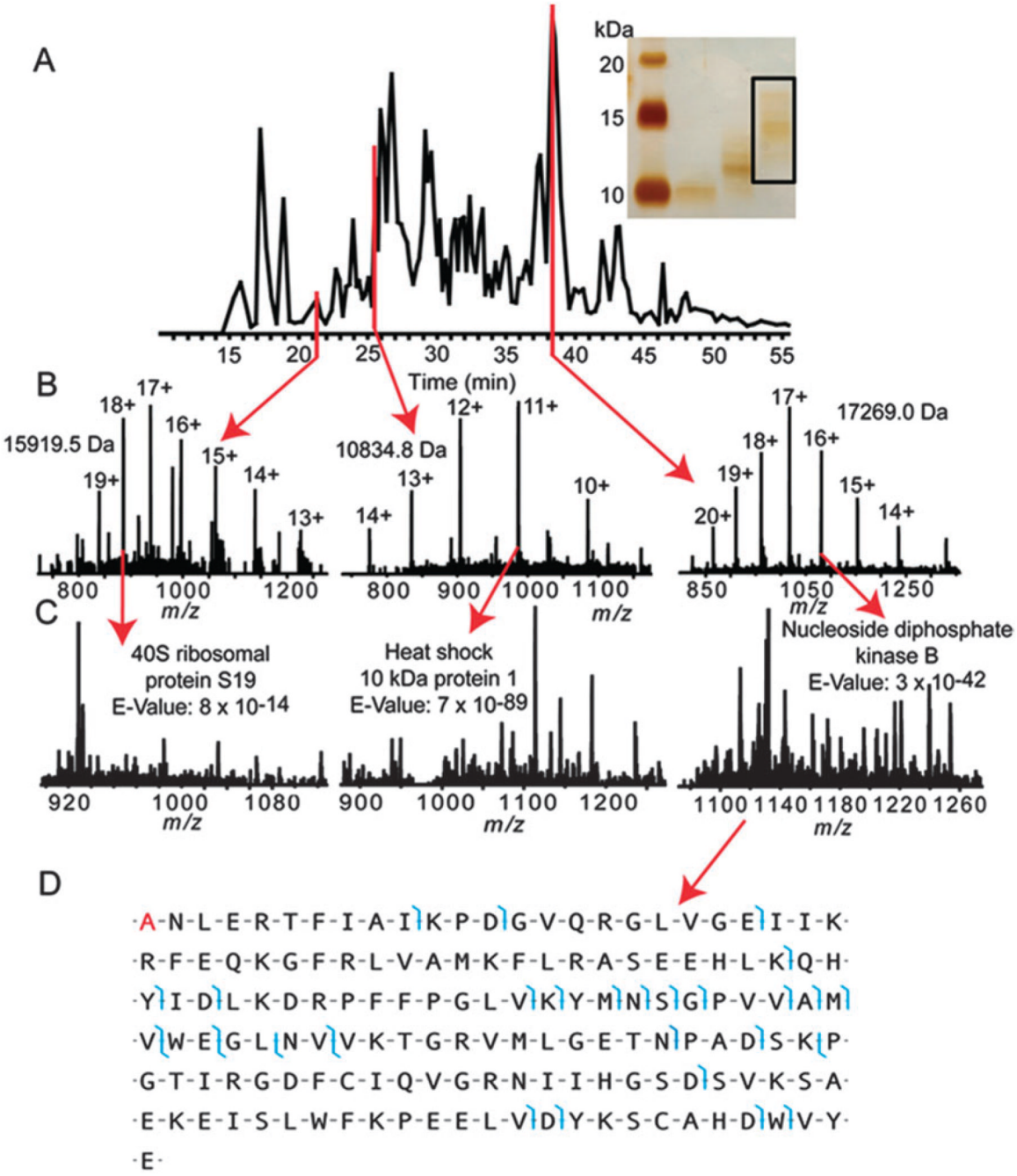


Fig. 3.

Examples selected from an LC-MS/MS injection of fraction #3 from a GELFrEE run. A base-peak chromatogram is shown (A), with charge state distributions from selected retention times shown in B. In panel C, abundant charge states (above the arrows) were targeted for fragmentation. Fragmentation mass spectra for each protein are shown along with the corresponding identifications and E-values. A fragmentation map (D) results from the matching fragment ions of nucleoside diphosphate kinase B found in HeLa cells. The protein is N-terminally acetylated.

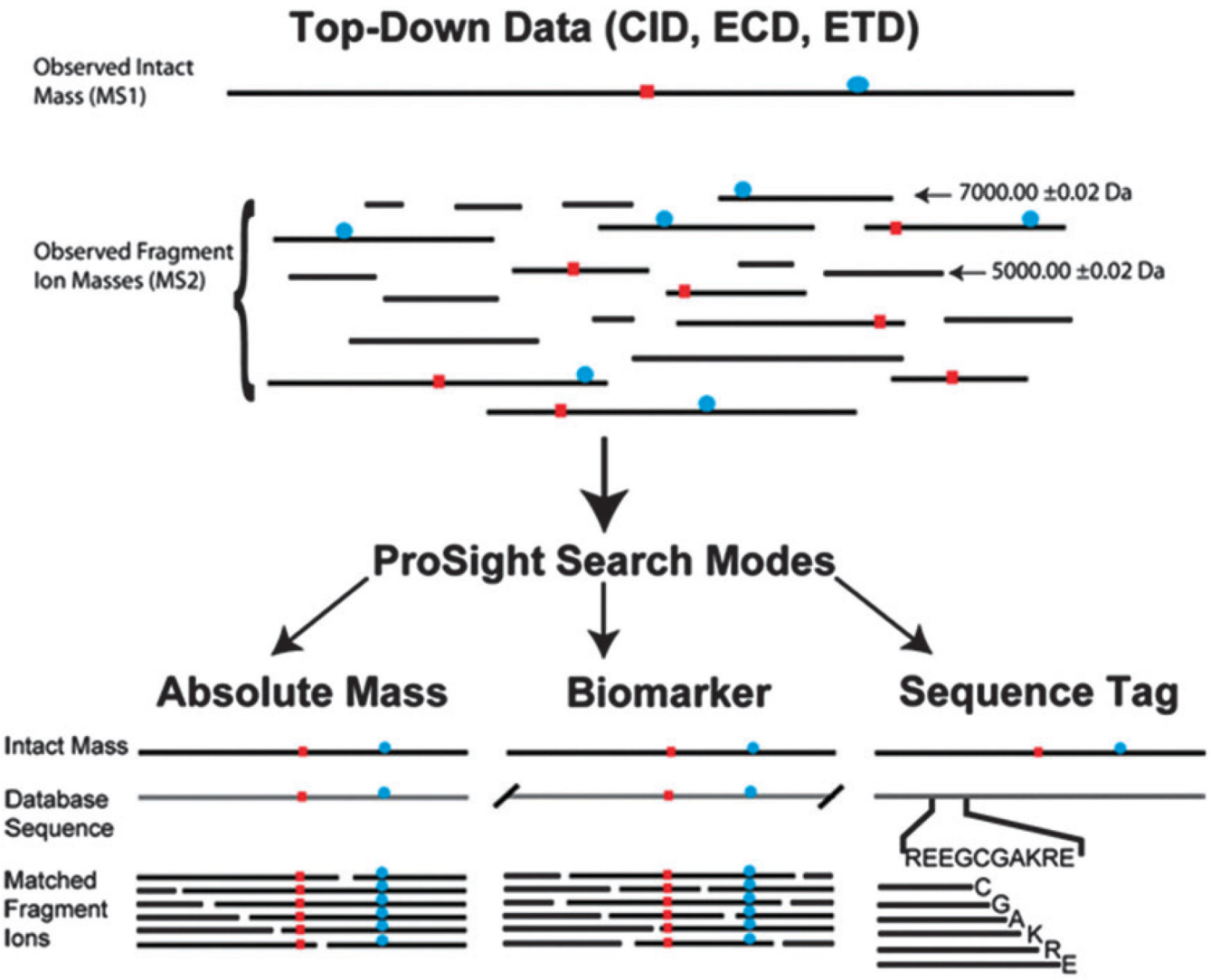


Fig. 4.

Schematic of the Top Down identification process. From the intact mass and fragmentation data (top), three different search algorithms can identify and characterize the observed protein form. These algorithms function either by matching mass values from intact proteins (absolute mass), a subset of larger sequence (biomarker), or a series of unique amino acids (sequence tag) to a database sequence. In this depiction, the power of Top Down for identification of post-translational modifications is indicated by the phosphorylation (blue) and methylation (red) on this hypothetical protein.

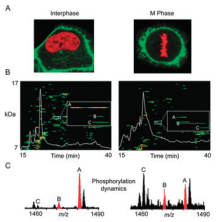


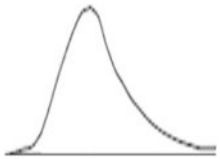
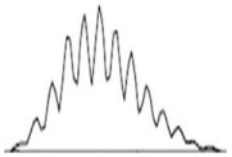
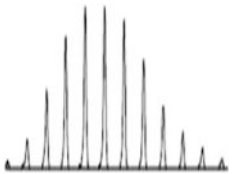
Fig. 5.

New LC-MS visualization software displays a map of masses detected as a function of LC retention time. Confocal microscopy images of live HeLa cells expressing markers for chromatin (red, histone 2B fused to red fluorescent protein) and plasma membrane (green, myristoylated/palmitoylated GFP) are shown (A).‡ In panel B, two visualizations are shown for GELFrEE fractions of similar molecular weight from unsynchronized interphase and M-Phase-arrested HeLa cells. Total-ion chromatogram traces are overlaid in each map. Post-translational modifications (phosphorylations, shown in red) are detected (C) and highlighted in panel B (inset).

‡Images provided by D. W. Gerlich.

Table 1

Characteristics of mass analyzers for use in Top Down protein analysis

Mass analyzer	Suitable for Top Down	Spectral acquisition time/s	Resolution/Da	Mass accuracy (ppm)	Performance at 8 kDa	Available fragmentation
Ion trap	+	0.05–0.3	1000	100–200		CID ETD
TOF						CID ISD
TOF-TOF Q-TOF	++	<0.01	10 000	5–20		PSD
FT Orbitrap	+++	0.1–1	60 000	3–10		CID ETD HCD CID ECD
FTICR	+++	0.1–1	200000	1–3		IRMPD

Abbreviations used: TOF, Time of Flight; Q, Quadrupole; FT, Fourier Transform; ICR, Ion Cyclotron Resonance; CID, Collision-Induced Dissociation; ETD, Electron Transfer Dissociation; ISD, In-Source Decay; PSD, Post-Source Decay; HCD, High Energy Collision; IRMPD, Infrared Multiphoton Dissociation. +, less suitable; ++, suitable; +++, very suitable.