Published in final edited form as:

Nat Biotechnol. 2016 October 11; 34(10): 1033–1034. doi:10.1038/nbt.3695.

Quantifying the human proteome

Amanda G. Paulovich and Jeffrey R. Whiteaker

Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109

Substantial advances in mass spectrometry (MS)-based proteomic profiling have enabled broad coverage of the expressed and modified proteome. However, confirmation and validation efforts to make biological findings from these “discovery” experiments actionable (in clinical or biological research) largely depend on conventional immunoassay platforms (e.g. Western blotting) and commercially offered antibodies, up to half of which fail to work as claimed¹. Targeted MS techniques, like selected or multiple reaction monitoring (SRM, MRM), are bridging the gap between discovery and validation by providing quantitative assays that can be standardized and highly characterized by conventional fit-for-purpose method validation approaches. In a recent issue of *Cell*, Kusebauch *et al*² describe the expansion of SRMATlas, a rich resource of *targeted* MS coordinates (i.e. peptides, fragments, retention times), covering the basic human proteome. This knowledge base will facilitate the development of targeted MS assays by providing a starting point to quantify the basic human proteome with high specificity and analytical rigor.

Even in the era of genomics, quantifying proteins is important. Not all genomic alterations are translated into alterations at the protein level, and genomic profiles are not reliable predictors of protein levels or activity^{3–5}. Because proteins execute physiological functions, most often proteins are the targets of therapeutics. Thus, the ability to quantify proteins and their modifications is critical across biomedical research, spanning basic to clinical studies.

Conventional methods widely employed to quantify proteins (e.g. Western blotting, immunohistochemistry, ELISA) rely on antibodies to capture target proteins from complex biological samples. Once captured, the concentration of the protein analyte is inferred from a surrogate signal arising from a molecular tag (e.g. fluorescent, enzymatic, nucleic acid, or mass tag) at the other end of the antibody. While these approaches have enabled many discoveries, they are unfortunately prone to several drawbacks. Most assays based on conventional platforms are semi-quantitative at best, difficult-to-impossible to multiplex, and are susceptible to interferences affecting the accuracy of quantification⁶. Indeed, it is extremely difficult to identify antibodies that are monospecific when applied to a range of biological specimens^{1, 6}, contributing to high costs and time required for assay development.

Because of these analytical limitations, a NextGen protein quantification platform based on targeted MS (SRM, MRM) is gaining acceptance. Unlike untargeted “shotgun” modes of MS widely deployed in discovery proteomic experiments, S/MRM is performed on specialized instruments that can be “tuned” to detect specific analytes of interest in complex biospecimens, based on their mass-to-charge ratio. In contrast to global isobaric labeling strategies, S/MRM targets signature peptides that are unique to a protein of interest, and a

synthetic or recombinant, stable isotope-labeled version of each target peptide or protein is spiked into biospecimens at a known concentration to enable precise, relative quantification.

Because the mass spectrometer is used to detect and quantify the *actual analyte of interest* (i.e. not a surrogate signal) relative to the spiked-in standard, specificity is in general much improved over conventional immunoassays, even when S/MRM is coupled to antibody-based analyte enrichment. S/MRM assays can be highly multiplexed, are precise across 3+ orders of magnitude linear range, and can be standardized and harmonized across laboratories, even on an international stage⁷. Indeed, mass spectrometers used to perform S/MRM assays are widely available in clinical and academic core laboratories, where they have been used for decades for quantifying small molecules (e.g., metabolites), and thus the infrastructure is in place for widespread adoption.

The SRMatlas described by Kusebauch *et al* provides a valuable knowledge base (i.e. coordinates) that can be used as the starting point for quantitative assay generation. A critical first step to S/MRM assay development is selection of proteotypic peptides that are unique to the protein of interest, reproducibly released by proteolysis, can be readily synthesized, and have favorable biophysical properties for chromatography and MS detection⁸. To assemble a proteome-wide resource of SRM coordinates, Kusebauch *et al* identified proteotypic peptides to all annotated proteins (including known isoforms, high frequency SNPs, N-glycosylated modifications, and pre/pro hormones) from existing empirical mass spectrometry datasets and prediction algorithms based on peptide chemical and physical properties (for those proteins with no empirical evidence). SRM coordinates (e.g. peptide m/z ratio, fragmentation data, relative retention time, and optimal instrument parameters) were obtained using synthetic peptides for the target sequences analyzed on two vendor instruments. In total, peptide sequences were identified and SRM coordinates characterized for an impressive 99.7% of the human proteome defined by UniProtKB/Swiss-Prot, and three or more peptides were selected for 95.4% of proteins. This impressive breadth of coverage provides a rich starting point for selecting peptide analytes for assay generation. The utility of the resource was demonstrated by developing SRM assays to profile relative changes in abundance for proteins in two proof-of-concept perturbation experiments: the effect of inhibition of cholesterol synthesis and the effect of docetaxel on prostate cancer cell lines. In each case, the results demonstrate the effectiveness of the resource to facilitate development of multiplexed SRM assays to confirm hypotheses.

SRMatlas is a first step in the process of developing targeted MS assays to all human proteins by providing the public with targeted MS coordinates. To move these SRM coordinates closer to the hands of biologists and clinical labs, real-world (fit-for-purpose) conditions must be utilized. Specifically, to be used for rigorous quantification, peptide coordinates from SRMatlas need to be converted into well characterized *assays* (Figure 1). SRM assay coordinates alone do not guarantee detection of the peptide in the sample of interest (indeed most selected peptides in SRMatlas were not observed in empirical data), especially for complex matrices like plasma or cell lysates, where many of the target peptides will require enrichment to enable detection of endogenous analyte. Likewise, interferences can be encountered when applying the SRM coordinates to measurement in complex biological matrices.

Fit-for-purpose guidelines for targeted MS-based assay characterization, analogous to well-established protocols in other areas of quantitative measurement, have been established by the community⁹. To have confidence in assay results, basic performance metrics of the assays must be characterized, such as imprecision, repeatability, reproducibility, bias, linearity, limit of quantification, matrix effects/selectivity, and analyte stability¹⁰. An open-source repository of well-characterized (i.e., fit-for-purpose) targeted MS assays is available¹¹ (assays.cancer.gov), as is open-source software for data analysis¹². Merging these tools together will undoubtedly expand the toolbox for all biological researchers, improving the way we study, diagnose, and treat disease. The addition of post-translational modifications beyond N-glycosylation (e.g., phosphorylation) will also greatly broaden the utility of the database.

References

1. Baker M. Reproducibility crisis: Blame it on the antibodies. *Nature*. 2015; 521:274–276. [PubMed: 25993940]
2. Kusebauch U, et al. Human SRMATlas: A Resource of Targeted Assays to Quantify the Complete Human Proteome. *Cell*. 2016; 166:766–778. [PubMed: 27453469]
3. Zhang H, et al. Integrated Proteogenomic Characterization of Human High-Grade Serous Ovarian Cancer. *Cell*. 2016; 166:755–765. [PubMed: 27372738]
4. Mertins P, et al. Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature*. 2016; 534:55–62. [PubMed: 27251275]
5. Zhang B, et al. Proteogenomic characterization of human colon and rectal cancer. *Nature*. 2014; 513:382–387. [PubMed: 25043054]
6. Hoofnagle AN, Wener MH. The fundamental flaws of immunoassays and potential solutions using tandem mass spectrometry. *J Immunol Methods*. 2009; 347:3–11. [PubMed: 19538965]
7. Kennedy JJ, et al. Demonstrating the feasibility of large-scale development of standardized assays to quantify human proteins. *Nat Methods*. 2014; 11:149–155. [PubMed: 24317253]
8. Hoofnagle AN, et al. Recommendations for the Generation, Quantification, Storage, and Handling of Peptides Used for Mass Spectrometry-Based Assays. *Clin Chem*. 2016; 62:48–69. [PubMed: 26719571]
9. Carr SA, et al. Targeted peptide measurements in biology and medicine: best practices for mass spectrometry-based assay development using a fit-for-purpose approach. *Mol Cell Proteomics*. 2014; 13:907–917. [PubMed: 24443746]
10. Grant RP, Hoofnagle AN. From lost in translation to paradise found: enabling protein biomarker method transfer by mass spectrometry. *Clin Chem*. 2014; 60:941–944. [PubMed: 24812416]
11. Whiteaker JR, et al. CPTAC Assay Portal: a repository of targeted proteomic assays. *Nat Methods*. 2014; 11:703–704. [PubMed: 24972168]
12. Sharma V, et al. Panorama: a targeted proteomics knowledge base. *J Proteome Res*. 2014; 13:4205–4210. [PubMed: 25102069]

Figure 1. Workflow for development and implementation of quantitative proteomic assays
Developing quantitative MS assays to protein targets requires identification of suitable peptides and configuration of instrument and sample preparation parameters for optimal detection. Once identified, these coordinates are converted into quantitative assays by fit-for-purpose characterization and validation. Quantitative rigor is maintained through the use of QA/QC standards during assay deployment.