HHS Public Access

Author manuscript

J Proteomics. Author manuscript; available in PMC 2019 March 20.

Published in final edited form as:

J Proteomics. 2018 March 20; 175: 3–4. doi:10.1016/j.jprot.2017.02.002.

Top-down proteomics: where we are, where we are going?

Luca Fornelli, **Timothy K. Toby**, **Luis F. Schachner**, **Peter F. Doubleday**, **Kristina Srzentić**, **Caroline J. DeHart**, and **Neil L. Kelleher***

Departments of Chemistry and Molecular Biosciences, Northwestern University, 2170 Campus Drive, Evanston, Illinois, 60208

Increasing Momentum

After some years spent increasing metrics of proteome coverage, the field of proteomics has increasingly focused on the *quality* of information generated during interrogation of living systems. Another aspect trending presently is to integrate proteomics with data from other "-omics" in order to gain deeper insights into cellular and disease biology. From this perspective, it is apparent that the analysis of intact proteoforms, or top-down proteomics,[1] presents additional advantages. In providing precise compositional information, TDP can add molecular details lost when proteoforms are dissected into proteolytic peptides used in bottom-up proteomics. Although BUP can identify and localize post-translational modifications (PTMs) on proteins, the well-known 'protein inference problem' greatly complicates the elucidation of their global patterns or cross-talk, aspects which can be captured using TDP. This is exemplified by the so-called "histone code", where PTMs comprise combinatorial and highly dynamic patterns resulting from concerted interactions between prior PTMs and histone-modifying enzymes. These patterns govern the reading of histone marks and myriad biomolecular activities, and can be comprehensively described at the proteoform level to help assign the functions of PTM patterns.

It is true that top-down proteomics is technologically challenging, yet perceptions about this are often historical and not updated quickly in the minds of experts or those far afield. Rapid advances in instrumentation by most manufacturers over the decade have rendered targeted and high-throughput top-down feasible—even on the benchtop mass spectrometers of today. Moreover, a collection of top-down proteomics practitioners have founded the Consortium for Top Down Proteomics (**CTDP**, http://topdownproteomics.org/), many of which have contributed articles to this special issue. **CTDP** has the aims of accelerating impactful and collaborative research,[2] increasing the visibility of top-down proteomics within the community, and promulgating knowledge and best practices for newcomers to the field. Notably, the first **CTDP** manuscript rallied interest and focus around the "proteoform" and now has over 220 citations.[3] As further recognition of the accelerating progress in top-

* n-kelleher@northwestern.edu, Phone: 847-467-4362, Fax: 847-467-3276.

The authors declare no conflicts of interest in preparing this piece.

down proteomics, the UniProt Knowledgebase has begun cross-referencing their accession numbers with the permanent PFR identifiers in the Proteoform Atlas hosted by the **CTDP** (http://repository.topdownproteomics.org/).

Denaturing and Native Top-Down MS

Typically, top-down proteomics has been performed under denaturing conditions, following workflows originally developed for bottom-up proteomics. Namely, proteins are isolated from cells in detergent-containing buffers, and subjected to pre-fractionation to reduce sample complexity. In recent years, combined with concurrent advances in dedicated high-throughput data analysis platforms, top-down proteomics of proteins <30 kDa has been sufficiently optimized to enable a degree of qualitative and quantitative proteome coverage more typically achieved by bottom-up. Today, high-confidence identification and characterization of several thousand proteoforms is feasible.[4–6] However, the number of proteoforms identified by denaturing top-down proteomics drops off at molecular weights (MW) exceeding ~50–70 kDa due to technical issues associated with both the limitations of mass spectrometry and the need for improved separations. Solutions to extend the MW range of proteoforms identified during denaturing top-down experiments based on Fourier-transform mass spectrometry include the use of new generation instruments such as the 21T FT-ICR mass spectrometers,[7,8] or novel data acquisition strategies aimed at increasing the spectral signal-to-noise of large proteoforms by recording short time-domain transients.

Another option for high mass proteins derives from native top-down mass spectrometry, which preserves non-covalent interactions, labile PTMs, cofactors and physiological stoichiometry.[9] Due to these unique features, native top-down mass spectrometry has traditionally been employed for the targeted analysis of highly purified protein complexes, remarkably allowing the analysis of macromolecular assemblies with masses up to the megaDalton range, like in the case of the intact ribosome[10,11] or viral capside assemblies.[12] Furthermore, native top-down technology can provide unique insights into the perturbations in proteoform-proteoform interactions which regulate specific cellular processes.[13] Alternatively to its complementary role in structural biology,[14,15] the high-MW capabilities of native top-down might someday be employed on-line with the appropriate protein separation techniques such as ion exchange chromatography, potentially increasing the analysis throughput and steering native top-down toward “discovery mode” proteomics.[16] Along with the instrumentation development, this possible direction for native mass spectrometry will also encourage the creation of new bioinformatics, particularly when the analysis will focus on multiproteoform complexes, requiring database search algorithms that consider both the intact assembly and the single subunits.[17] With many analytical benefits and increased accessibility to biologically relevant protein structures and interactions, native top-down mass spectrometry has the potential to not only access the higher-MW regions of the human proteome but also to provide enhanced information and molecular detail.

When will a tipping point be reached?

As it stands today, bottom-up and top-down proteomics stand as highly complementary in terms of value proposition to both researchers and stakeholders. While bottom-up remains unmatched in depth of proteome coverage and quantitation of peptides, top-down proteomics characterizes proteoforms and their variants directly. Native top-down can hone in on endogenous protein complexes.[18] As techniques for denaturing and native top-down continue to evolve, one could extrapolate toward a tipping point where the analysis of intact proteins becomes far more widespread. The “Why?” for this change is becoming more clear — great progress in proteoform-resolved measurements has the value of high molecular specificity when it comes to protein molecules, often with strong mechanistic connections to pathology. As an example, a recent report of phosphorylated alpha-synuclein proteoforms in Parkinson’s Disease[19] demonstrates the new kind of information being provided by proteoform-resolved measurements. The “When?” is the crux of the issue: when will the value of these measurements lead to wider adoption and deployment of the technology in academic and clinical research? There is a common assertion that top-down is several years behind bottom-up in terms of acceptance and implementation in the field. If true, then perhaps increasing numbers of practitioners will adopt the direct interrogation of intact proteins and their protein complexes. Advancements in top-down proteomics like those enclosed in this special issue will be the driving agents of that change, and the implications for the study of protein molecules and the translation of this knowledge to understanding disease hold a good deal of unrealized potential.

Acknowledgments

The authors would like to acknowledge the W.M. Keck foundation for generous support and funding (DT061512). In addition, this work was partially supported the National Resource for Translational and Developmental Proteomics under Grant P41 GM108569 from the National Institute of General Medical Sciences, National Institutes of Health.

References

1. Toby TK, Fornelli L, Kelleher NL. Progress in Top-Down Proteomics and the Analysis of Proteoforms. Annu Rev Anal Chem. 2016; 9:499–519.
2. Dang X, Scotcher J, Wu S, Chu RK, Tolic N, Ntai I, Thomas PM, Fellers RT, Early BP, Zheng Y, Durbin KR, Leduc RD, Wolff JJ, Thompson CJ, Pan J, Han J, Shaw JB, Salisbury JP, Easterling M, Borchers CH, Brodbelt JS, Agar JN, Pasa-Tolic L, Kelleher NL, Young NL. The first pilot project of the consortium for top-down proteomics: a status report. Proteomics. 2014; 14:1130–1140. [PubMed: 24644084]
3. Smith LM, Kelleher NL. Consortium for Top Down Proteomics. Proteoform: a single term describing protein complexity. Nat Methods. 2013; 10:186–187. [PubMed: 23443629]
4. Durbin KR, Fornelli L, Fellers RT, Doubleday PF, Narita M, Kelleher NL. Quantitation and Identification of Thousands of Human Proteoforms below 30 kDa. J Proteome Res. 2016; 15:976–982. [PubMed: 26795204]
5. Anderson JC, Wan Y, Kim YM, Pasa-Tolic L, Metz TO, Peck SC. Decreased abundance of type III secretion system-inducing signals in Arabidopsis mkp1 enhances resistance against Pseudomonas syringae. Proc Natl Acad Sci U S A. 2014; 111:6846–6851. [PubMed: 24753604]
6. Vorontsov EA, Rensen E, Prangishvili D, Krupovic M, Chamot-Rooke J. Abundant Lysine Methylation and N-Terminal Acetylation in Sulfolobus islandicus Revealed by Bottom-Up and Top-Down Proteomics. Mol Cell Proteomics. 2016; 15:3388–3404. [PubMed: 27555370]

7. Shaw JB, Lin TY, Leach FE 3rd, Tolmachev AV, Tolic N, Robinson EW, Koppenaal DW, Pasa-Tolic L. 21 Tesla Fourier Transform Ion Cyclotron Resonance Mass Spectrometer Greatly Expands Mass Spectrometry Toolbox. J Am Soc Mass Spectrom. 2016; 27:1929–1936. [PubMed: 27734325]

8. Hendrickson CL, Quinn JP, Kaiser NK, Smith DF, Blakney GT, Chen T, Marshall AG, Weisbrod CR, Beu SC. 21 Tesla Fourier Transform Ion Cyclotron Resonance Mass Spectrometer: A National Resource for Ultrahigh Resolution Mass Analysis. J Am Soc Mass Spectrom. 2015; 26:1626–1632. [PubMed: 26091892]

9. Loo JA. Electrospray ionization mass spectrometry: a technology for studying noncovalent macromolecular complexes. Int J Mass Spectrom. 2000; 200:175–186.

10. Rostom AA, Fucini P, Benjamin DR, Juenemann R, Nierhaus KH, Hartl FU, Dobson CM, Robinson CV. Detection and selective dissociation of intact ribosomes in a mass spectrometer. Proc Natl Acad Sci U S A. 2000; 97:5185–5190. [PubMed: 10805779]

11. Loo JA, Berhane B, Kaddis CS, Wooding KM, Xie Y, Kaufman SL, Chernushevich IV. Electrospray ionization mass spectrometry and ion mobility analysis of the 20S proteasome complex. J Am Soc Mass Spectrom. 2005; 16:998–1008. [PubMed: 15914020]

12. Snijder J, Rose RJ, Veesler D, Johnson JE, Heck AJ. Studying 18 MDa virus assemblies with native mass spectrometry. Angew Chem Int Ed. 2013; 52:4020–4023.

13. Snijder J, Burnley RJ, Wiegard A, Melquiond ASJ, Bonvin AMJJ, Axmann IM, Heck AJR. Insight into cyanobacterial circadian timing from structural details of the KaiB-KaiC interaction. Proc Natl Acad Sci U S A. 2014; 111:1379–1384. [PubMed: 24474762]

14. Benesch JLP, Ruotolo BT. Mass spectrometry: come of age for structural and dynamical biology. Curr Opin Struc Biol. 2011; 21:641–649.

15. Dyachenko A, Gruber R, Shimon L, Horovitz A, Sharon M. Allosteric mechanisms can be distinguished using structural mass spectrometry. Proc Natl Acad Sci U S A. 2013; 110:7235–7239. [PubMed: 23589876]

16. Muneeruddin K, Nazzaro M, Kaltashov IA. Characterization of Intact Protein Conjugates and Biopharmaceuticals Using Ion-Exchange Chromatography with Online Detection by Native Electrospray Ionization Mass Spectrometry and Top-Down Tandem Mass Spectrometry. Anal Chem. 2015; 87:10138–10145. [PubMed: 26360183]

17. Skinner OS, Havugimana PC, Haverland NA, Fornelli L, Early BP, Greer JB, Fellers RT, Durbin KR, Do Vale LHF, Melani RD, Seckler HS, Nelp MT, Belov ME, Horning SR, Makarov AA, LeDuc RD, Bandarian V, Compton PD, Kelleher NL. An informatic framework for decoding protein complexes by top-down mass spectrometry. Nat Methods. 2016; 13:237–240. [PubMed: 26780093]

18. Ntai I, LeDuc RD, Fellers RT, Erdmann-Gilmore P, Davies SR, Rumsey J, Early BP, Thomas PM, Li S, Compton PD, Ellis MJ, Ruggles KV, Fenyo D, Boja ES, Rodriguez H, Townsend RR, Kelleher NL. Integrated Bottom-Up and Top-Down Proteomics of Patient-Derived Breast Tumor Xenografts. Mol Cell Proteomics. 2016; 15:45–56. [PubMed: 26503891]

19. Kellie JF, Higgs RE, Ryder JW, Major A, Beach TG, Adler CH, Merchant K, Knierman MD. Quantitative measurement of intact alpha-synuclein proteoforms from post-mortem control and Parkinson's disease brain tissue by intact protein mass spectrometry. Sci Rep. 2014; 4:5797. [PubMed: 25052239]