

OPINION

# The emergence of top-down proteomics in clinical research

John P Savaryn<sup>1</sup>, Adam D Catherman<sup>1</sup>, Paul M Thomas<sup>1</sup>, Michael M Abecassis<sup>2</sup> and Neil L Kelleher<sup>1,3\*</sup>

## Abstract

Proteomic technology has advanced steadily since the development of 'soft-ionization' techniques for mass-spectrometry-based molecular identification more than two decades ago. Now, the large-scale analysis of proteins (proteomics) is a mainstay of biological research and clinical translation, with researchers seeking molecular diagnostics, as well as protein-based markers for personalized medicine. Proteomic strategies using the protease trypsin (known as bottom-up proteomics) were the first to be developed and optimized and form the dominant approach at present. However, researchers are now beginning to understand the limitations of bottom-up techniques, namely the inability to characterize and quantify intact protein molecules from a complex mixture of digested peptides. To overcome these limitations, several laboratories are taking a whole-protein-based approach, in which intact protein molecules are the analytical targets for characterization and quantification. We discuss these top-down techniques and how they have been applied to clinical research and are likely to be applied in the near future. Given the recent improvements in mass-spectrometry-based proteomics and stronger cooperation between researchers, clinicians and statisticians, both peptide-based (bottom-up) strategies and whole-protein-based (top-down) strategies are set to complement each other and help researchers and clinicians better understand and detect complex disease phenotypes.

predicting disease prognosis and identifying druggable targets for new therapeutics. Diagnostic or companion diagnostic biomarkers are greatly sought after. The holy grail of biomarker discovery, however, is proteomic biomarkers that predict that a given phenotype will develop. Great progress has been made toward these goals over the past 20 years, and proteomics has been a powerful tool for providing information about a broad range of diseases and clinical phenotypes. However, compared with the discoveries that rapidly followed completion of the Human Genome Project, the translation of proteomic information into medical advances has been slower than expected. A plethora of biological information has been obtained, yet the data have minimal clinical relevance. This type of discovery-based protein analysis has, therefore, been associated with a high cost and a low return on investment. Despite the modest use of proteomics within clinical applications, many in the field are optimistic that proteomics, which is still evolving, will play an important part in 21st century medicine [1,2].

Proteomic research has mostly been dominated by bottom-up techniques. Such techniques involve *in vitro* enzymatic digestion of the sample and mass spectrometry (MS)-based analysis of the resultant peptide mixture. Inferences are then drawn about the protein composition of the sample. Over the last 20 years, such bottom-up methods have been developed into extremely sensitive and selective methods capable of identifying >5,000 proteins within a single sample. These methods follow in the footsteps of many 'small-molecule' liquid chromatography (LC)-MS assays that have been approved by the US Food and Drug Administration (for example, those for vitamin D3, glycosphingolipids and thyroglobulin) and are poised to augment this capability in the clinical research laboratory [3].

Bottom-up technology has produced a myriad of proteomic data for many living systems [4-6], enabled innovative ways for understanding disease [7] and provided new leads for clinical diagnostics [8]; however, the complete proteomic tool kit for 21st century research will consist of orthogonal methods that allow analysis at multiple levels: the peptide, whole-protein and intact

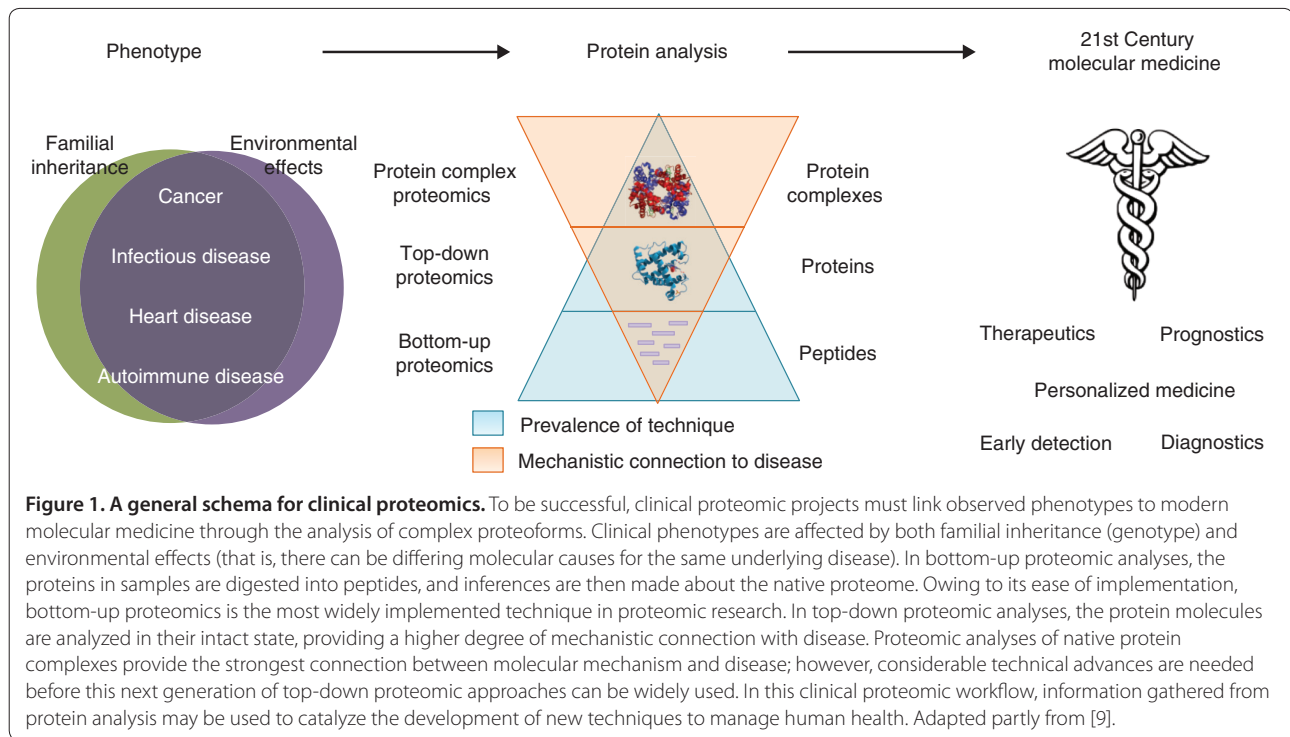
## Proteomics in 21st century medicine

The main goals of using proteomics in translational research include detecting disease in the early stages,

\*Correspondence: n-kelleher@northwestern.edu

<sup>1</sup>Departments of Chemistry and Molecular Biosciences, the Chemistry of Life Processes Institute, the Proteomics Center of Excellence, 2145 N. Sheridan Dr, Evanston, IL 60208, USA

Full list of author information is available at the end of the article



protein complex levels [9]. Although bottom-up proteomic technology is well developed, the technology for analyzing whole proteins (known as top-down proteomics) and intact protein complexes (known as next-generation top-down proteomics or protein complex proteomics) is less so (Figure 1, center). Notwithstanding the nascent technology, biological research will benefit greatly from a combined proteomic approach that can take advantage of the individual strengths of all three approaches to complement the deficiencies inherent in each. We propose that such a combination approach will result in an increased return on investment for MS-based proteomics in the next decade or two and therefore a greater impact on human health (Figure 1).

### State-of-the-art bottom-up proteomics in clinical research

Most clinical proteomic research focuses on identifying the molecular signatures of specific diseases or disease phenotypes from relevant biological samples from patients. When found, these molecular signatures, or biomarkers, provide novel ways to detect, understand and, perhaps, treat disease. Much of the search for biomarkers has been conducted on human serum or plasma. Although plasma is readily obtainable, it is daunting in its proteomic complexity, owing to a vast dynamic range of component concentrations within a single sample that spans more than ten orders of magnitude [10]. Not surprisingly, thorough analysis of

the protein composition of plasma is a challenge. Nevertheless, techniques for carrying out targeted measurements in human serum have been developed.

One such technique is an antibody-based enrichment strategy termed SISCAPA (stable isotope standards and capture by antipeptide antibodies). Whiteaker *et al.* [11] used SISCAPA to achieve a >1,000-fold enrichment of target peptides within plasma and to detect analytes in the nanogram per milliliter range using an ion-trap mass spectrometer. Another technique that has now been widely implemented is multiple reaction monitoring (MRM), which measures targeted peptides within complex mixtures and can be used for absolute quantification of these peptides [12]. For example, by optimizing sample preparation and measurement conditions, Keshishian *et al.* [13] used MRM and achieved limits of quantification (LOQs) in the low nanogram per milliliter range without the need for antibody-based enrichment. Although the antibody-based methods used in clinical laboratories can achieve much lower LOQs, in the picogram to femtogram per milliliter range, as is the case for cardiac troponin and prostate-specific antigen [14,15], optimized MRM assays coupled with SISCAPA could represent the future of biomarker validation assays [16].

Examples of MRM successes in clinical research include the following: the quantification of proteins in the cerebrospinal fluid to aid understanding of the later stages of multiple sclerosis [17]; the development of quantitative validation techniques for plasma biomarkers,

with LOQs reaching picograms per milliliter [13]; and the demonstration of robust targeted assays for cancer-associated protein quantification in both plasma and urine samples from patients [18]. In the first example, Jia *et al.* [17] used MRM to quantify 26 proteins from the cerebrospinal fluid of patients with secondary progressive multiple sclerosis. They included patients with a non-inflammatory neurological disorder and healthy humans as controls. The many significant differences in the abundance of certain proteins between patient groups may hold true upon further sampling and could yield important insight and provide a new method for multiple sclerosis research [17]. In the second example, Keshishian *et al.* [13] performed important empirical testing of serum-processing options and provided a method for achieving an LOQ appropriate for current serum biomarkers (low nanogram per milliliter), even while multiplexing the assay to monitor multiple analytes. In the third example, Huttenhain *et al.* [18] extended this empirical testing to develop MRM assays for over 1,000 cancer-associated proteins in both serum and urine. They extended their results to monitor, using MS, the levels of four biomarkers that are currently used to assess ovarian cancer risk (apolipoprotein A1, transferrin,  $\beta_2$ -microglobulin and transthyretin; using Quest Diagnostics' OVA1 enzyme-linked immunosorbent assay (ELISA) panel). In a panel of 83 serum samples, they found significant differences in the abundance of these proteins between patients with ovarian cancer and those with benign ovarian tumors, and these differences were consistent with prior results obtained from immunoassays. This study exemplifies the strength of MRM for multiplexed quantification of peptide biomarkers in complex clinical samples.

MRM offers unrivaled utility for sensitive and accurate detection of target peptides in clinical samples (information that is subsequently used to infer the presence and level of proteins in the sample). However, the proteome harbors more complexity than typical MRM assays can interrogate. This analytical mismatch confounds the diagnostic accuracy of the MRM-based assays in ways that are not possible to overcome by using bottom-up MS-based proteomic technology alone.

One issue with MRM is that it is a targeted assay and relies on *a priori* knowledge of the protein to be measured. At present, most of that knowledge is obtained from bottom-up, discovery-type proteomic studies, in which enzymatic digestion precedes the peptide-based analysis of proteins in complex mixtures. Herein lies the key limitation of bottom-up strategies. With enzymatic digestion, the information describing individual intact proteins is lost, preventing complete characterization of all of the protein forms expressed at one time for any given protein-coding gene. As a result, clinical conclusions

are based on potentially inaccurate protein expression levels, because these levels are derived from quantifying peptides that may not be representative of all of the diverse forms of protein molecules present. (For example, the peptide sequence is common to many forms of a protein molecule; however, some forms are post-translationally modified on amino acids within the same stretch of sequence.) The net effect of a bottom-up strategy is that MRM peptides report only generally on protein expression of a gene, because modified peptides that represent individual protein molecules are unlikely to be discovered upon enzymatic digestion in an untargeted fashion.

Measuring the expression of protein-coding genes at the protein level is important; however, in a living system, it is the individual protein molecules that are likely to correlate more tightly with (aberrant) molecular functions. Because these individual protein molecules (which, for example, contain coding polymorphisms, mutations, splicing variations and post-translational modifications) are likely to perform different functions from other modified versions of the same parent protein [19], it becomes imperative to measure protein expression with a precision that will distinguish between even closely related intact protein forms. Top-down proteomics offers this precision.

### Top-down proteomic approaches

Top-down MS-based proteomic technology provides the highest molecular precision for analyzing primary structures by examining proteins in their intact state, without the use of enzymatic digestion. In doing so, top-down proteomic techniques can fully characterize the composition of individual protein molecules (these intact protein molecules were recently coined 'proteoforms' [20]). Traditionally, the top-down strategy consisted of two-dimensional protein separation involving isoelectric focusing and PAGE followed by visualization of the protein spots within the gel, a technique known as two-dimensional gel electrophoresis. Both two-dimensional gel electrophoresis [21] and difference gel electrophoresis [22] facilitate a 'birds'-eye' view of the proteins in a sample in one or more biological states. Salient proteome features are then further investigated by identifying the proteins of interest using bottom-up MS. These techniques provide a large visual representation of the proteome and have been applied in disease research, such as cancer research [23,24]; however, several technical challenges have impeded the universal adoption of this top-down approach. First, there are limitations on proteome resolution, leading to the co-migration of multiple proteins to the same location on the gel. Second, there are issues with gel-to-gel reproducibility. Third, this approach is labor intensive. Last, the enzymatic digestion

required for MS identification prevents full molecular characterization [25,26].

An alternative method for top-down proteomics, and the front-runner for becoming the technique of choice for top-down proteomics, is LC electrospray ionization tandem MS (LC-ESI-MS/MS). This soft-ionization technique can be applied to intact proteins of up to approximately 50 kDa using hybrid instruments offering Fourier-transform-based high-resolution measurements [27]. The high-resolution LC-ESI-MS/MS approach to top-down proteomics has recently proven to be capable of truly high-throughput protein identification [28] and is now appreciated as a viable option for proteome discovery [29].

We hypothesize, as do many researchers in the top-down proteomics field, that the information obtained from precise, comprehensive whole-protein analysis will be connected more directly to complex disease phenotypes than information gained from bottom-up analyses. As a result, studying proteomes at the whole-protein level will provide a more efficient translation of proteomic data into phenotypic understanding and early detection of disease. At present, top-down proteomic techniques are less sensitive than bottom-up strategies, which presents concerns for biomarker studies. Nonetheless, there is a need for a combined approach to translational proteomics that uses both top-down and bottom-up strategies. Figure 2 depicts the positioning of whole-protein (top-down) analysis and peptide-based (bottom-up) protein analysis in the space of complex human disease. With complete protein characterization afforded by top-down analyses, sensitive MRM assays with LOQs in the nanogram per milliliter range can be developed to target the exact proteoforms that are most closely connected to the disease phenotype of interest. When proteoforms are larger than the current limit for top-down proteomics, which is approximately 50 kDa, an intermediate technique called middle-down proteomics can be used. With this technique, targeted enzymatic digestion occurs minimally throughout the protein to produce large peptides with an average size of about 6 kDa [30]. These large stretches of polypeptide can facilitate partial characterization of large proteins (>50 kDa) and allows better proteoform specificity in MRM assay development.

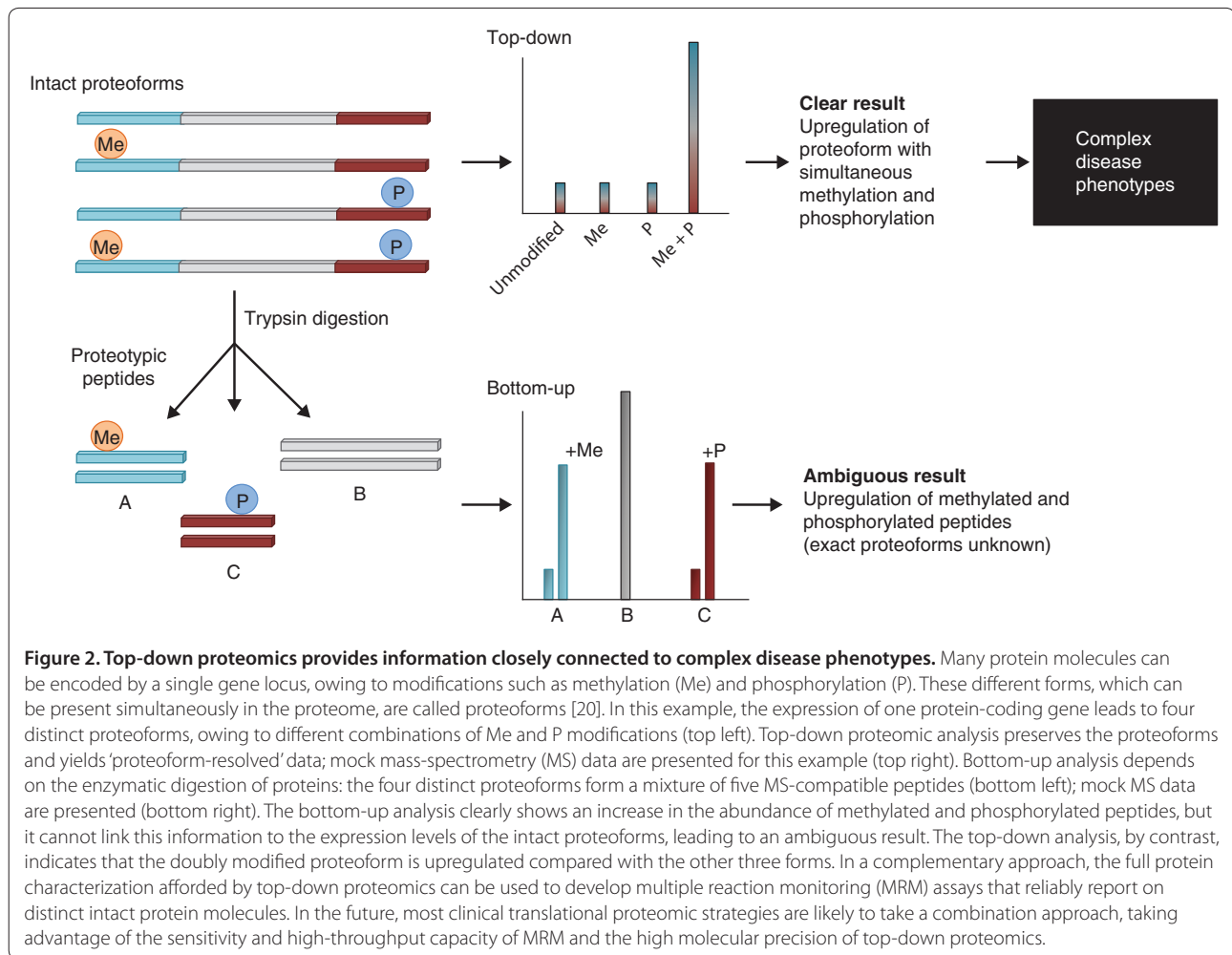
### Recent advances in top-down proteomic implementation

At present, proteomic approaches in clinical research can be grouped into two categories: protein-profiling approaches, and protein identification and characterization using the 'grind and find' strategy. In addition to the two-dimensional gel electrophoresis and difference gel electrophoresis methods described above, another

historical profiling approach was surface-enhanced laser desorption/ionization time-of-flight MS (SELDI-TOF MS). In SELDI-TOF MS, a solid-phase enrichment step is used to bind proteins in complex mixtures, most often serum or plasma, reducing the sample complexity by compressing the dynamic range of the sample to be analyzed. Then, laser desorption is used to ionize the proteins from the surface directly into a time-of-flight mass analyzer for MS profiling. With its ability to decrease the daunting complexity of plasma [10] to make it more amenable to protein profiling, SELDI-TOF analysis was once a highly touted technique for plasma proteomic studies, particularly for biomarker discovery assays. One of the main early arguments in favor of such an approach was offered by Petricoin and Liotta [31]. They argued that although SELDI-TOF was purely an MSI profiling technique, which does not provide enough mass or chemical selectivity to ensure that a differentially expressed mass is a unique entity, comparison of the collective profile of disease and non-disease samples could uncover genuine biomarker signatures, and it would be the signatures rather than the identification of any one biomarker that would have an impact on medicine.

MS imaging (MSI) is a protein-profiling technique that is similar in certain respects to SELDI-TOF and is rapidly gaining popularity because of its innovative pairing with topological information at both the tissue and cellular levels. Sweedler and Caprioli are pioneers of MSI using matrix-assisted laser desorption/ionization (MALDI) MS, and they have applied this approach to answer many biological questions. For all applications, researchers are finding much value in being able to pinpoint protein MS profiles to certain locations within a tissue slice or organism, depending on the type of sample at which the experiment is aimed. One striking use of MSI has been to identify biomarker profiles of renal cell carcinoma in kidney tissue [32] (Table 1). Progress in this burgeoning area of clinical research will involve identifying and precisely characterizing the proteoforms detected by MSI-based profiling approaches.

In the protein characterization mode of analysis, top-down proteomics has been applied in several high-profile translational research projects (Table 1). In contrast to the proteome profiling of modern MS-based imaging techniques, top-down proteomics offers protein identification, molecular characterization (often complete) and relative quantification of related protein species. For example, Chamot-Rooke and colleagues [33] are taking advantage of top-down proteomics to identify factors associated with the invasiveness of the bacterium *Neisseria meningitidis*. They used precision MS to quantify the expression of proteoforms in type IV pili, implicating these structures in the detachment of bacteria from



meningitis-associated tissue [33]. In a similar manner, Ge and colleagues have been performing top-down analyses on intact cardiac troponin I proteoforms to gain insight into myocardial dysfunction. In a recent study, the Ge group observed an increase in phosphorylation in the failing human myocardium by examining the proteoforms of intact cardiac troponin I [34]. Interestingly, they also unambiguously localized the phosphorylation events within the protein and uncovered information that is important for gaining a mechanistic understanding of myocardial failure. In another example of proteoform-resolved top-down analysis, Hendrickson and Yates and colleagues [35] identified, characterized and quantified multiple proteoforms of apolipoprotein CIII within human blood, including those with O-linked glycosylation. Their research is important not only because it extends the concept of proteoform quantification but also because apolipoprotein CIII is associated with coronary artery disease.

Other groups are using MS coupled with hydrogen-deuterium (H-D)-exchange chemistry to study the

dynamics of intact proteins. In a potent application of H-D-exchange mass-spectrometry, Agar and colleagues [36] studied the protein dynamics of superoxide dismutase 1 variants associated with familial amyotrophic lateral sclerosis. In the variants analyzed, they found a common structural and dynamic change within the electrostatic loop of the protein [36]. Their data provide important molecular mechanistic insight into this inherited form of motor neuron disease and further exemplify the utility of proteoform-resolved data from intact proteins for informing clinical research.

### The future of top-down strategies in clinical proteomics

Support for using top-down proteomics in clinical research is growing with each publication that features its use. The examples described above were hard won by early adopters of the technique and illustrate the application of whole-protein analysis to a diverse range of disease-related questions that can be answered with proteoform-resolved information (Table 1). However, even

**Table 1. Selected applications of intact protein analysis in translational research**

Laboratory <sup>a</sup>	Disease or condition	Application description	Reference
Chamot-Rooke	Bacterial meningitis	Relative quantification of intact <i>Neisseria meningitidis</i> type IV pilus proteoforms	[33]
Agar	Neurodegeneration	H-D exchange-enabled analysis of fALS SOD1 variant protein dynamics	[36]
Ge	Myocardial dysfunction	Relative quantification of intact cardiac troponin I proteoforms	[34]
Caprioli	Renal carcinoma	Tissue profiling of intact proteins in cancerous versus healthy kidneys	[32]
Hendrickson and Yates	Coronary artery disease	Relative quantification of intact apolipoprotein CIII proteoforms	[35]
Nelson	Diabetes	Relative quantification of proteoforms in plasma from healthy individuals and diabetics[41]	

<sup>a</sup>Examples of laboratories applying top-down proteomic strategies to clinically related research are presented here, with references to their recent work. In this diverse research, top-down proteomics is being used to understand the dynamics of intact proteins, to measure the relative abundances of intact proteoforms and to provide mass spectrometry profiles of intact proteins directly from human tissue. In all of these cases, the information obtained by studying whole proteins has led to significant insight into human disease. fALS, familial amyotrophic lateral sclerosis; H-D, hydrogen-deuterium; SOD1, superoxide dismutase 1.

with these tangible examples of top-down proteomics providing an unmatched level of analytical resolution, the technique is not as widespread as its bottom-up counterpart. One of the main reasons why top-down proteomics is somewhat esoteric at present is that it took longer to develop into a high-throughput assay. It was not until 2011 that top-down proteomics was shown to be applicable to large-scale experiments [28]. Before then, its use was limited to a focused approach for characterizing targeted proteins within samples. Much of the top-down proteomic research described above fits into this category. However, now that top-down proteomics can be performed on Orbitrap MS instruments without the need for a superconducting magnet, as recently demonstrated by Ahlf *et al.* [37] and Tian *et al.* [38], it is expected that more laboratories will begin to apply high-throughput top-down techniques regularly without needing collaborators. In fact, a new Consortium for Top Down Proteomics has formed, with the mission ‘to promote innovative research, collaboration and education accelerating the comprehensive analysis of intact proteins’ [39].

As top-down proteomics becomes more widespread, we can expect to see certain clinical research topics illuminated. One aspect of disease biology that is ripe for top-down analysis is the immune system. The immune system is connected to many human diseases in various ways and consists of a range of cell types, with close to 300 distinct populations in the blood alone [40]. To date, information within the immune system that is associated with disease mechanisms, progression and biomarkers has gone untouched by top-down proteomic approaches. We believe that a search for disease-associated biomarkers using gene- and cell-specific proteomics will substantially benefit from the application of whole-protein analysis to the proteomes of the immune cell populations associated with individual diseases. This idea combines the high analytical precision of top-down proteomics with a layer of precision from individual cell-type resolution.

The analysis of disease-associated immune cell populations (for example, sorted by flow cytometry) using top-down proteomics will have an integral role in shaping the future of clinical proteomic research. In the ideal situation, certain disease studies will begin with top-down proteomic analyses to characterize the intact proteins in each immune cell type in the peripheral blood. Peripheral blood cells can be isolated from patients by the same routine procedure used for obtaining whole blood, serum, and plasma and thus serve as prime candidates for clinical studies of samples directly obtained from patients. The top-down characterization of proteins in immune cell populations will provide proteoform-resolved data that report on the expression profile of proteins within these cell types. The profiles will be readily comparable with ‘healthy’ human cell proteomes by applying the technique to samples isolated from patients without the disease under study. Then, taking a hybrid approach to clinical proteomic research, the discovery phase of top-down proteomics, with its proteoform-resolved data, can then be used to guide the development of proteoform-specific peptides for follow-up, large-scale MRM validation trials.

We believe that the single-cell analysis capabilities of flow cytometry will couple well with proteoform-resolved top-down data. In general, flow cytometry is a common and well-developed procedure for analyzing the cell-by-cell expression of particular proteins using antibodies targeting these proteins. However, without proteoform-resolved information to guide the development and selection of antibodies for monitoring, the information from a flow cytometry experiment could be confusing, with the same protein inference problem that limits the specificity of MRM (Figure 2). In other words, neither technique can accurately describe distinct proteoforms when used alone.

With the pairing of top-down proteomics and flow cytometry, individual proteoforms can be targeted by antibodies that bind only to those distinct forms of the protein. In this manner, the flow cytometry information

will also be proteoform-resolved. Adding this layer of precision to both the MRM and the flow-cytometry follow-up assays will provide a considerable advance toward understanding and diagnosing complex phenotypes, especially when the data are paired with cell-by-cell information from disease-associated immune cells. Ultimately, pairing proteoform-resolved information from top-down proteomics with sensitive and standardized MRM assays and similarly sensitive and standardized targeted flow cytometry assays will provide two promising options for the development of validated clinical diagnostic assays for early disease-phenotype detection.

We hope that in the near future more clinical proteomic pursuits will begin with top-down proteomics discovery that will drive the research with proteoform-resolved precision. One clear benefit of the spread of top-down technology to many laboratories would be a collective increase in the precision of data collection and reporting compared with the prototypic information that bottom-up proteomics is currently providing (Figure 2). Another advantage would be global 'beta testing' of the technique. Inevitably, the more people who use top-down proteomics, the more demand there will be for improved instrumentation and data acquisition (plus the critical software). This type of increased demand will guide the industrial development of top-down platform tools that will benefit the research community directly, by allowing more robust and capable analysis. Thus, a positive feedback loop will commence that will mirror the robust growth cycle experienced by bottom-up technologies over the past 20 years. Having seen the improvements over that time, it is exciting to imagine where top-down technology will be in the near future.

Finally, the overall goal for using top-down proteomics in clinical research is not to take the place of the well-developed, optimized assays that are used in diagnostic laboratories around the world (for example, targeted RNA measurements, DNA sequencing and ELISAs). Rather, the goal is to inform the development and implementation of more-sensitive, more-selective diagnostic tests. By correlating the exact proteoforms with a given disease phenotype, diagnostic laboratories will be able to design assays to perform routine analyses in a proteoform-specific manner.

#### Abbreviations

ELISA, enzyme-linked immunosorbent assay; fALS, familial amyotrophic lateral sclerosis; H-D, hydrogen-deuterium; kDa, kiloDaltons; LC, liquid chromatography; LC-ESI-MS/MS, LC electrospray ionizations tandem MS; LOQ, limit of quantification; MALDI, matrix-assisted laser desorption/ionization; Me, methylation; MRM, multiple reaction monitoring; MS, mass spectrometry; MSI, MS imaging; P, phosphorylation; SELDI-TOF MS, surface-enhanced laser desorption/ionization time of flight MS; SISCAPA, stable isotope standards and capture by antipeptide antibodies; SOD1, superoxide dismutase 1.

#### Competing interests

The authors declare that they have no competing interests.

#### Acknowledgements

Work in the authors' laboratories was supported by the National Institutes of Health under grant R01GM067193 (NLK). JPS is supported by the Northwestern University Comprehensive Transplant Center through grant U19AI063603.

#### Author details

<sup>1</sup>Departments of Chemistry and Molecular Biosciences, the Chemistry of Life Processes Institute, the Proteomics Center of Excellence, 2145 N. Sheridan Dr, Evanston, IL 60208, USA. <sup>2</sup>The Comprehensive Transplant Center, 676 N. St Clair, Chicago IL 60611, USA. <sup>3</sup>The Robert H Lurie Comprehensive Cancer Center, 303 E. Superior, Chicago, IL 60611, USA.

Published: 27 June 2013

#### References

1. Chen R, Snyder M: **Promise of personalized omics to precision medicine.** *Wiley Interdiscip Rev Syst Biol Med* 2013, **5**:73-82.
2. Mann M: **Proteomics for biomedicine: a half-completed journey.** *EMBO Mol Med* 2012, **4**:75-77.
3. Bystrom CE, Salameh W, Reitz R, Clarke NJ: **Plasma renin activity by LC-MS/MS: development of a prototypical clinical assay reveals a subpopulation of human plasma samples with substantial peptidase activity.** *Clin Chem* 2010, **56**:1561-1569.
4. Boersema PJ, Geiger T, Wisniewski JR, Mann M: **Quantification of the N-glycosylated secretome by super-SILAC during breast cancer progression and in human blood samples.** *Mol Cell Proteomics* 2013, **12**:158-171.
5. Wisniewski JR, Dus K, Mann M: **Proteomic workflow for analysis of archival formalin-fixed and paraffin-embedded clinical samples to a depth of 10 000 proteins.** *Proteomics Clin Appl* 2013, **7**:225-233.
6. Zanivan S, Meves A, Behrendt K, Schoof EM, Neilson LJ, Cox J, Tang HR, Kalna G, van Ree JH, van Deursen JM, Trempus CS, Machesky LM, Lindling R, Wickström SA, Fässler R, Mann M: **In vivo SILAC-based proteomics reveals phosphoproteome changes during mouse skin carcinogenesis.** *Cell Rep* 2013, **3**:552-566.
7. Dendrou CA, Bell JI, Fugger L: **Weighing in on autoimmune disease: big data tip the scale.** *Nat Med* 2013, **19**:138-139.
8. Anderson L, Hunter CL: **Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins.** *Mol Cell Proteomics* 2006, **5**:573-588.
9. Compton, PD, Kelleher NL: **Spinning up mass spectrometry for whole protein complexes.** *Nat Methods* 2012, **9**:1065-1066.
10. Anderson NL, Anderson NG: **The human plasma proteome: history, character, and diagnostic prospects.** *Mol Cell Proteomics* 2002, **1**:845-867.
11. Whiteaker JR, Zhao L, Zhang HY, Feng LC, Piening BD, Anderson L, Paulovich AG: **Antibody-based enrichment of peptides on magnetic beads for mass-spectrometry-based quantification of serum biomarkers.** *Anal Biochem* 2007, **362**:44-54.
12. Huillet C, Adrait A, Lebert D, Picard G, Trauchessec M, Louwagie M, Dupuis A, Hittinger L, Ghaleh B, Le Corvoisier P, Jaquinod M, Garin J, Bruley C, Brun V: **Accurate quantification of cardiovascular biomarkers in serum using Protein Standard Absolute Quantification (PSAQ) and selected reaction monitoring.** *Mol Cell Proteomics* 2012, **11**:M111.008235.
13. Keshishian H, Addona T, Burgess M, Kuhn E, Carr SA: **Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution.** *Mol Cell Proteomics* 2007, **6**:2212-2229.
14. Apple FS: **A new season for cardiac troponin assays: it's time to keep a scorecard.** *Clin Chem* 2009, **55**:1303-1306.
15. Lepor H, Cheli CD, Thiel RP, Taneja SS, Laze J, Chan DW, Sokoll LJ, Mangold L, Partin AW: **Clinical evaluation of a novel method for the measurement of prostate-specific antigen, AccuPSA™, as a predictor of 5-year biochemical recurrence-free survival after radical prostatectomy: results of a pilot study.** *BJU Int* 2012, **109**:1770-1775.
16. Carr SA, Anderson L: **Protein quantitation through targeted mass spectrometry: the way out of biomarker purgatory?** *Clinical Chem* 2008, **54**:1749-1752.
17. Jia Y, Wu T, Jelinek CA, Bielekova B, Chang L, Newsome S, Gnanapavan S, Giovannoni G, Chen D, Calabresi PA, Nath A, Cotter RJ: **Development of protein biomarkers in cerebrospinal fluid for secondary progressive**

- multiple sclerosis using selected reaction monitoring mass spectrometry (SRM-MS). *Clin Proteomics* 2012, **9**:9.
18. Huttenhain R, Soste M, Selevsek N, Rost H, Sethi A, Carapito C, Farrah T, Deutsch EW, Kusebauch U, Moritz RL, Nimeus-Malmstrom E, Rinner O, Aebersold R: **Reproducible quantification of cancer-associated proteins in body fluids using targeted proteomics.** *Sci Transl Med* 2012, **4**:142ra194.
  19. Musselman CA, Lalonde ME, Côté J, Kutateladze TG: **Perceiving the epigenetic landscape through histone readers.** *Nat Struct Mol Biol* 2012, **19**:1218-1227.
  20. Smith LM, Kelleher NL; Consortium for Top Down Proteomics: **Proteoform: a single term describing protein complexity.** *Nat Methods* 2013, **10**:186-187.
  21. O'Farrell PH: **High resolution two-dimensional electrophoresis of proteins.** *J Biol Chem* 1975, **250**:4007-4021.
  22. Unlu M, Morgan ME, Minden JS: **Difference gel electrophoresis: a single gel method for detecting changes in protein extracts.** *Electrophoresis* 1997, **18**:2071-2077.
  23. Bloom GC, Eschrich S, Zhou JX, Coppola D, Yeatman TJ: **Elucidation of a protein signature discriminating six common types of adenocarcinoma.** *Int J Cancer* 2007, **120**:769-775.
  24. Zhou G, Li H, DeCamp D, Chen S, Shu H, Gong Y, Flaig M, Gillespie JW, Hu N, Taylor PR, Emmert-Buck MR, Liotta LA, Petricoin EF 3rd, Zhao Y: **2D Differential in-gel electrophoresis for the identification of esophageal scans cell cancer-specific protein markers.** *Mol Cell Proteomics* 2002, **1**:117-124.
  25. Gygi SP, Corthals GL, Zhang Y, Rochon Y, Aebersold R: **Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology.** *Proc Natl Acad Sci U S A* 2000, **97**:9390-9395.
  26. Ong SE, Pandey A: **An evaluation of the use of two-dimensional gel electrophoresis in proteomics.** *Biomol Eng* 2001, **18**:195-205.
  27. Kellie JF, Catherman AD, Durbin KR, Tran JC, Tipton JD, Norris JL, Witkowski CE 2nd, Thomas PM, Kelleher NL: **Robust analysis of the yeast proteome under 50 kDa by molecular-mass-based fractionation and top-down mass spectrometry.** *Anal Chem* 2012, **84**:209-215.
  28. Tran JC, Zamdborg L, Ahlf DR, Lee JE, Catherman AD, Durbin KR, Tipton JD, Vellaichamy A, Kellie JF, Li M, Wu C, Sweet SM, Early BP, Siuti N, LeDuc RD, Compton PD, Thomas PM, Kelleher NL: **Mapping intact protein isoforms in discovery mode using top-down proteomics.** *Nature* 2011, **480**:254-258.
  29. Auffray C, Caulfield T, Khoury MJ, Lupski JR, Schwab M, Veenstra T: **2012 Highlights in translational 'omics.** *Genome Med* 2013, **5**:10.
  30. Wu C, Tran JC, Zamdborg L, Durbin KR, Li M, Ahlf DR, Early BP, Thomas PM, Sweedler JV, Kelleher NL: **A protease for 'middle-down' proteomics.** *Nat Methods* 2012, **9**:822-824.
  31. Petricoin E 3rd, Liotta LA: **The vision for a new diagnostic paradigm.** *Clin Chem* 2003, **49**:1276-1278.
  32. Oppenheimer SR, Mi D, Sanders ME, Caprioli RM: **Molecular analysis of tumor margins by MALDI mass spectrometry in renal carcinoma.** *J Proteome Res* 2010, **9**:2182-2190.
  33. Chamot-Rooke J, Mikaty G, Malosse C, Soyler M, Dumont A, Gault J, Imhaus AF, Martin P, Trellet M, Clary G, Chafey P, Camoin L, Nilges M, Nassif X, Dumenil G: **Posttranslational modification of pili upon cell contact triggers *N. meningitidis* dissemination.** *Science* 2011, **331**:778-782.
  34. Dong X, Sumandea CA, Chen YC, Garcia-Cazarin ML, Zhang J, Balke CW, Sumandea MP, Ge Y: **Augmented phosphorylation of cardiac troponin I in hypertensive heart failure.** *J Biol Chem* 2012, **287**:848-857.
  35. Mazur MT, Cardasis HL, Spellman DS, Liaw A, Yates NA, Hendrickson RC: **Quantitative analysis of intact apolipoproteins in human HDL by top-down differential mass spectrometry.** *Proc Natl Acad Sci U S A* 2010, **107**:7728-7733.
  36. Molnar KS, Karabacak NM, Johnson JL, Wang Q, Tiwari A, Hayward LJ, Coales SJ, Hamuro Y, Agar JN: **A common property of amyotrophic lateral sclerosis-associated variants: destabilization of the copper/zinc superoxide dismutase electrostatic loop.** *J Biol Chem* 2009, **284**:30965-30973.
  37. Ahlf DR, Compton PD, Tran JC, Early BP, Thomas PM, Kelleher NL: **Evaluation of the compact high-field orbitrap for top-down proteomics of human cells.** *J Proteome Res* 2012, **11**:4308-4314.
  38. Tian Z, Tolić N, Zhao R, Moore RJ, Hengel SM, Robinson EW, Stenoien DL, Wu S, Smith RD, Paša-Tolić L: **Enhanced top-down characterization of histone post-translational modifications.** *Genome Biol* 2012, **13**:R86.
  39. Consortium for Top Down Proteomics [<http://www.topdownproteomics.org>]
  40. Bendall SC, Simonds EF, Qiu P, Amir ED, Krutzik PO, Finck R, Bruggner RV, Melamed R, Trejo A, Ornatsky OI, Balderas RS, Plevritis SK, Sachs K, Pe'er D, Tanner SD, Nolan GP: **Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum.** *Science* 2011, **332**:687-696.
  41. Borges CR, Oran PE, Buddi S, Jarvis JW, Schaab MR, Rehder DS, Rogers SP, Taylor T, Nelson RW: **Building multidimensional biomarker views of type 2 diabetes on the basis of protein microheterogeneity.** *Clin Chem* 2011, **57**:719-728.

doi:10.1186/gm457

Cite this article as: Savaryn JP, et al.: The emergence of top-down proteomics in clinical research. *Genome Medicine* 2013, **5**:53.