HHS Public Access

Author manuscript

J Am Soc Mass Spectrom. Author manuscript; available in PMC 2021 July 25.

Published in final edited form as:

J Am Soc Mass Spectrom. 2021 June 02; 32(6): 1278–1294. doi:10.1021/jasms.1c00099.

Novel Strategies to Address the Challenges in Top-Down Proteomics

Jake A. Melby[#], **David S. Roberts**[#], **Eli J. Larson**

Department of Chemistry, University of Wisconsin–Madison, Madison, Wisconsin 53706, United States;

Kyle A. Brown

Department of Chemistry, University of Wisconsin–Madison, Madison, Wisconsin 53706, United States; Department of Surgery, University of Wisconsin–Madison, Madison, Wisconsin 53705, United States

Elizabeth F. Bayne, **Song Jin**

Department of Chemistry, University of Wisconsin–Madison, Madison, Wisconsin 53706, United States;

Ying Ge

Department of Chemistry, University of Wisconsin–Madison, Madison, Wisconsin 53706, United States; Department of Cell and Regenerative Biology and Human Proteomics Program, University of Wisconsin–Madison, Madison, Wisconsin 53705, United States;

[#] These authors contributed equally to this work.

Abstract

Top-down mass spectrometry (MS)-based proteomics is a powerful technology for comprehensively characterizing proteoforms to decipher post-translational modifications (PTMs) together with genetic variations and alternative splicing isoforms toward a proteome-wide understanding of protein functions. In the past decade, top-down proteomics has experienced rapid growth benefiting from groundbreaking technological advances, which have begun to reveal the potential of top-down proteomics for understanding basic biological functions, unraveling disease mechanisms, and discovering new biomarkers. However, many challenges remain to be comprehensively addressed. In this Account & Perspective, we discuss the major challenges currently facing the top-down proteomics field, particularly in protein solubility, proteome dynamic range, proteome complexity, data analysis, proteoform–function relationship, and

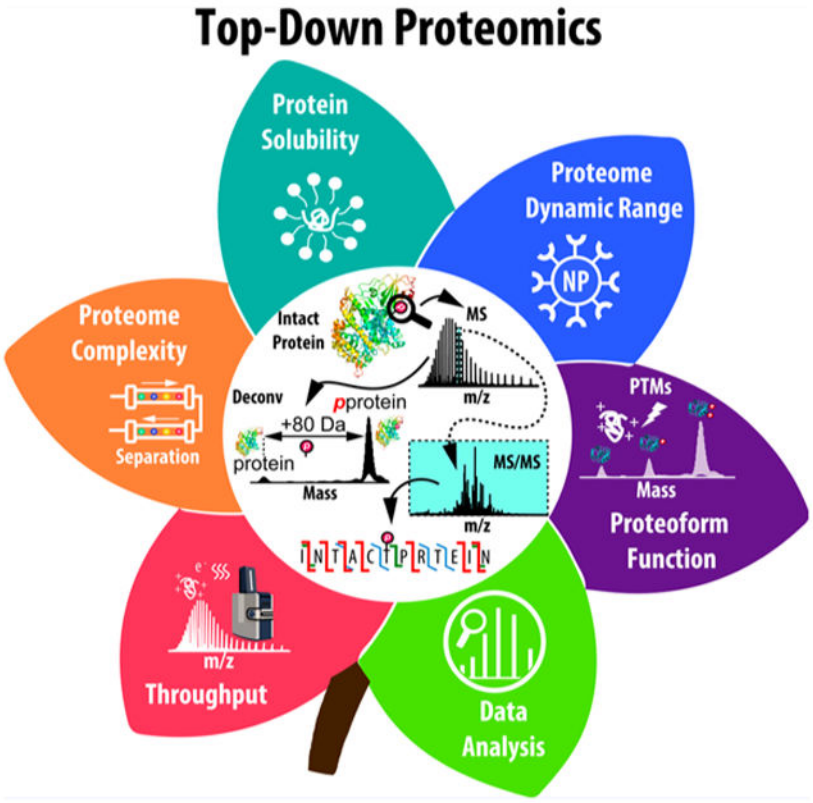
Corresponding Author: **Ying Ge** – Department of Chemistry, University of Wisconsin–Madison, Madison, Wisconsin 53706, United States; Department of Cell and Regenerative Biology and Human Proteomics Program, University of Wisconsin–Madison, Madison, Wisconsin 53705, United States; Phone: 608-265-4744; ying.ge@wisc.edu.

Complete contact information is available at: https://pubs.acs.org/10.1021/jasms.1c00099

The authors declare the following competing financial interest(s): The University of Wisconsin, Madison, has filed a provisional patent application P180335US01, US serial number 62/682,027 (June 7, 2018) on the basis of the photocleavable surfactant work. Y.G., S.J., and K.A.B. are named as inventors on the provisional patent application. The University of Wisconsin, Madison, has filed a provisional patent application serial number 62/949,869 (December 18, 2019) on the basis of the nanoparticles for cardiac troponin enrichment. Y.G., S.J., and D.S.R. are named as the inventors on the provisional patent application. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

analytical throughput for precision medicine. We specifically review the major technology developments addressing these challenges with an emphasis on our research group's efforts, including the development of top-down MS-compatible surfactants for protein solubilization, functionalized nanoparticles for the enrichment of low-abundance proteoforms, strategies for multidimensional chromatography separation of proteins, and a new comprehensive user-friendly software package for top-down proteomics. We have also made efforts to connect proteoforms with biological functions and provide our visions on what the future holds for top-down proteomics.

Graphical Abstract



1. INTRODUCTION

In the post-genomics era, a comprehensive analysis of "proteoforms"[1,2]—all protein products that arise from a single gene due to genetic variations, alternative splicing, and post-translational modifications (PTMs)—is essential for understanding biological systems at a functional level and for dissecting complex molecular systems with consideration of individual variability toward precision medicine.[3–7] The well-established "bottom-up" proteomics[8] analyzes peptides from protein digests and has become invaluable for the high-throughput identification and quantification of proteins as well as exploration of proteome structure and function,[9,10] but it is suboptimal for characterizing PTMs and sequence variants due to the "peptide-to-protein" inference problem.[2,11] In contrast, top-down mass

spectrometry (MS)-based proteomics analyzes intact proteins, thus making it the premier technology for comprehensive analysis of proteoforms to decipher PTM codes together with genetic mutations and alternative splicing.[1–7]

Top-down MS was first coined by McLafferty and co-workers in 1999,[12] and the early top-down MS studies generally took a targeted approach focusing on offline purified single proteins or simple protein mixtures due to the inherent limitations of intact protein analysis.[12–17] Top-down proteomics, pioneered by Kelleher and co-workers, includes the front-end fractionation of intact proteins, high-resolution MS,[18] and back-end informatics,[19] making it possible to identify proteins from complex mixtures.[6,20,21] In the past decade, top-down proteomics has experienced rapid growth due to many technological advances,[22–32] thereby enabling the applications of top-down proteomics to complex protein mixtures and for understanding basic biological functions, unraveling disease mechanisms, and discovering new biomarkers.[33–45] However, many challenges still exist for top-down MS-based proteomics, especially protein solubility, dynamic range, proteome complexity, data analysis, proteoform–function relationships, and analytical throughput of top-down proteomics for precision medicine (Figure 1). Previous reviews on top-down proteomics have already described various aspects of top-down proteomics such as technological advancements and biomedical applications, among others.[4–6,17,21,46–53] In this Account & Perspective, we will focus on the most pressing challenges currently facing the top-down proteomics field. We illustrate how recent and future technological innovation could help address these challenges to advance top-down proteomics toward the mainstream. Given the focus of this special issue, we specifically detail the major strides our research group has made to address these current challenges, and we will provide our vision of what the future holds for top-down proteomics.

2. ADDRESSING THE CHALLENGE OF PROTEIN SOLUBILITY

Protein solubility represents a major challenge in top-down MS-based proteomics, especially for membrane proteins[54,55] and extracellular matrix (ECM) proteins,[56,57] which are difficult to extract and solubilize from tissues and cells but play significant biological roles.[5,48,58–60] Despite the success of several important top-down proteomics studies of membrane proteins,[25,54,61–65] the time-consuming preparations and labor-intensive workups prior to MS analysis leave significant room for technological improvement in this area. Moreover, early work generally focused on targeting a specific protein of interest or the mitochondrial membrane subproteome.[25,54,61–65] Surfactants (also known as detergents) are typically used in protein extraction buffers to effectively solubilize proteins from cells and tissues.[66,67] However, conventional surfactants that are capable of solubilizing and denaturing proteins with high efficiency, such as sodium dodecyl sulfate (SDS), are not compatible with downstream MS analysis due to the signal suppression of proteins and thus require extensive cleanup that can result in protein loss and irreproducibility.[66,68] Although non-ionic surfactants, such as dodecyl β-D-maltoside (DDM) and octyl β-D-glucopyranoside (OG), are compatible with MS when used at lower concentrations,[68,69] they are relatively mild surfactants with limited solubilization ability.[67] Acid-cleavable surfactants have been previously developed for bottom-up proteomics, such as RapiGest (RG)[70] (also known as ALS),[20] ProteaseMax,[71] and MS-compatible slowly degradable surfactant (MaSDeS),[72] to

enable protein solubilization and improve in-gel or in-solution digestion efficiency. However, none of these surfactants are directly compatible with top-down MS-based proteomics.[59]

To address this problem, it is essential to develop novel top-down MS-compatible surfactants that can quickly degrade into innocuous non-surfactant byproducts, facilitating easy surfactant removal before MS analysis. We recently identified a novel photocleavable surfactant, 4-hexylphenylazosulfonate ("Azo"), which effectively solubilizes proteins with similar performance to SDS but can rapidly degrade upon exposure to UV irradiation, thus enabling top-down proteomics (Figure 2A–D).[59] A systematic comparison of Azo with commonly used surfactants (i.e., SDS, RapiGest, ProteaseMax, MaSDeS, DDM, OG, digitonin) was performed to evaluate their ability to solubilize proteins from the insoluble cardiac tissue pellets, which confirmed that Azo is the only surfactant that can effectively solubilize proteins while maintaining top-down MS compatibility (Figure 2E–G). Moreover, Azo can greatly enhance access to proteins from many cellular components including the nucleus, mitochondria, plasma membrane, endoplasmic reticulum, cytoplasm, cytoskeleton, and extracellular region using top-down proteomics.[59] Importantly, the use of Azo enables the facile solubilization of membrane proteins from human cardiac tissue for comprehensive characterization and localization of PTMs.[59] For example, the PTMs of phospholamban, a well-known cardiac regulatory protein, were characterized and a palmitoylation modification to cysteine 36 within the transmembrane region was localized.[59] Moreover, 46 subunits of the electron transport chain and 51 proteins with transmembrane domains, including all of the subunits of the ATP synthase complex, were confidently identified with high mass accuracy directly from cardiac tissue. Since the synthesis of Azo is very simple and Azo can be used as a general SDS replacement in SDS-PAGE, this surfactant has a broader impact beyond MS-based proteomics.[59]

We further showed that Azo is a versatile proteomics surfactant compatible with both top-down and bottom-up proteomics workflows.[59,73] Rapid enzymatic digestion by trypsin (~1 h) was observed due to the highly denaturing nature of Azo facilitating the opportunity for high-throughput proteomics. Additionally, an Azo-aided bottom-up method was developed to study integral membrane proteins, providing over a 6-fold increase in the number of identified integral membrane proteins when compared to conventional sample preparation strategies.[73] The unique ability of Azo to facilitate both top-down and bottom-up proteomics, including the analysis of membrane proteins, makes it a valuable tool for proteomics sample preparation (Figure 2E). Recently, we expanded the use of Azo to extracellular matrix (ECM) proteomics.[60] The ECM provides an architectural meshwork for surrounding cells, and the dysregulation of ECM proteins directly contributes to various pathologies and progression in human diseases such as cancers and cardiovascular diseases.[56,57,74,75] However, the biochemical/proteomic characterization of ECM proteins is extremely challenging due to the nearly insoluble nature of the ECM.[35,36] To address these challenges, we established a new decellularization/extraction method enabled by Azo[59,73] for the efficient enrichment of ECM proteins, minimizing the sample cleanup before LC-MS/MS analysis.[60] Overall, we have shown that Azo could enhance proteomic analysis through simple protein sample preparation and streamlined proteomic workflows, so we anticipate Azo will be extended to other proteomics applications. Nonetheless, it should be

noted that certain protein classes are currently more amenable to acid/base extractions, such as acid-soluble histones.[76,77]

MS-based proteomics experiments are commonly performed under denaturing conditions to obtain the primary structural analysis of proteins.[78] In contrast, native MS performed under non-denaturing conditions seeks to preserve whole protein structures and their complexes, which enables the study of protein–protein interactions and ligand/drug binding.[79–81] The recent integration of native MS with top-down proteomics ("native top-down proteomics") allowed the structural characterization of macromolecular complexes together with comprehensive sequence mapping and proteoform characterization.[26,82–86] In particular, the rapidly developing native membrane proteomics has huge potential for unlocking structural and functional information on these important proteins.[26,69] Solubilizing membrane proteins for native MS requires keeping proteins stable and as "native" as possible. Toward this end, modular surfactants were designed and employed by Urner et al. for improved purification and native MS of G-protein-coupled receptors.[87] In an effort to further increase the "native" environment, Chorev et al. developed a method to analyze protein complexes ejected directly from native *E. coli* outer membranes.[88] Finally, one of the most significant advancements was the development of "nativeomics", wherein native MS was used to precisely identify ligands bound to membrane protein assemblies, providing insight into their functional role.[82] Conceivably, the continued developments in the area of native membrane proteomics, including new membrane mimics and purification strategies, will be needed for analyzing endogenous membrane proteins/protein complexes (particularly as it shifts toward studying eukaryotic membrane proteins).[84]

3. ADDRESSING THE CHALLENGE IN THE HIGH DYNAMIC RANGE OF THE HUMAN PROTEOME

The high dynamic range of the proteome (10^8–10^{12}) makes detection of low-abundance proteoforms extremely difficult.[89] Front-end protein enrichment strategies are therefore required to specifically capture and enrich low-abundance proteoforms from the complex biological milieu before MS analysis.[90,91] Antibodies have historically been the dominant affinity reagents for protein capture and quantification in biomedical research.[91–93] Antibody-based affinity purification (AP) has been a preferred approach in the targeted analysis of intact proteins and protein complexes.[4,94,95] For example, AP has been utilized to couple with top-down mass spectrometry (MS) for the purification of cardiac troponin from both human and animal heart tissues to identify disease-associated proteoform alterations.[34,35] Nonetheless, antibody-based technologies have significant limitations, including the high cost of antibody production, difficulty in the generation of antibodies with high specificity, limited availability of high-quality antibodies, batch-to-batch variability of the antibodies, and relatively low stability.[96–99] Hence, there is an urgent need to develop next-generation affinity reagents.[92,100,101]

To address this challenge, we have developed surface-functionalized multivalent superparamagnetic nanoparticles (NPs) as a general affinity platform to capture and enrich low-abundance proteoforms with high specificity, provided the NPs are functionalized with a

suitable affinity reagent.[102–104] Functionalized NPs have several unique advantages for protein enrichment: (a) NPs possess comparable size and diffusion kinetics to antibodies, allowing for better penetration through complex protein mixtures and promoting effective capture of low-abundance proteins; (b) NPs can be produced with simple, fast, and scalable synthesis, in contrast to the long time required to produce antibodies; (c) NPs have high surface area to volume ratios due to their nanoscale size (5–20 nm) which translates into high surface ligand density and increased protein interactions; (d) NPs can have multiple flexible binding sites allowing multivalent ligand interactions which can increase the overall binding affinity and ensure the efficient capture of lower-abundance proteins.[102–104]

The first generation of functionalized superparamagnetic iron oxide (Fe_3O_4) NPs was developed to universally enrich phosphoproteins regardless of the specific phosphorylation site (Ser, Thr, or Tyr phosphorylation).[102] Subsequently, an integrated top-down proteomics workflow has been established for comprehensive phosphoprotein characterization enabled by coupling specific intact phosphoprotein enrichment by cobalt ferrite ($CoFe_2O_4$) NPs with online LC-MS/MS analysis.[104] Roberts et al. have further developed a reproducible large-scale synthesis of surface silanized Fe_3O_4 NPs as an enabling nanoproteomics platform and demonstrated its applications for the highly specific enrichment of the human heart phosphoproteome.[103] In this work, a reproducible and large-scale (200 mg) synthesis of surface-silanized magnetite (Fe_3O_4) NPs for biological applications was developed. As an alternative to antibodies, these NPs were surface functionalized with a dinuclear Zn(II)-dipicolylamine (Zn-DPA) complex as a phosphoprotein affinity ligand to specifically enrich intact phosphoproteins.[103] By further integrating with a LC-MS/MS workflow suitable for intact protein analysis, this highly scalable platform was used for the global enrichment, identification, and characterization of endogenous phosphoproteins from human cardiac tissues.

Recently, we have exploited this nanoproteomics platform to enrich cardiac troponin I (cTnI), the gold-standard biomarker for acute myocardial infarction (also known as a heart attack), from human serum with high specificity and sensitivity.[30] The human blood proteome is enormously complex, possessing a huge dynamic range (10^{12}) dominated by the presence of numerous highly abundant proteins, such as human serum albumin (HSA), making the analysis of low-abundance proteins extraordinarily challenging.[89,90] Therefore, a nanoproteomics strategy using peptide-functionalized NPs was developed to specifically capture and enrich low-abundance cTnI (<1 ng/mL) from human serum with high specificity and reproducibility (Figure 3A). By coupling this nanoproteomics strategy with top-down MS, diverse cTnI proteoforms were enriched and characterized with high reproducibility (Figure 3B).[30] Importantly, these carefully designed peptide-functionalized NPs preserved all endogenous cTnI proteoform distributions and faithfully retained the endogenous cTnI PTM profiles at every step of the enrichment process without introduction of artifactual modifications (Figure 3C). The ability to accurately quantify multiple proteoforms concurrently marks a unique advantage that the top-down MS approach provides over bottom-up MS and conventional immuno-based techniques.[4,98,99,105,106] Moreover, these NPs can directly capture and enrich cTnI from human serum with high sensitivity (LOD of ~0.06 ng/mL) and high specificity, showing near-complete resistance to nonspecific binding of the highly abundant HSA. These NPs not only outperformed conventional monoclonal

antibody platforms for serum cTnI enrichment but also faithfully and holistically preserve all endogenous cTnI proteoforms (Figure 3C). We are currently applying this nanoproteomics platform to human patient blood samples to establish the relationships between the cTnI proteoforms and underlying disease etiology for precision medicine.

This top-down nanoproteomics platform has a unique combination of specificity, scalability, reproducibility, and efficiency for the capture and enrichment of low-abundance proteoforms and can be widely applied to many other biological systems of interest to enable proteoform-resolved analysis of low-abundance proteins in general. Future applications synergizing the advances in aptamer[107,108] developments with nanotechnologies[109,110] may enable the application of alternative affinity reagents to new nanoparticle platforms for robust intact protein analysis by top-down MS. Nonetheless, advances in *in silico* methods for affinity reagent design and virtual screening can lead to the further development of affinity reagents such as affimers and improvements in phage display can help address these challenges.[30] Moreover, it should be noted that enrichment strategies may introduce systemic bias in protein sampling and requires careful evaluation of the enrichment method and data analysis to ensure quantitative reproducibility.[48,111] We anticipate continuous technology advancements in the coming years for the enrichment of the low-abundance proteoform and the subsequent quantitative top-down MS analysis.

4. ADDRESSING THE CHALLENGES IN PROTEOME COMPLEXITY

The complexity of the proteome remains a significant challenge to top-down proteomics, which necessitates the separation of intact proteins prior to the MS analysis.[22,50,112,113] As MS instruments capable of top-down analysis have become more widely available, the demand for the further development of intact protein separations has grown.[5] Despite the central roles played by liquid chromatography (LC) in the development of bottom-up proteomics,[114] chromatographic separation of proteins has been considered very challenging and underdeveloped.[5,112] However, recent improvements to liquid chromatographs,[115,116] advances in coreshell and monolithic stationary phases,[117,118] and the development of new column chemistry and selectivities[119] have greatly improved the separation resolution and efficiency of intact protein mixtures. In this section, we will discuss size-based chromatography separation, non-denaturing separations, ion mobility spectrometry, and the creative integration of several orthogonal separation modes in multidimensional liquid chromatography (MDLC), which have shown great potentials to enhance top-down proteomics (Figure 4A).

The broad range of protein molecular weights within a proteome can exceed 5 orders of magnitude,[120] making global top-down coverage of the proteome difficult. This challenge is most acute for large proteins because, as protein size increases, the signal-to-noise ratio decreases due to an increased number of charge states and isotopomers.[121] To address this problem, Tran and Doucette developed gel-eluted liquid fraction entrapment electrophoresis (GELFrEE) as an offline technique to bin proteins into size-selected fractions.[122] Another method, termed Passively Eluting Proteins from Polyacrylamide gels as Intact species for MS (PEPPI-MS), has been developed for size-based proteome fractionation prior to LC-MS/MS.[24] The PEPPI-MS method enables recovery of separated protein bands from SDS-

PAGE gels using equipment common to biological and biochemical laboratories. Recently, our group developed serial size exclusion chromatography (sSEC) for size-based proteome fractionation followed by RPLC enabling the detection of proteoforms up to 223 kDa using a Q-TOF mass spectrometer (Figure 4B).[29] In contrast to conventional SEC methods which use only one separation column,[123] the sSEC method links multiple columns of differing porosity to more effectively separate smaller proteoforms from large proteoforms, facilitating the characterization of high molecular weight proteins.[29] Proteins in excess of 100 kDa that were previously only detectable using ultrahigh-resolution MS, such as cardiac myosin binding protein C (140 kDa)[124] and myosin heavy chain (223 kDa),[125] could be characterized using online RPLC-MS/MS using a Q-TOF mass spectrometer.[29] We further extended this method for characterization of large cardiac proteoforms, such as swine aconitate hydratase (82.5 kDa), using a 12T FTICR-MS without front-end separation.[126] Unlike GELFrEE or PEPPI-MS, sSEC can fractionate proteins in both offline and online modes due to MS compatibility of the mobile phases used. This flexibility allows much greater sample throughput than other size-based separations.

Mobile phase compatibility with ESI is also essential to the development of new separation selectivities.[127] Protein separation techniques such as ion-exchange chromatography (IEX) and hydrophobic interaction chromatography (HIC)[128] rely on high concentrations of buffer salt in the mobile phase to stabilize the tertiary structure of the protein and influence the selectivity of the separation, but the conventional buffers such as phosphate, sulfate, or citrate salts are non-volatile.[129,130] Recent work has allowed online coupling of both IEX and HIC with MS using the volatile buffer ammonium acetate.[130,131] In collaboration with Andy Alpert from PolyLC, we developed a series of more hydrophobic HIC materials that can separate proteins using MS-compatible concentrations of ammonium acetate preserving protein structure rather than promoting retention. We have shown for the first time the MS-compatible HIC method, separating a complex *E. coli* cell lysate in only 25 min (Figure 4C), demonstrating the potential of online HIC-MS for both qualitative and quantitative top-down proteomics.[130] The stabilizing effect of the volatile salt and limited time on column resulted in native-like charge state distributions for proteins, aiding the detection of large protein species up to 206 kDa in the cell lysate.[130] Further development of HIC-MS by Chen et al. separated two intact IgG species, offering great potential for HIC-MS as a biotherapeutic characterization technique.[23]

There is also significant interest in the development of capillary electrophoresis-MS (CE-MS) as a non-denaturing separation technique.[132–135] The wide variety of orthogonality of separation selectivity to conventional LC-MS methods and low sample volume requirements have made CE-MS a very attractive technique for native separations.[136,137] As new commercial systems and discrete devices are developed,[138,139] we anticipate that CE-MS will rise to increased prominence in top-down proteomics.

In addition to the growing number of liquid-phase separation techniques, the use of gas-phase ion mobility to separate intact proteins prior to top-down MS has gained significant attention recently.[140–142] Ion mobility spectrometry (IMS) is a gas-phase separation with selectivity based on a molecule's drift time or mobility, a unique property of each molecule based on how it behaves in the drift tube.[143] IMS has experienced continual growth

throughout the years, yielding novel techniques such as drift tube ion mobility, traveling wave ion mobility, trapped ion mobility, field asymmetric ion mobility, and differential mobility ion mobility among others.[143] High-resolution ion mobility separation has the potential to rapidly separate proteoforms with high sequence homology with limited front-end separation. Recently, the Ruotolo group has exploited traveling wave ion mobility for the top-down characterization of chemically derivatized native protein molecules.[144] We believe advances in ion mobility technologies will attract enormous attention from the top-down proteomics community in the future as new methods and applications showing detailed molecular characterization are established.

Unlike peptides, proteins have a much more diverse range of physiochemical properties; therefore, MDLC by combining multiple orthogonal separation modalities is necessary to separate various proteins in a proteome.[5,112,145] Although extensively used in bottom-up separations, the use of two-dimensional LC (2DLC) coupled to MS in top-down proteomics was more recent.[29,146,147] Much of the 2DLC-MS work for top-down proteomics has used a comprehensive “offline” coupling strategy, collecting fractions across the first dimension of separation (^{1}D) and then analyzing all fractions using the second dimension of separation (^{2}D) interfaced directly with the MS. Offline ^{1}D separation allows samples to be buffer-exchanged or concentrated between dimensions, allowing the coupling of MS-incompatible separation techniques. Our group has developed an offline 2DLC strategy coupling HIC and RPLC (Figure 4D).[146] The orthogonality of HIC and RPLC separations greatly extended the range of separable protein polarities in an *E. coli* cell lysate.[146] To enable better characterization of large proteoforms, we used a 2DLC approach coupling sSEC to separate proteins by size and then used RPLC-MS to separate by hydrophobicity.[29] We further developed three-dimensional liquid chromatography (3DLC), coupling HIC-IEX-RPC with offline ^{1}D HIC and ^{2}D IEX separation before ^{3}D online RPLC-MS (Figure 4D).[148] We achieved a 14-fold improvement in protein identifications of a human embryonic kidney cell lysate using 3D HIC-IEX-RPLC-MS compared to 2D IEX-RPLC-MS.[148]

Though the advantages of MDLC are clear, the process is time-consuming and labor-intensive. Online MDLC strategies can alleviate this problem through automation, using a valve interface to automate transfer of eluent from ^{1}D to ^{2}D.[145] While many coupling techniques for online 2DLC exist, the comprehensive mode, which transfers all eluent from ^{1}D to ^{2}D, has the greatest potential to aid global top-down proteomics. Several groups have employed comprehensive online 2DLC for top-down proteomic analyses.[28,149,150] Moore and Jorgenson developed an online comprehensive 3DLC system for peptide separation, which has not yet been extended for use with intact proteins.[151] In the coming years, we anticipate the power of MDLC and the attraction of automation will lead to new combinations of separation modalities. We also expect the use of more non-denaturing techniques in separation strategies as the interest in native MS grows within the top-down community.

5. ADDRESSING THE CHALLENGE IN TOP-DOWN PROTEOMICS DATA ANALYSIS

Analysis of top-down proteomics data can be complex and labor-intensive, and represents a major barrier for new users.[5,152] Thus, software solutions capable of deconvoluting spectral data and searching generated results against a database have been developed. Top-down proteomics software, ProSight PC, was first introduced by the Kelleher group in 2001, and several generations of this software have been developed since its inception.[153–155] Moreover, the MS-Align[156] and TopPIC[157] database search algorithms introduced by Liu and colleagues have proven to be valuable tools for top-down proteomics data analysis. Recently, Informed Proteomics, a software developed by Park et al., is an open-source top-down proteomics software that uses a LC-MS feature-finding algorithm (ProMex) and a database search algorithm (MSPathFinder).[158] Other database search algorithms frequently used for top-down data analysis include pTOP,[159] Mascot Top Down,[32] and Proteoform Suite.[160] Nonetheless, most data analysis tools for top-down proteomics either focus on protein identification, lack a user-friendly interface, or are proprietary to instrument vendors and thus reduce accessibility to users. To address these challenges, our group developed MASH Suite Pro, an open-source software package for top-down proteomics that combines data analysis tools including protein identification, quantification, and characterization into a simple package.[152,161] However, this software is limited to processing Thermo raw data files and uses one deconvolution and one search algorithm.

To improve flexibility in data processing and expand access to multiple vendor formats, our group created MASH Explorer (Figure 5A), a top-down proteomics software that processes targeted- and discovery-based MS data for unambiguous proteoform characterization and identification (Figure 5A–C).[31] This software incorporates several deconvolution and database search algorithms in a user-friendly software environment and can process data from multiple vendor formats.[31] Moreover, it provides visualization tools for simple validation and correction of computational outputs, which is an essential step to ensure the reliability of proteomics data.[31] MASH Explorer has proven to be an indispensable tool for the top-down proteomics community, with a global user base of 2086 MASH users as of March 1, 2021 (64.5% from North America, 0.6% from South America, 17.6% from Europe, and 7.8% from Asia) (Figure 5D).

Importantly, MASH Explorer has the flexibility to incorporate new and updated software developed by computational scientists. McIlwain et al. developed an ensemble machine learning technique that can process and combine the resultant peak lists from multiple deconvolution algorithms, thus improving the accuracy and confidence in protein identifications from database searches.[162] Hence, we are currently incorporating this machine learning method in the next generation of MASH Explorer to enhance proteoform identification and characterization. We have also recently integrated the FlashDeconv deconvolution algorithm[163] to MASH Explorer. FlashDeconv is a rapid deconvolution algorithm, outpacing other deconvolution tools in the top-down proteomics field by applying a logarithmic transformation to the mass-to-charge ratio allowing for spectral pattern matching. Using this approach, Jeong et al. observed more detected masses and fewer

deconvolution artifacts compared to other methods used. With the addition of FlashDeconv, the throughput of MASH Explorer has been greatly improved.

Outside of being a universal and user-friendly software environment for top-down proteomics, future generations of MASH Explorer will be made amenable to other proteomics applications such as native top-down proteomics. In particular, the incorporation of UniDec[164] into MASH Explorer will be an exciting feature for the native MS community. UniDec developed by Marty et al. is a universal deconvolution algorithm that applies a Bayesian algorithm to separate the mass and charge dimension of a complex spectrum.[164] This software provides fast and robust deconvolution of both isotopically and non-isotopically resolved mass spectra while also incorporating ion mobility-mass spectra. Adding UniDec into the MASH Explorer software is anticipated to enable native protein identification with simultaneous visualization of protein fragment ions.

As the top-down proteomics community continues to grow, the need for universal, comprehensive, and globally accessible top-down proteomics software will increase tremendously. Future software will also need to address the challenges in the identification and characterization of large proteoforms because of the difficulties in obtaining sufficient MS/MS sequence coverage. Combination of multiple MS/MS techniques including but not limited to activated ion electron capture[165]/transfer dissociation,[166] and more recently developed ultraviolet photodissociation,[167] and electron ionization dissociation,[168,169] will likely have great potential for proteoform characterization and therefore will require the necessary software tools for effective data analysis to be available to the top-down proteomics community.

6. ADDRESSING THE CHALLENGE IN LINKING PROTEOFORMS TO FUNCTION

Proteins are the primary effectors of function in biology, and thus, the top-down proteomics field is actively seeking to link proteoform-resolved information with functional characteristics in various biological systems.[170] Chamot-Rooke et al. revealed the functionally critical PilE proteoforms carrying phosphoglycerol groups on serines were tightly associated with crossing of the epithelial barrier and access to the bloodstream in cerebrospinal meningitis (*N. meningitidis*).[33] Paša-Tolić and co-workers showed a unique protein S-thiolation switch as a result of infection-like conditions in *Salmonella typhimurium*.[36] Carel et al. employed top-down AP-MS to decipher the functional role of *O*-mycoloylation in targeting the outer membrane proteins to the mycomembrane in bacteria.[171] Kelleher and co-workers demonstrated the crosstalk between mutation and PTMs after immunoaffinity purification and top-down MS analysis of the KRAS4b proteoforms, which provides new insights in KRAS biology in cancer.[40] Agar and co-workers applied top-down MS to ALS-patient immunopurified super-oxide dismutase-1 (SOD1) and revealed that, among the three prevalent SOD1 proteoforms, one had a neurotoxic function and another one had a neuroprotective function.[172]

Our group has previously combined top-down proteomics and functional measurements to study sarcopenia, the deterioration of muscle mass and strength as a result of aging, a major

burden on the healthcare system, but the molecular mechanism underlying sarcopenia-associated muscle dysfunction remains largely unknown.[39,173] We uncovered a progressive age-related decrease in the phosphorylation of myosin regulatory light chain (RLC), a critical regulatory protein in muscle contraction, in the skeletal muscle of aging rats using top-down targeted proteomics (Figure 6A).[39] The sites of decreasing phosphorylation were localized to Ser14/15, of which Ser14 phosphorylation was a previously unidentified phosphorylation site in RLC from fast-twitch skeletal muscle. The functional measurement of the mechanical properties of single fast-twitch fibers indicated that age-related decreases in the contractile properties of sarcopenic fast-twitch muscles could be accounted for by the decreased RLC phosphorylation. This study provides strong evidence for the functional role of RLC phosphorylation at Ser14/15 in sarcopenia-related muscle dysfunction and the potential therapeutic modulation of RLC phosphorylation as a potential avenue to treat sarcopenia.[39] In another study, we developed a top-down proteomics method that allowed for simultaneous quantification of sarcomeric proteoforms and contractile measurements from rat fast- and slow-twitch skeletal muscle tissues to correlate muscle function with sarcomeric proteoform expression.[173] Importantly, there was a connection with the sarcomeric proteoforms found in the aging rats, which was consistent with the functional phenotype. Age-related changes were observed in the phosphorylation levels of Z-disc proteins, which represented the first report of PTMs in the Z-disc associated with aging. The use of proteoform measurements with muscle contraction properties will continue to prove to be essential in the relationship between proteoforms and function in human muscle diseases.

Recently, our lab has been developing methods that integrate omics data with functional properties of human cardiomyocytes differentiated from induced pluripotent stem cells (hiPSC-CM).[43,45] We established an unbiased method combining proteomics with functional measurements for the accurate and comprehensive assessment of hiPSC-CM maturation.[43] In addition to the detection of known maturation-associated contractile protein alterations, for the first time, contractile protein PTMs such as decreased phosphorylation of α-tropomyosin were identified as potential novel hiPSC-CMs maturation markers. While most hiPSC-CM research is performed using 2D monolayer systems, there has been a significant effort to develop three-dimensional (3D) engineered cardiac tissues (ECTs) because ECTs more faithfully reflect the relevant cardiac physiology of endogenous cardiac tissues.[45] However, many of the metrics used to assess hiPSC-ECTs are often performed on separate samples due to incompatibilities between the various measurements, which is problematic due to the highly heterogeneous nature of stem-cell-derived CMs. To address this issue, Melby et al. developed a functionally integrated top-down proteomics method for a standardized assessment of hiPSC-ECTs (Figure 6B).[45] Isometric twitch force measurements were performed on two groups of hiPSC-ECTs, and using the same samples, top-down proteomics was implemented to analyze sarcomeric proteoforms. This integrated approach allowed for the determination of direct correlations between functional properties and sarcomeric proteoforms. We envision that this method will be highly valuable in future studies using hiPSC-ECTs for disease modeling, cardiotoxicity screenings, and regenerative therapies. Although all studies have linked proteoforms with functions in various biological systems and diseases to some extent, there remains a challenge to establish a direct causal

relationship between proteoform changes and the corresponding functional properties to further understand the cause and effect relationship.

7. THROUGHPUT OF TOP-DOWN PROTEOMICS FOR PRECISION MEDICINE

Precision medicine, wherein disease treatment and prevention consider individual variability, is increasingly recognized as the future of modern health care.[174,175] High-throughput omics technologies are becoming enabling forces for precision medicine with the promise to characterize patients at an extraordinarily detailed molecular level.[176,177] In the post-genomics era, a comprehensive analysis of proteoforms including PTMs and mutations/alternative splicing is essential to gain a transformative understanding of disease mechanisms and identifying new therapeutic targets.[3,178–180]

While top-down proteomics is not yet fully incorporated in the clinic, there are several translational studies that eloquently demonstrate the power that this technology could provide for clinical applications.[4,30,34,40–42,48] Top-down proteomics has been applied to important clinical samples, such as cardiac,[34,44] liver,[181] and tumor[40] biopsy samples, blood,[41] and serum,[30] to understand the role of proteoforms in human disease. For example, Kelleher and co-workers applied top-down proteomics for the analysis of proteoform signatures from acute rejection fractions of circulating tumor cells and determined there was a distinct proteoform signature between patients that received a liver transplant versus the control group.[181] Our group has linked altered cardiac proteoforms to disease phenotypes in heart failure using both animal models and human clinical samples.[34,35,38] Phosphorylated proteoforms of cardiac troponin I (cTnI), a gold standard biomarker for detection of cardiac injury, were identified by top-down quantitative proteomics to be candidate biomarkers for chronic heart failure.[34] Top-down proteomics has also identified actin isoforms as a potential cardiac disease marker[182] and discovered novel phosphorylation of a critical Z-disc protein, Enigma Homologue isoform 2 in acute myocardial infarction.[38]

Recently, we applied top-down proteomics to septal myectomy samples collected from hypertrophic cardiomyopathy (HCM) patients with severe outflow tract obstruction and revealed a shared sarcomeric proteoform pattern across HCM tissues when compared to non-failing donor heart tissues.[44] HCM is a genetically heterogeneous disease with nearly 1400 mutations identified in at least 11 genes that encode proteins within the sarcomere.[183] However, regardless of various HCM-causing mutations, the detected sarcomeric proteoform profiles of these samples were highly consistent. This finding suggests that common signaling pathways are activated in the myocardium of obstructive HCM patients with similar phenotypes requiring septal myectomy intervention and opens the door for future therapeutic interventions targeting HCM phenotypes, which are more accurately reflected by the proteoform profiles rather than the specific genotype.[44] These various examples demonstrate the importance of proteoform-level molecular details, which provide critical insights into disease progression, in a variety of clinical research areas.

To fully reach the potential of top-down proteomics in precision medicine and translate to the clinic, several challenges mentioned in this perspective must be addressed. Future

technology developments that can enable the routine proteoform-resolved characterization of proteins >50 kDa are of particular interest, as over half of the proteome contains proteins of this size.[29] To develop a top-down proteomics assay in the clinic, it is essential to analyze biofluids, such as blood and serum, which still suffer from the immense challenge of finding the low-abundance biomarkers of interest among an ocean of other highly abundant blood proteins in the vast biological milieu.[30] Important progress has been made with standardized top-down proteomics protocols for the analysis of peripheral blood mononuclear cells and other common biofluids; however, challenges related to proteome complexity remain.[42] We anticipate that novel approaches such as nanoproteomics,[30] and other enrichment techniques, will prove to be critical to addressing this major challenge. Moreover, rigor and reproducibility need to be established to ensure robust and unbiased experimental design, analysis, interpretation, and reporting of results, which is especially critical for the further development of top-down quantitative proteomics toward its clinical use. Top-down proteomics recently has made major strides in the reproducible and quantitative analysis of proteforms.[4,43,44,48,106] Arguably, it has great potential to be more rigorous and reproducible than other approaches such as conventional bottom-up MS- and antibody-based techniques. Nonetheless, further method development and technology improvements will be required to extend the application of top-down MS-based quantitation toward complex PTMs or heavily modified proteins.

Another pressing issue as it pertains to the utility of proteoform analysis in clinical applications is the throughput of current top-down proteomics workflows. Great progress has been made in high-throughput clinical applications for bottom-up proteomics pioneered by several groups such as Regnier,[184] Van Eyk,[185,186] and Carr.[187] Collectively, these studies demonstrate both the feasibility and benefits of using proteomics for the analysis of biomolecules from clinical samples. However, except for matrix assisted laser desorption ionization-TOF-based intact protein assays,[7,188] the typical discovery-based top-down proteomics still suffers from low throughput and there is an urgent need to optimize the entire workflow including sample preparation, separation, data acquisition, and analysis. To speed up sample preparation, the use of automation has the potential to revolutionize the application of top-down proteomics in a clinical setting by greatly enhancing accuracy, reproducibility, and time spent in preparing samples. Although online LC methods are common-place in current top-down MS workflows, given the high speed of ion mobility separations,[143] further technology development with improved front-end ion mobility separations may enable rapid and robust protein separation. Similarly, the Evosep system implements sensitive nanoflow liquid separations using a rapid reverse phase gradient proving to be a valuable clinical tool.[189,190] We anticipate a greater array of new column chemistry will enable this technology to be applied for proteoform-resolved clinical sample analysis. Furthermore, for discovery-based top-down proteomics experiments, the process of deconvoluting the highly complex data and then searching the peak list against a database generally takes several hours to fully complete. Discovery-based searches are far slower than target-based searches but do not require *a priori* knowledge of the proteoform of interest, which for some clinical applications is not a possibility. Conceivably, top-down targeted proteomics focusing on one or multiple protein targets might have great advantages to be used in the clinic. Currently, the nanoproteomics cTnI-NP assay and the HCM clinical

patient analysis by top-down proteomics require ~2.5 h total analysis time, when accounting for sample preparation, workup, and final data acquisition.[30,44] We anticipate future technology developments in both software and hardware will enhance the throughput of top-down proteomics, enabling the use of top-down proteomics for precision medicine and translation to the clinic in the near future.

8. CONCLUSION AND FUTURE DIRECTIONS

In this Account & Perspective, we have provided our view of the major challenges affiliated with top-down proteomics at present and new strategies (including those already developed and to be developed) to address these challenges. We have focused on the most pressing areas including protein solubility, dynamic range of the proteome, proteome complexity, data analysis, linking proteoforms to biological function, and throughput of top-down proteomics for precision medicine. There are also other ongoing challenges associated with top-down proteomics such as the sensitivity, rigor, and reproducibility of top-down proteomics. As an invited contribution for the Special Focus Issue of JASMS that is dedicated to Ying Ge in honor of her 2020 Biemann Medal, this Account & Perspective focused more on the contributions of our laboratory to the rapidly growing top-down proteomics field. We greatly appreciate and value the contributions by other top-down proteomics laboratories all over the world, some of which are also discussed here, but it is not possible to comprehensively review all advances in this single Account & Perspective. We envision that the Consortium of Top-down Proteomics (www.topdownproteomics.org) will continue to develop and grow as a community and foster international collaborations to further advance top-down proteomics. It is also essential to establish interdisciplinary teams with diverse expertise in different fields to continuously address the remaining challenges for top-down proteomics to reach its full potential. We firmly believe that future advancements in the many strategies discussed herein, together with the continued improvement in MS instrumentation, device miniaturization to improve sensitivity, automation for high-throughput sample processing, data acquisition, and analysis, will enable the ubiquitous application of top-down proteomics in biomedical research for precision medicine.

ACKNOWLEDGMENTS

We would like to acknowledge the funding from the National Institutes of Health (NIH) R01s, HL096971, HL109810, GM117058, GM125085, and S10 OD018475. J.A.M. would like to acknowledge support from the Training Program in Translational Cardiovascular Science, T32 HL007936-20. D.S.R. would like to acknowledge support from the American Heart Association Predoctoral Fellowship Grant No. 832615/David S. Roberts/2021. K.A.B. would like to acknowledge the Training Program in Translational Cardiovascular Science, T32 HL007936-19, and the Vascular Surgery Research Training Program grant, T32HL110853. Y.G. would like to acknowledge all of the current and former group members as well as our collaborators who have made significant contributions to the Ge lab's development of top-down proteomics which culminated in this Biemann Medal Award.

REFERENCES

(1). Smith LM; Kelleher NL Proteoform: A Single Term Describing Protein Complexity. Nat. Methods 2013, 10, 186–187. [PubMed: 23443629]

(2). Smith LM; Kelleher NL Proteoforms as the next Proteomics Currency: Identifying Precise Molecular Forms of Proteins Can Improve Our Understanding of Function. Science 2018, 359 (6380), 1106–1107. [PubMed: 29590032]

(3). Aebersold R; Agar JN; Amster IJ; Baker MS; Bertozzi CR; Boja ES; Costello CE; Cravatt BF; Fenselau C; Garcia BA; Ge Y; Gunawardena J; Hendrickson RC; Hergenrother PJ; Huber CG; Ivanov AR; Jensen ON; Jewett MC; Kelleher NL; Kiessling LL; Krogan NJ; Larsen MR; Loo JA; Ogorzalek Loo RR; Lundberg E; MacCoss MJ; Mallick P; Mootha VK; Mrksich M; Muir TW; Patrie SM; Pesavento JJ; Pitteri SJ; Rodriguez H; Saghatelian A; Sandoval W; Schlüter H; Sechi S; Slavoff SA; Smith LM; Snyder MP; Thomas PM; Uhlén M; Van Eyk JE; Vidal M; Walt DR; White FM; Williams ER; Wohlschlager T; Wysocki VH; Yates NA; Young NL; Zhang B How many human proteoforms are there? Nat. Chem. Biol 2018, 14, 206–214. [PubMed: 29443976]

(4). Cai W; Tucholski TM; Gregorich ZR; Ge Y Top-down Proteomics: Technology Advancements and Applications to Heart Diseases. Expert Rev. Proteomics 2016, 13, 717–730. [PubMed: 27448560]

(5). Chen B; Brown KA; Lin Z; Ge Y Top-Down Proteomics: Ready for Prime Time? Anal. Chem 2018, 90, 110–127. [PubMed: 29161012]

(6). Siuti N; Kelleher NL Decoding Protein Modifications Using Top-down Mass Spectrometry. Nat. Methods 2007, 4 (10), 817–821. [PubMed: 17901871]

(7). Kelleher NL; Thomas PM; Ntai I; Compton PD; Leduc RD Deep and Quantitative Top-down Proteomics in Clinical and Translational Research. Expert Rev. Proteomics 2014, 11, 649–651. [PubMed: 25347991]

(8). Zhang Y; Fonslow BR; Shan B; Baek MC; Yates JR Protein Analysis by Shotgun/Bottom-up Proteomics. Chem. Rev 2013, 113 (4), 2343–2394. [PubMed: 23438204]

(9). Aebersold R; Mann M Mass-Spectrometric Exploration of Proteome Structure and Function. Nature 2016, 537, 347–355. [PubMed: 27629641]

(10). Cravatt BF; Simon GM; Yates JR The Biological Impact of Mass-Spectrometry-Based Proteomics. Nature 2007, 450 (7172), 991–1000. [PubMed: 18075578]

(11). Chait BT Mass Spectrometry: Bottom-up or Top-Down? Science 2006, 314 (5796), 65–66. [PubMed: 17023639]

(12). Kelleher NL; Lin HY; Valaskovic GA; Aaserud DJ; Fridriksson EK; McLafferty FW Top down versus Bottom up Protein Characterization by Tandem High-Resolution Mass Spectrometry. J. Am. Chem. Soc 1999, 121 (4), 806–812.

(13). Loo JA; Edmonds CG; Smith RD Primary Sequence Information from Intact Proteins by Electrospray Ionization Tandem Mass Spectrometry. Science 1990, 248 (4952), 201–204. [PubMed: 2326633]

(14). Ge Y; Lawhorn BG; ElNaggar M; Strauss E; Park JH; Begley TP; McLafferty FW Top down Characterization of Larger Proteins (45 KDa) by Electron Capture Dissociation Mass Spectrometry. J. Am. Chem. Soc 2002, 124 (4), 672–678. [PubMed: 11804498]

(15). Ge Y; ElNaggar M; Sze SK; Oh HB; Begley TP; McLafferty FW; Boshoff H; Barry CE Top Down Characterization of Secreted Proteins from Mycobacterium Tuberculosis by Electron Capture Dissociation Mass Spectrometry. J. Am. Soc. Mass Spectrom 2003, 14 (3), 253–261. [PubMed: 12648932]

(16). Han X; Jin M; Breuker K; McLafferty FW Extending Top-down Mass Spectrometry to Proteins with Masses Great than 200 Kilodaltons. Science 2006, 314 (5796), 109–112. [PubMed: 17023655]

(17). Reid GE; McLuckey SA ‘Top down’ Protein Characterization via Tandem Mass Spectrometry. J. Mass Spectrom 2002, 37 (7), 663–675. [PubMed: 12124999]

(18). Marshall AG; Hendrickson CL High-Resolution Mass Spectrometers. Annu. Rev. Anal. Chem 2008, 1 (1), 579–599.

(19). Meng F; Cargile BJ; Miller LM; Forbes AJ; Johnson JR; Kelleher NL Informatics and Multiplexing of Intact Protein Identification in Bacteria and the Archaea. Nat. Biotechnol 2001, 19 (10), 952–957. [PubMed: 11581661]

(20). Meng F; Cargile BJ; Patrie SM; Johnson JR; McLoughlin SM; Kelleher NL Processing Complex Mixtures of Intact Proteins for Direct Analysis by Mass Spectrometry. Anal. Chem 2002, 74 (13), 2923–2929. [PubMed: 12141648]

(21). Kelleher NL Peer Reviewed: Top-Down Proteomics. Anal. Chem 2004, 76 (11), 196 A–203 A.

(22). Tran JC; Zamdborg L; Ahlf DR; Lee JE; Catherman AD; Durbin KR; Tipton JD; Vellaichamy A; Kellie JF; Li M; Wu C; Sweet SMM; Early BP; Siuti N; LeDuc RD; Compton PD; Thomas PM; Kelleher NL Mapping Intact Protein Isoforms in Discovery Mode Using Top-down Proteomics. Nature 2011, 480 (7376), 254–258. [PubMed: 22037311]

(23). Chen B; Lin Z; Alpert AJ; Fu C; Zhang Q; Pritts WA; Ge Y Online Hydrophobic Interaction Chromatography-Mass Spectrometry for the Analysis of Intact Monoclonal Antibodies. Anal. Chem 2018, 90 (12), 7135–7138. [PubMed: 29846060]

(24). Takemori A; Butcher DS; Harman VM; Brownridge P; Shima K; Higo D; Ishizaki J; Hasegawa H; Suzuki J; Yamashita M; Loo JA; Ogorzalek Loo RR; Beynon RJ; Anderson LC; Takemori N PEPPI-MS: Polyacrylamide-Gel-Based Prefractionation for Analysis of Intact Proteoforms and Protein Complexes by Mass Spectrometry. J. Proteome Res 2020, 19 (9), 3779–3791. [PubMed: 32538093]

(25). Catherman AD; Durbin KR; Ahl DR; Early BP; Fellers RT; Tran JC; Thomas PM; Kelleher NL Large-Scale Top-down Proteomics of the Human Proteome: Membrane Proteins, Mitochondria, and Senescence. Mol. Cell. Proteomics 2013, 12 (12), 3465–3473. [PubMed: 24023390]

(26). Li H; Nguyen HH; Ogorzalek Loo RR; Campuzano IDG; Loo JA An Integrated Native Mass Spectrometry and Top-down Proteomics Method That Connects Sequence to Structure and Function of Macromolecular Complexes. Nat. Chem 2018, 10 (2), 139. [PubMed: 29359744]

(27). McCool EN; Lubeckyj RA; Shen X; Chen D; Kou Q; Liu X; Sun L Deep Top-Down Proteomics Using Capillary Zone Electrophoresis-Tandem Mass Spectrometry: Identification of 5700 Proteoforms from the *Escherichia Coli* Proteome. Anal. Chem 2018, 90 (9), 5529–5533. [PubMed: 29620868]

(28). Wang Z; Yu D; Cupp-Sutton KA; Liu X; Smith K; Wu S Development of an Online 2D Ultrahigh-Pressure Nano-LC System for High-PH and Low-PH Reversed Phase Separation in Top-Down Proteomics. Anal. Chem 2020, 92 (19), 12774–12777. [PubMed: 32857493]

(29). Cai W; Tucholski T; Chen B; Alpert AJ; McIlwain S; Kohmoto T; Jin S; Ge Y Top-Down Proteomics of Large Proteins up to 223 KDa Enabled by Serial Size Exclusion Chromatography Strategy. Anal. Chem 2017, 89 (10), 5467–5475. [PubMed: 28406609]

(30). Tiambeng TN; Roberts DS; Brown KA; Zhu Y; Chen B; Wu Z; Mitchell SD; Guardado-Alvarez TM; Jin S; Ge Y Nanoproteomics Enables Proteoform-Resolved Analysis of Low-Abundance Proteins in Human Serum. Nat. Commun 2020, 11 (1), 3903. [PubMed: 32764543]

(31). Wu Z; Roberts DS; Melby JA; Wenger K; Wetzel M; Gu Y; Ramanathan SG; Bayne EF; Liu X; Sun R; Ong IM; McIlwain SJ; Ge Y MASH Explorer: A Universal Software Environment for Top-Down Proteomics. J. Proteome Res 2020, 19 (9), 3867–3876. [PubMed: 32786689]

(32). Karabacak NM; Li L; Tiwari A; Hayward LJ; Hong P; Easterling ML; Agar JN Sensitive and Specific Identification of Wild Type and Variant Proteins from 8 to 669 KDa Using Top-down Mass Spectrometry. Mol. Cell. Proteomics 2009, 8 (4), 846–856. [PubMed: 19074999]

(33). Chamot-Rooke J; Mikaty G; Malosse C; Soyer 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 (6018), 778–782. [PubMed: 21311024]

(34). Zhang J; Guy MJ; Norman HS; Chen YC; Xu Q; Dong X; Guner H; Wang S; Kohmoto T; Young KH; Moss RL; Ge Y Top-down Quantitative Proteomics Identified Phosphorylation of Cardiac Troponin i as a Candidate Biomarker for Chronic Heart Failure. J. Proteome Res 2011, 10 (9), 4054–4065. [PubMed: 21751783]

(35). 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 (2), 848–857. [PubMed: 22052912]

(36). Ansong C; Wu S; Meng D; Liu X; Brewer HM; Kaiser BLD; Nakayasu ES; Cort JR; Pevzner P; Smith RD; Heffron F; Adkins JN; Paša-Tolić L Top-down Proteomics Reveals a Unique Protein S-Thiolation Switch in Salmonella Typhimurium in Response to Infection-like Conditions. Proc. Natl. Acad. Sci. U. S. A 2013, 110 (25), 10153–10158. [PubMed: 23720318]

(37). Bourgoin-Voillard S; Leymarie N; Costello CE Top-down Tandem Mass Spectrometry on RNase A and B Using a Qh/FT-ICR Hybrid Mass Spectrometer. Proteomics 2014, 14 (10), 1174–1184. [PubMed: 24687996]

(38). Peng Y; Gregorich ZR; Valeja SG; Zhang H; Cai W; Chen YC; Guner H; Chen AJ; Schwahn DJ; Hacker TA; Liu X; Ge Y Top-down Proteomics Reveals Concerted Reductions in Myofilament and Z-Disc Protein Phosphorylation after Acute Myocardial Infarction. Mol. Cell. Proteomics 2014, 13 (10), 2752–2764. [PubMed: 24969035]

(39). Gregorich ZR; Peng Y; Cai W; Jin Y; Wei L; Chen AJ; McKiernan SH; Aiken JM; Moss RL; Diffee GM; Ge Y Top-down Targeted Proteomics Reveals Decrease in Myosin Regulatory Light-Chain Phosphorylation That Contributes to Sarcopenic Muscle Dysfunction. J. Proteome Res 2016, 15 (8), 2706–2716. [PubMed: 27362462]

(40). Ntai I; Fornelli L; DeHart CJ; Hutton JE; Doubleday PF; LeDuc RD; van Nispen AJ; Fellers RT; Whiteley G; Boja ES; Rodriguez H; Kelleher NL Precise Characterization of KRAS4b Proteoforms in Human Colorectal Cells and Tumors Reveals Mutation/Modification Cross-Talk. Proc. Natl. Acad. Sci. U. S. A 2018, 115 (16), 4140–4145. [PubMed: 29610327]

(41). He L; Rockwood AL; Agarwal AM; Anderson LC; Weisbrod CR; Hendrickson CL; Marshall AG Diagnosis of Hemoglobinopathy and β-Thalassemia by 21 T Fourier Transform Ion Cyclotron Resonance Mass Spectrometry and Tandem Mass Spectrometry of Hemoglobin from Blood. Clin. Chem 2019, 65 (8), 986–994. [PubMed: 31040099]

(42). Toby TK; Fornelli L; Srzentić K; DeHart CJ; Levitsky J; Friedewald J; Kelleher NL A Comprehensive Pipeline for Translational Top-down Proteomics from a Single Blood Draw. Nat. Protoc 2019, 14 (1), 119–152. [PubMed: 30518910]

(43). Cai W; Zhang J; de Lange WJ; Gregorich ZR; Karp H; Farrell ET; Mitchell SD; Tucholski T; Lin Z; Biermann M; McIlwain SJ; Ralphe JC; Kamp TJ; Ge Y An Unbiased Proteomics Method to Assess the Maturation of Human Pluripotent Stem Cell–Derived Cardiomyocytes. Circ. Res 2019, 125 (11), 936–953. [PubMed: 31573406]

(44). Tucholski T; Cai W; Gregorich ZR; Bayne EF; Mitchell SD; McIlwain SJ; de Lange WJ; Wrobbel M; Karp H; Hite Z; Vikhorev PG; Marston SB; Lal S; Li A; Dos Remedios C; Kohmoto T; Hermsen J; Ralphe JC; Kamp TJ; Moss RL; Ge Y Distinct Hypertrophic Cardiomyopathy Genotypes Result in Convergent Sarcomeric Proteoform Profiles Revealed by Top-down Proteomics. Proc. Natl. Acad. Sci. U. S. A 2020, 117 (40), 24691–24700. [PubMed: 32968017]

(45). Melby JA; De Lange WJ; Zhang J; Roberts DS; Mitchell SD; Tucholski T; Kim G; Kyrvasilis A; McIlwain SJ; Kamp TJ; Ralphe JC; Ge Y Functionally Integrated Top-Down Proteomics for Standardized Assessment of Human Induced Pluripotent Stem Cell-Derived Engineered Cardiac Tissues. J. Proteome Res 2021, 20, 1424–1433. [PubMed: 33395532]

(46). Catherman AD; Skinner OS; Kelleher NL Top Down Proteomics: Facts and Perspectives. Biochem. Biophys. Res. Commun 2014, 445 (4), 683–693. [PubMed: 24556311]

(47). Toby TK; Fornelli L; Kelleher NL Progress in Top-Down Proteomics and the Analysis of Proteoforms. Annu. Rev. Anal. Chem 2016, 9, 499–519.

(48). Brown KA; Melby JA; Roberts DS; Ge Y Top-down Proteomics: Challenges, Innovations, and Applications in Basic and Clinical Research. Expert Rev. Proteomics 2020, 17 (10), 719–733. [PubMed: 33232185]

(49). Tucholski T; Ge Y Fourier-Transform Ion Cyclotron Resonance Mass Spectrometry for Characterizing Proteoforms. Mass Spectrom. Rev 2020, DOI: 10.1002/mas.21653.

(50). Gregorich ZR; Ge Y Top-down Proteomics in Health and Disease: Challenges and Opportunities. Proteomics 2014, 14 (10), 1195–1210. [PubMed: 24723472]

(51). Armirotti A; Damonte G Achievements and Perspectives of Top-down Proteomics. Proteomics 2010, 10 (20), 3566–3576. [PubMed: 20859958]

(52). Cui W; Rohrs HW; Gross ML Top-down Mass Spectrometry: Recent Developments, Applications and Perspectives. Analyst 2011, 136 (19), 3854–3864. [PubMed: 21826297]

(53). Schaffer LV; Millikin RJ; Miller RM; Anderson LC; Fellers RT; Ge Y; Kelleher NL; LeDuc RD; Liu X; Payne SH; Sun L; Thomas PM; Tucholski T; Wang Z; Wu S; Wu Z; Yu D; Shortreed MR; Smith LM Identification and Quantification of Proteoforms by Mass Spectrometry. Proteomics 2019, 19 (10), 1800361.

(54). Whitelegge JP Integral Membrane Proteins and Bilayer Proteomics. Anal. Chem 2013, 85 (5), 2558–2568. [PubMed: 23301778]

(55). Wu CC; Yates JR The Application of Mass Spectrometry to Membrane Proteomics. Nat. Biotechnol 2003, 21 (3), 262–267. [PubMed: 12610573]

(56). Naba A; Clauser KR; Hoersch S; Liu H; Carr SA; Hynes RO The Matrisome: In Silico Definition and in Vivo Characterization by Proteomics of Normal and Tumor Extracellular Matrices. Mol. Cell. Proteomics 2012, 11 (4), M111.014647.

(57). Lindsey ML; Jung M; Hall ME; DeLeon-Pennell KY Proteomic Analysis of the Cardiac Extracellular Matrix: Clinical Research Applications. Expert Rev. Proteomics 2018, 15 (2), 105–112. [PubMed: 29285949]

(58). Barrera NP; Robinson CV Advances in the Mass Spectrometry of Membrane Proteins: From Individual Proteins to Intact Complexes. Annu. Rev. Biochem 2011, 80, 247–271. [PubMed: 21548785]

(59). Brown KA; Chen B; Guardado-Alvarez TM; Lin Z; Hwang L; Ayaz-Guner S; Jin S; Ge Y A Photocleavable Surfactant for Top-down Proteomics. Nat. Methods 2019, 16 (5), 417–420. [PubMed: 30988469]

(60). Knott SJ; Brown KA; Josyer H; Carr A; Inman D; Jin S; Friedl A; Ponik SM; Ge Y Photocleavable Surfactant-Enabled Extracellular Matrix Proteomics. Anal. Chem 2020, 92 (24), 15693–15735. [PubMed: 33232116]

(61). Carroll J; Altman MC; Fearnley IM; Walker JE Identification of Membrane Proteins by Tandem Mass Spectrometry of Protein Ions. Proc. Natl. Acad. Sci. U. S. A 2007, 104 (36), 14330–14335. [PubMed: 17720804]

(62). Ryan CM; Souda P; Bassilian S; Ujwal R; Zhang J; Abramson J; Ping P; Durazo A; Bowie JU; Hasan SS; Baniulis D; Cramer WA; Faull KF; Whitelegge JP Post-Translational Modifications of Integral Membrane Proteins Resolved by Top-down Fourier Transform Mass Spectrometry with Collisionally Activated Dissociation. Mol. Cell. Proteomics 2010, 9 (5), 791–803. [PubMed: 20093275]

(63). Skinner OS; Catherman AD; Early BP; Thomas PM; Compton PD; Kelleher NL Fragmentation of Integral Membrane Proteins in the Gas Phase. Anal. Chem 2014, 86 (9), 4627–4634. [PubMed: 24689519]

(64). Catherman AD; Li M; Tran JC; Durbin KR; Compton PD; Early BP; Thomas PM; Kelleher NL Top down Proteomics of Human Membrane Proteins from Enriched Mitochondrial Fractions. Anal. Chem 2013, 85 (3), 1880–1888. [PubMed: 23305238]

(65). Whitelegge JP; Gundersen CB; Faull KF Electrospray-Ionization Mass Spectrometry of Intact Intrinsic Membrane Proteins. Protein Sci. 1998, 7 (6), 1423–1430. [PubMed: 9655347]

(66). Wiśniewski JR; Zougman A; Nagaraj N; Mann M Universal Sample Preparation Method for Proteome Analysis. Nat. Methods 2009, 6 (5), 359–362. [PubMed: 19377485]

(67). Speers AE; Wu CC Proteomics of Integral Membrane Proteins - Theory and Application. Chem. Rev 2007, 107, 3687–3714. [PubMed: 17683161]

(68). Loo RRO; Dales N; Andrews PC Surfactant Effects on Protein Structure Examined by Electrospray Ionization Mass Spectrometry. Protein Sci. 1994, 3 (11), 1975–1983. [PubMed: 7703844]

(69). Laganowsky A; Reading E; Hopper JTS; Robinson CV Mass Spectrometry of Intact Membrane Protein Complexes. Nat. Protoc 2013, 8 (4), 639–651. [PubMed: 23471109]

(70). Yu YQ; Gilar M; Lee PJ; Bouvier ESP; Gebler JC Enzyme-Friendly, Mass Spectrometry-Compatible Surfactant for In-Solution Enzymatic Digestion of Proteins. Anal. Chem 2003, 75 (21), 6023–6028. [PubMed: 14588046]

(71). Saveliev SV; Woodroofe CC; Sabat G; Adams CM; Klaubert D; Wood K; Urh M Mass Spectrometry Compatible Surfactant for Optimized In-Gel Protein Digestion. Anal. Chem 2013, 85 (2), 907–914. [PubMed: 23256507]

(72). Chang YH; Gregorich ZR; Chen AJ; Hwang L; Guner H; Yu D; Zhang J; Ge Y New Mass-Spectrometry-Compatible Degradable Surfactant for Tissue Proteomics. J. Proteome Res 2015, 14 (3), 1587–1599. [PubMed: 25589168]

(73). Brown KA; Tucholski T; Eken C; Knott S; Zhu Y; Jin S; Ge Y High-Throughput Proteomics Enabled by a Photocleavable Surfactant. Angew. Chem 2020, 132 (22), 8484–8488.

(74). Hynes RO The Extracellular Matrix: Not Just Pretty Fibrils. Science 2009, 326 (5957), 1216–1219. [PubMed: 19965464]

(75). Mouw JK; Ou G; Weaver VM Extracellular Matrix Assembly: A Multiscale Deconstruction. Nat. Rev. Mol. Cell Biol 2014, 15 (12), 771–785. [PubMed: 25370693]

(76). Garcia BA; Pesavento JJ; Mizzen CA; Kelleher NL Pervasive Combinatorial Modification of Histone H3 in Human Cells. Nat. Methods 2007, 4 (6), 487–489. [PubMed: 17529979]

(77). Holt MV; Wang T; Young NL Expeditious Extraction of Histones from Limited Cells or Tissue Samples and Quantitative Top-Down Proteomic Analysis. Curr. Protoc 2021, 1 (2), No. e26. [PubMed: 33534192]

(78). Donnelly DP; Rawlins CM; DeHart CJ; Fornelli L; Schachner LF; Lin Z; Lippens JL; Aluri KC; Sarin R; Chen B; Lantz C; Jung W; Johnson KR; Koller A; Wolff JJ; Campuzano IDG; Auclair JR; Ivanov AR; Whitelegge JP; Pasa-Tolic L; Chamot-Rooke J; Danis PO; Smith LM; Tsybin YO; Loo JA; Ge Y; Kelleher NL; Agar J Best Practices and Benchmarks for Intact Protein Analysis for Top-down Mass Spectrometry. Nat. Methods 2019, 16 (7), 587–594. [PubMed: 31249407]

(79). Ruotolo BT; Giles K; Campuzano I; Sandercock AM; Bateman RH; Robinson CV Biochemistry: Evidence for Macromolecular Protein Rings in the Absence of Bulk Water. Science 2005, 310 (5754), 1658–1661. [PubMed: 16293722]

(80). Bolla JR; Agasid MT; Mehmood S; Robinson CV Membrane Protein-Lipid Interactions Probed Using Mass Spectrometry. Annu. Rev. Biochem 2019, 88, 85–111. [PubMed: 30901263]

(81). Leney AC; Heck AJR Native Mass Spectrometry: What Is in the Name? J. Am. Soc. Mass Spectrom 2017, 28, 5–13.

(82). Gault J; Liko I; Landreh M; Shutin D; Bolla JR; Jefferies D; Agasid M; Yen HY; Ladds MJGW; Lane DP; Khalid S; Mullen C; Remes PM; Huguet R; McAlister G; Goodwin M; Viner R; Syka JEP; Robinson CV Combining Native and 'Omics' Mass Spectrometry to Identify Endogenous Ligands Bound to Membrane Proteins. Nat. Methods 2020, 17 (5), 505–508. [PubMed: 32371966]

(83). Skinner OS; Haverland NA; Fornelli L; Melani RD; Do Vale LHF; Seckler HS; Doubleday PF; Schachner LF; Srzentić K; Kelleher NL; Compton PD Top-down Characterization of Endogenous Protein Complexes with Native Proteomics. Nat. Chem. Biol 2018, 14 (1), 36–41. [PubMed: 29131144]

(84). Keener JE; Zhang G; Marty MT Native Mass Spectrometry of Membrane Proteins. Anal. Chem 2021, 93, 583–597. [PubMed: 33115234]

(85). Schachner LF; Jooß K; Morgan MA; Piunti A; Meiners MJ; Kafader JO; Lee AS; Iwanaszko M; Cheek MA; Burg JM; Howard SA; Keogh M; Shilatifard A; Kelleher NL Decoding the Protein Composition of Whole Nucleosomes with Nuc-MS. Nat. Methods 2021, 18, 303. [PubMed: 33589837]

(86). Zhou M; Lantz C; Brown KA; Ge Y; Pasa-Tolic L; Loo JA; Lermyte F Higher-Order Structural Characterisation of Native Proteins and Complexes by Top-down Mass Spectrometry. Chem. Sci 2020, 11, 12918–12936. [PubMed: 34094482]

(87). Urner LH; Liko I; Yen HY; Hoi KK; Bolla JR; Gault J; Almeida FG; Schweder MP; Shutin D; Ehrmann S; Haag R; Robinson CV; Pagel K Modular Detergents Tailor the Purification and Structural Analysis of Membrane Proteins Including G-Protein Coupled Receptors. Nat. Commun 2020, 11 (1), 564. [PubMed: 31992701]

(88). Chorev DS; Baker LA; Wu D; Beilsten-Edmands V; Rouse SL; Zeev-Ben-Mordehai T; Jiko C; Samsudin F; Gerle C; Khalid S; Stewart AG; Matthews SJ; Grunewald K; Robinson CV Protein Assemblies Ejected Directly from Native Membranes Yield Complexes for Mass Spectrometry. Science 2018, 362 (6416), 829–834. [PubMed: 30442809]

(89). Anderson NL; Anderson NG The Human Plasma Proteome: History, Character, and Diagnostic Prospects. Mol. Cell Proteomics 2002, 1 (11), 845–867. [PubMed: 12488461]

(90). Rifai N; Gillette MA; Carr SA Protein Biomarker Discovery and Validation: The Long and Uncertain Path to Clinical Utility. Nat. Biotechnol 2006, 24 (8), 971–983. [PubMed: 16900146]

(91). Xie S; Moya C; Bilgin B; Jayaraman A; Walton SP Emerging Affinity-Based Techniques in Proteomics. Expert Rev. Proteomics 2009, 6 (5), 573–583. [PubMed: 19811078]

(92). Lollo B; Steele F; Gold L Beyond Antibodies: New Affinity Reagents to Unlock the Proteome. Proteomics 2014, 14 (6), 638–644. [PubMed: 24395722]

(93). Uhlen M; Fagerberg L; Hallstrom BM; Lindskog C; Oksvold P; Mardinoglu A; Sivertsson A; Kampf C; Sjostedt E; Asplund A; Olsson I; Edlund K; Lundberg E; Navani S; Szigyarto CA; Odeberg J; Djureinovic D; Takanen JO; Hober S; Alm T; Edqvist P; Berling H; Tegel H; Mulder J; Rockberg J; Nilsson P; Schwenk JM; Hamsten M; von Feilitzen K; Forsbergm M; Persson L; Johansson F; Zwahlenm M; von Heijne G; Nielsen J; Ponten F Tissue-Based Map of the Human Proteome. Science 2015, 347 (6220), 1260419–1260419. [PubMed: 25613900]

(94). Gregorich ZR; Chang YH; Ge Y Proteomics in Heart Failure: Top-down or Bottom-Up? Pfluegers Arch. 2014, 466 (6), 1199–1209. [PubMed: 24619480]

(95). Bauer A; Kuster B Affinity Purification-Mass Spectrometry. Eur. J. Biochem 2003, 270 (4), 570–578. [PubMed: 12581197]

(96). Baker M Blame It on the Antibodies. Nature 2015, 521 (7552), 274–276. [PubMed: 25993940]

(97). Gilda JE; Ghosh R; Cheah JX; West TM; Bodine SC; Gomes AV Western Blotting Inaccuracies with Unverified Antibodies: Need for a Western Blotting Minimal Reporting Standard (WBMRS). PLoS One 2015, 10 (8), e0135392. [PubMed: 26287535]

(98). Bradbury A; Plückthun A Reproducibility: Standardize Antibodies Used in Research. Nature 2015, 518 (7537), 27–29. [PubMed: 25652980]

(99). Janes KA Fragile Epitopes-Antibody's Guess Is as Good as Yours. Sci. Signaling 2020, 13 (616), No. eaaz8130.

(100). Smith JG; Gerszten RE Emerging Affinity-Based Proteomic Technologies for Large-Scale Plasma Profiling in Cardiovascular Disease. Circulation 2017, 135 (17), 1651–1664. [PubMed: 28438806]

(101). Marx V Calling the next Generation of Affinity Reagents. Nat. Methods 2013, 10 (9), 829–833. [PubMed: 23985727]

(102). Hwang L; Ayaz-Guner S; Gregorich ZR; Cai W; Valeja SG; Jin S; Ge Y Specific Enrichment of Phosphoproteins Using Functionalized Multivalent Nanoparticles. J. Am. Chem. Soc 2015, 137 (7), 2432–2435. [PubMed: 25655481]

(103). Roberts DS; Chen B; Tiambeng TN; Wu Z; Ge Y; Jin S Reproducible Large-Scale Synthesis of Surface Silanized Nanoparticles as an Enabling Nanoproteomics Platform: Enrichment of the Human Heart Phosphoproteome. Nano Res. 2019, 12 (6), 1473–1481. [PubMed: 31341559]

(104). Chen B; Hwang L; Ochowicz W; Lin Z; Guardado-Alvarez TM; Cai W; Xiu L; Dani K; Colah C; Jin S; Ge Y Coupling Functionalized Cobalt Ferrite Nanoparticle Enrichment with Online LC/MS/MS for Top-down Phosphoproteomics. Chem. Sci 2017, 8 (6), 4306–4311. [PubMed: 28660060]

(105). Chen Y; Hoover ME; Dang X; Shomo AA; Guan X; Marshall AG; Freitas MA; Young NL Quantitative Mass Spectrometry Reveals That Intact Histone H1 Phosphorylations Are Variant Specific and Exhibit Single Molecule Hierarchical Dependence. Mol. Cell. Proteomics 2016, 15 (3), 818–833. [PubMed: 26209608]

(106). Lin Z; Wei L; Cai W; Zhu Y; Tucholski T; Mitchell SD; Guo W; Ford SP; Diffee GM; Ge Y Simultaneous Quantification of Protein Expression and Modifications by Top-down Targeted Proteomics: A Case of the Sarcomeric Subproteome. Mol. Cell. Proteomics 2019, 18 (3), 594–605.

(107). Ellington AD; Szostak JW In Vitro Selection of RNA Molecules That Bind Specific Ligands. Nature 1990, 346 (6287), 818–822. [PubMed: 1697402]

(108). Tuerk C; Gold L Systematic Evolution of Ligands by Exponential Enrichment: RNA Ligands to Bacteriophage T4 DNA Polymerase. Science 1990, 249 (4968), 505–510. [PubMed: 2200121]

(109). Wang T; Chen C; Larcher LM; Barrero RA; Veedu RN Three Decades of Nucleic Acid Aptamer Technologies: Lessons Learned, Progress and Opportunities on Aptamer Development. Biotechnol. Adv 2019, 37 (1), 28–50. [PubMed: 30408510]

(110). Yang L; Zhang X; Ye M; Jiang J; Yang R; Fu T; Chen Y; Wang K; Liu C; Tan W Aptamer-Conjugated Nanomaterials and Their Applications. Adv. Drug Delivery Rev 2011, 63 (14), 1361–1370.

(111). Bantscheff M; Lemeer S; Savitski MM; Kuster B Quantitative Mass Spectrometry in Proteomics: Critical Review Update from 2007 to the Present. Anal. Bioanal. Chem 2012, 404 (4), 939–965. [PubMed: 22772140]

(112). Doucette AA; Tran JC; Wall MJ; Fitzsimmons S Intact Proteome Fractionation Strategies Compatible with Mass Spectrometry. Expert Rev. Proteomics 2011, 8 (6), 787–800. [PubMed: 22087661]

(113). Zhang Z; Wu S; Stenoien DL; Paša-Tolić L High-Throughput Proteomics. Annu. Rev. Anal. Chem 2014, 7 (1), 427–454.

(114). Washburn MP; Wolters D; Yates JR Large-Scale Analysis of the Yeast Proteome by Multidimensional Protein Identification Technology. Nat. Biotechnol 2001, 19 (3), 242–247. [PubMed: 11231557]

(115). Wang Y; Olesik SV Enhanced-Fluidity Liquid Chromatography-Mass Spectrometry for Intact Protein Separation and Characterization. Anal. Chem 2019, 91 (1), 935–942. [PubMed: 30523683]

(116). Eschelbach JW; Jorgenson JW Improved Protein Recovery in Reversed-Phase Liquid Chromatography by the Use of Ultrahigh Pressures. Anal. Chem 2006, 78 (5), 1697–1706. [PubMed: 16503625]

(117). Liang Y; Jin Y; Wu Z; Tucholski T; Brown KA; Zhang L; Zhang Y; Ge Y Bridged Hybrid Monolithic Column Coupled to High-Resolution Mass Spectrometry for Top-Down Proteomics. Anal. Chem 2019, 91 (3), 1743–1747. [PubMed: 30668094]

(118). Hayes R; Ahmed A; Edge T; Zhang H Core-Shell Particles: Preparation, Fundamentals and Applications in High Performance Liquid Chromatography. J. Chromatogr. A 2014, 1357, 36–52. [PubMed: 24856904]

(119). Astefanei A; Dapic I; Camenzuli M Different Stationary Phase Selectivities and Morphologies for Intact Protein Separations. Chromatographia 2017, 80 (5), 665–687. [PubMed: 28529348]

(120). Picotti P; Bodenmiller B; Mueller LN; Domon B; Aebersold R Full Dynamic Range Proteome Analysis of S. Cerevisiae by Targeted Proteomics. Cell 2009, 138 (4), 795–806. [PubMed: 19664813]

(121). Compton PD; Zamdborg L; Thomas PM; Kelleher NL On the Scalability and Requirements of Whole Protein Mass Spectrometry. Anal. Chem 2011, 83 (17), 6868–6874. [PubMed: 21744800]

(122). Tran JC; Doucette AA Multiplexed Size Separation of Intact Proteins in Solution Phase for Mass Spectrometry. Anal. Chem 2009, 81 (15), 6201–6209. [PubMed: 19572727]

(123). Chen X; Ge Y Ultrahigh Pressure Fast Size Exclusion Chromatography for Top-down Proteomics. Proteomics 2013, 13(17), 2563–2566. [PubMed: 23794208]

(124). Ge Y; Rybakova IN; Xu Q; Moss RL Top-down High-Resolution Mass Spectrometry of Cardiac Myosin Binding Protein C Revealed That Truncation Alters Protein Phosphorylation State. Proc. Natl. Acad. Sci. U. S. A 2009, 106 (31), 12658–12663. [PubMed: 19541641]

(125). Jin Y; Wei L; Cai W; Lin Z; Wu Z; Peng Y; Kohmoto T; Moss RL; Ge Y Complete Characterization of Cardiac Myosin Heavy Chain (223 KDa) Enabled by Size-Exclusion Chromatography and Middle-Down Mass Spectrometry. Anal. Chem 2017, 89 (9), 4922–4930. [PubMed: 28366003]

(126). Tucholski T; Knott SJ; Chen B; Pistono P; Lin Z; Ge Y A Top-Down Proteomics Platform Coupling Serial Size Exclusion Chromatography and Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. Anal. Chem 2019, 91 (6), 3835–3844. [PubMed: 30758949]

(127). García MC The Effect of the Mobile Phase Additives on Sensitivity in the Analysis of Peptides and Proteins by High-Performance Liquid Chromatography-Electrospray Mass Spectrometry. J. Chromatogr. B: Anal. Technol. Biomed. Life Sci 2005, 825 (2), 111–123.

(128). Alpert AJ High-Performance Hydrophobic-Interaction Chromatography of Proteins on a Series of Poly(Alkyl Aspart-Amide)-Silicas. J. Chromatogr. A 1986, 359 (C), 85–97.

(129). Queiroz JA; Tomaz CT; Cabral JMS Hydrophobic Interaction Chromatography of Proteins. J. Biotechnol 2001, 87 (2), 143–159. [PubMed: 11278038]

(130). Chen B; Peng Y; Valeja SG; Xiu L; Alpert AJ; Ge Y Online Hydrophobic Interaction Chromatography-Mass Spectrometry for Top-Down Proteomics. Anal. Chem 2016, 88 (3), 1885–1891. [PubMed: 26729044]

(131). 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 (19), 10138–10145. [PubMed: 26360183]

(132). Mehaffey MR; Xia Q; Brodbelt JS Uniting Native Capillary Electrophoresis and Multistage Ultraviolet Photodissociation Mass Spectrometry for Online Separation and Characterization of Escherichia Coli Ribosomal Proteins and Protein Complexes. Anal. Chem 2020, 92 (22), 15202–15211. [PubMed: 33156608]

(133). Shen X; Kou Q; Guo R; Yang Z; Chen D; Liu X; Hong H; Sun L Native Proteomics in Discovery Mode Using Size-Exclusion Chromatography-Capillary Zone Electrophoresis-Tandem Mass Spectrometry. Anal. Chem 2018, 90 (17), 10095–10099. [PubMed: 30085653]

(134). Jooß K; McGee JP; Melani RD; Kelleher NL Electrophoresis 2021, DOI: 10.1002/elps.202000317.

(135). Shen X; Liang Z; Xu T; Yang Z; Wang Q; Chen D; Pham L; Du W; Sun L Investigating Native Capillary Zone Electrophoresis-Mass Spectrometry on a High-End Quadrupole-Time-of-flight Mass Spectrometer for the Characterization of Monoclonal Antibodies. Int. J. Mass Spectrom 2021, 462, 116541. [PubMed: 33642939]

(136). Li Y; Champion MM; Sun L; Champion PAD; Wojcik R; Dovichi NJ Capillary Zone Electrophoresis-Electrospray Ionization-Tandem Mass Spectrometry as an Alternative Proteomics Platform to Ultraperformance Liquid Chromatography-Electrospray Ionization-Tandem Mass Spectrometry for Samples of Intermediate Complexity. Anal. Chem 2012, 84 (3), 1617–1622. [PubMed: 22182061]

(137). Stolz A; Jooß K; Höcker O; Römer J; Schlecht J; Neusüß C Recent Advances in Capillary Electrophoresis-Mass Spectrometry: Instrumentation, Methodology and Applications. Electrophoresis 2019, 40 (1), 79–112. [PubMed: 30260009]

(138). Fussl F; Trappe A; Carillo S; Jakes C; Bones J Comparative Elucidation of Cetuximab Heterogeneity on the Intact Protein Level by Cation Exchange Chromatography and Capillary Electrophoresis Coupled to Mass Spectrometry. Anal. Chem 2020, 92 (7), 5431–5438. [PubMed: 32105056]

(139). Mack S; Arnold D; Bogdan G; Bousse L; Danan L; Dolnik V; Ducusin M; Gwerder E; Herring C; Jensen M; Ji J; Lacy S; Richter C; Walton I; Gentalen E A Novel Microchip-based Imaged CIEF-MS System for Comprehensive Characterization and Identification of Biopharmaceutical Charge Variants. Electrophoresis 2019, 40 (23–24), 3084–3091. [PubMed: 31663138]

(140). Baker ES; Burnum-Johnson KE; Ibrahim YM; Orton DJ; Monroe ME; Kelly RT; Moore RJ; Zhang X; Théberge R; Costello CE; Smith RD Enhancing Bottom-up and Top-down Proteomic Measurements with Ion Mobility Separations. Proteomics 2015, 15 (16), 2766–2776. [PubMed: 26046661]

(141). Zinnel NF; Pai PJ; Russell DH Ion Mobility-Mass Spectrometry (IM-MS) for Top-down Proteomics: Increased Dynamic Range Affords Increased Sequence Coverage. Anal. Chem 2012, 84 (7), 3390–3397. [PubMed: 22455956]

(142). Nshanian M; Lantz C; Wongkongkathep P; Schrader T; Klärner F-G; Blümke A; Despres C; Ehrmann M; Smet-Nocca C; Bitan G; Loo JA Native Top-Down Mass Spectrometry and Ion Mobility Spectrometry of the Interaction of Tau Protein with a Molecular Tweezer Assembly Modulator. J. Am. Soc. Mass Spectrom 2019, 30 (1), 16–23. [PubMed: 30062477]

(143). Dodds JN; Baker ES Ion Mobility Spectrometry: Fundamental Concepts, Instrumentation, Applications, and the Road Ahead. J. Am. Soc. Mass Spectrom 2019, 30 (11), 2185–2195. [PubMed: 31493234]

(144). Polasky DA; Dixit SM; Keating MF; Gadkari VV; Andrews PC; Ruotolo BT Pervasive Charge Solvation Permeates Native-like Protein Ions and Dramatically Influences Top-down Sequencing Data. J. Am. Chem. Soc 2020, 142 (14), 6750–6760. [PubMed: 32203657]

(145). Pirok BWJ; Stoll DR; Schoenmakers PJ Recent Developments in Two-Dimensional Liquid Chromatography: Fundamental Improvements for Practical Applications. Anal. Chem 2019, 91 (1), 240–263. [PubMed: 30380827]

(146). Xiu L; Valeja SG; Alpert AJ; Jin S; Ge Y Effective Protein Separation by Coupling Hydrophobic Interaction and Reverse Phase Chromatography for Top-down Proteomics. Anal. Chem 2014, 86 (15), 7899–7906. [PubMed: 24968279]

(147). Wang Z; Ma H; Smith K; Wu S Two-Dimensional Separation Using High-PH and Low-PH Reversed Phase Liquid Chromatography for Top-down Proteomics. Int. J. Mass Spectrom 2018, 427, 43–51. [PubMed: 31097918]

(148). Valeja SG; Xiu L; Gregorich ZR; Guner H; Jin S; Ge Y Three Dimensional Liquid Chromatography Coupling Ion Exchange Chromatography/Hydrophobic Interaction Chromatography/Reverse Phase Chromatography for Effective Protein Separation in Top-down Proteomics. Anal. Chem 2015, 87 (10), 5363–5371. [PubMed: 25867201]

(149). Gargano AFG; Shaw JB; Zhou M; Wilkins CS; Fillmore TL; Moore RJ; Somsen GW; Paša-Tolić L Increasing the Separation Capacity of Intact Histone Proteoforms Chromatography Coupling Online Weak Cation Exchange-HILIC to Reversed Phase LC UVPD-HRMS. J. Proteome Res 2018, 17 (11), 3791–3800. [PubMed: 30226781]

(150). Baghdady YZ; Schug KA Online Comprehensive High PH Reversed Phase × Low PH Reversed Phase Approach for Two-Dimensional Separations of Intact Proteins in Top-Down Proteomics. Anal. Chem 2019, 91 (17), 11085–11091. [PubMed: 31366196]

(151). Moore AW; Jorgenson JW Comprehensive Three-Dimensional Separation of Peptides Using Size Exclusion Chromatography/Reversed Phase Liquid Chromatography/Optically Gated Capillary Zone Electrophoresis. Anal. Chem 1995, 67 (19), 3456–3463. [PubMed: 8686894]

(152). Guner H; Close PL; Cai W; Zhang H; Peng Y; Gregorich ZR; Ge Y MASH Suite: A User-Friendly and Versatile Software Interface for High-Resolution Mass Spectrometry Data Interpretation and Visualization. J. Am. Soc. Mass Spectrom 2014, 25 (3), 464–470. [PubMed: 24385400]

(153). Meng F; Cargile BJ; Miller LM; Forbes AJ; Johnson JR; Kelleher NL Informatics and Multiplexing of Intact Protein Identification in Bacteria and the Archaea. Nat. Biotechnol 2001, 19 (10), 952–957. [PubMed: 11581661]

(154). Zamdborg L; LeDuc RD; Glowacz KJ; Kim Y-B; Viswanathan V; Spaulding IT; Early BP; Bluhm EJ; Babai S; Kelleher NL ProSight PTM 2.0: Improved Protein Identification and Characterization for Top down Mass Spectrometry. Nucleic Acids Res. 2007, 35, W701. [PubMed: 17586823]

(155). LeDuc RD; Taylor GK; Kim Y-B; Januszyk TE; Bynum LH; Sola JV; Garavelli JS; Kelleher NL ProSight PTM: An Integrated Environment for Protein Identification and Characterization by Top-down Mass Spectrometry. Nucleic Acids Res. 2004, 32, W340–350. [PubMed: 15215407]

(156). Liu X; Sirotkin Y; Shen Y; Anderson G; Tsai YS; Ting YS; Goodlett DR; Smith RD; Bafna V; Pevzner PA Protein Identification Using Top-Down. Mol. Cell. Proteomics 2012, 11 (6), M111.008524.

(157). Kou Q; Xun L; Liu X TopPIC: A Software Tool for Top-down Mass Spectrometry-Based Proteoform Identification and Characterization. Bioinformatics 2016, 32 (22), 3495–3497. [PubMed: 27423895]

(158). Park J; Piehowski PD; Wilkins C; Zhou M; Mendoza J; Fujimoto GM; Gibbons BC; Shaw JB; Shen Y; Shukla AK; Moore RJ; Liu T; Petyuk VA; Tolic N; Pasa-Tolic L; Smith RD; Payne SH; Kim S Informed-Proteomics: Open-Source Software Package for Top-down Proteomics. Nat. Methods 2017, 14 (9), 909–914. [PubMed: 28783154]

(159). Sun RX; Luo L; Wu L; Wang RM; Zeng WF; Chi H; Liu C; He SM PTop 1.0: A High-Accuracy and High-Efficiency Search Engine for Intact Protein Identification. Anal. Chem 2016, 88 (6), 3082–3090. [PubMed: 26844380]

(160). Cesnik AJ; Shortreed MR; Schaffer LV; Knoener RA; Frey BL; Scalf M; Solntsev SK; Dai Y; Gasch AP; Smith LM Proteoform Suite: Software for Constructing, Quantifying, and Visualizing Proteoform Families. J. Proteome Res 2018, 17 (1), 568–578. [PubMed: 29195273]

(161). Cai W; Guner H; Gregorich ZR; Chen AJ; Ayaz-Guner S; Peng Y; Valeja SG; Liu X; Ge Y MASH Suite pro: A Comprehensive Software Tool for Top-down Proteomics. Mol. Cell. Proteomics 2016, 15 (2), 703–714. [PubMed: 26598644]

(162). McIlwain SJ; Wu Z; Wetzel M; Belongia D; Jin Y; Wenger K; Ong IM; Ge Y Enhancing Top-Down Proteomics Data Analysis by Combining Deconvolution Results through a Machine Learning Strategy. J. Am. Soc. Mass Spectrom 2020, 31 (5), 1104–1113. [PubMed: 32223200]

(163). Jeong K; Kim J; Gaikwad M; Hidayah SN; Heikaus L; Schlüter H; Kohlbacher O FLASHDeconv: Ultrafast, High-Quality Feature Deconvolution for Top-Down Proteomics. Cell Syst. 2020, 10 (2), 213–218. [PubMed: 32078799]

(164). Marty MT; Baldwin AJ; Marklund EG; Hochberg GKA; Benesch JLP; Robinson CV Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles. Anal. Chem 2015, 87 (8), 4370–4376. [PubMed: 25799115]

(165). Horn DM; Ge Y; McLafferty FW Activated Ion Electron Capture Dissociation for Mass Spectral Sequencing of Larger (42 kDa) Proteins. Anal. Chem 2000, 72 (20), 4778–4784. [PubMed: 11055690]

(166). Riley NM; Westphall MS; Coon JJ Activated Ion-Electron Transfer Dissociation Enables Comprehensive Top-Down Protein Fragmentation. J. Proteome Res 2017, 16 (7), 2653–2659. [PubMed: 28608681]

(167). Shaw JB; Li W; Holden DD; Zhang Y; Griep-Raming J; Fellers RT; Early BP; Thomas PM; Kelleher NL; Brodbelt JS Complete Protein Characterization Using Top-down Mass Spectrometry and Ultraviolet Photodissociation. J. Am. Chem. Soc 2013, 135 (34), 12646–12651. [PubMed: 23697802]

(168). Fung YME; Adams CM; Zubarev RA Electron Ionization Dissociation of Singly and Multiply Charged Peptides. J. Am. Chem. Soc 2009, 131 (29), 9977–9985. [PubMed: 19621955]

(169). Zenaidee MA; Lantz C; Perkins T; Jung W; Loo RRO; Loo JA Internal Fragments Generated by Electron Ionization Dissociation Enhance Protein Top-Down Mass Spectrometry. J. Am. Soc. Mass Spectrom 2020, 31 (9), 1896–1902. [PubMed: 32799534]

(170). Smith LM; Agar JN; Chamot-Rooke J; Danis PO; Ge Y; Loo JA; Paša-Tolić L; Tsybin YO; Kelleher NL The Human Proteoform Project: Bringing Proteoforms to Life A Plan to Define the Human Proteome. Preprints 2020, 2020100368 DOI: 10.20944/preprints202010.0368.v1.

(171). Carel C; Marcoux J; Réat V; Parra J; Latgé G; Laval F; Demange P; Burlet-Schiltz O; Milon A; Daffé M; Tropis MG; Renault MA Identification of Specific Posttranslational O-Mycoloylations Mediating Protein Targeting to the Mycomembrane. Proc. Natl. Acad. Sci. U. S. A 2017, 114 (16), 4231–4236. [PubMed: 28373551]

(172). Schmitt ND; Agar JN Parsing Disease-Relevant Protein Modifications from Epiphenomena: Perspective on the Structural Basis of SOD1-Mediated ALS. J. Mass Spectrom 2017, 52 (7), 480–491. [PubMed: 28558143]

(173). Wei L; Gregorich ZR; Lin Z; Cai W; Jin Y; McKiernan SH; McIlwain S; Aiken JM; Moss RL; Diffee GM; Ge Y Novel Sarcopenia-Related Alterations in Sarcomeric Protein Post-Translational Modifications (PTMs) in Skeletal Muscles Identified by Top-down Proteomics. Mol. Cell. Proteomics 2018, 17 (1), 134–145. [PubMed: 29046390]

(174). Hood L; Friend SH Predictive, Personalized, Preventive, Participatory (P4) Cancer Medicine. Nat. Rev. Clin. Oncol 2011, 8 (3), 184–187. [PubMed: 21364692]

(175). Van Eyk JE; Snyder MP Precision Medicine: Role of Proteomics in Changing Clinical Management and Care. J. Proteome Res 2018, 18 (1), 1–6. [PubMed: 30296097]

(176). Karczewski KJ; Snyder MP Integrative Omics for Health and Disease. Nat. Rev. Genet 2018, 19 (5), 299–310. [PubMed: 29479082]

(177). Chen R; Snyder M Promise of Personalized Omics to Precision Medicine. Wiley Interdiscip. Rev. Syst. Biol. Med 2013, 5 (1), 73–82. [PubMed: 23184638]

(178). Gajadhar AS; White FM System Level Dynamics of Post-Translational Modifications. Curr. Opin. Biotechnol 2014, 28, 83–87. [PubMed: 24441143]

(179). Krueger KE; Srivastava S Posttranslational Protein Modifications: Current Implications for Cancer Detection, Prevention, and Therapeutics. Mol. Cell. Proteomics 2006, 5 (10), 1799–1810. [PubMed: 16844681]

(180). Mann M; Jensen ON Proteomic Analysis of Post-Translational Modifications. Nat. Biotechnol 2003, 21 (3), 255–261. [PubMed: 12610572]

(181). Toby TK; Abecassis M; Kim K; Thomas PM; Fellers RT; LeDuc RD; Kelleher NL; Demetris J; Levitsky J Proteoforms in Peripheral Blood Mononuclear Cells as Novel Rejection Biomarkers in Liver Transplant Recipients. Am. J. Transplant 2017, 17 (9), 2458–2467. [PubMed: 28510335]

(182). Chen YC; Ayaz-Guner S; Peng Y; Lane NM; Locher MR; Kohmoto T; Larsson L; Moss RL; Ge Y Effective Top-Down LC/MS+ Method for Assessing Actin Isoforms as a Potential Cardiac Disease Marker. Anal. Chem 2015, 87 (16), 8399–8406. [PubMed: 26189812]

(183). Maron MS; Maron BJ Seminar Hypertrophic Cardiomyopathy. Lancet 2013, 381, 242–255. [PubMed: 22874472]

(184). Madian AG; Rochelle NS; Regnier FE Mass-Linked Immuno-Selective Assays in Targeted Proteomics. Anal. Chem 2013, 85 (2), 737–748. [PubMed: 22950521]

(185). Fu Q; Johnson CW; Wijayawardena BK; Kowalski MP; Kheradmand M; Van Eyk JE A Plasma Sample Preparation for Mass Spectrometry Using an Automated Workstation. J. Visualized Exp 2020, e59842.

(186). Van Den Broek I; Mastali M; Mastali M; Mouapi K; Bystrom C; Bairey Merz CN; Van Eyk JE; Van Eyk JE; Van Eyk JE Quality Control and Outlier Detection of Targeted Mass Spectrometry Data from Multiplex Protein Panels. J. Proteome Res 2020, 19 (6), 2278–2293. [PubMed: 32270680]

(187). Abbatiello SE; Schilling B; Mani DR; Zimmerman LJ; Hall SC; MacLean B; Albertolle M; Allen S; Burgess M; Cusack MP; Gosh M; Hedrick V; Held JM; Inerowicz HD; Jackson A; Keshishian H; Kinsinger CR; Lyssand J; Makowski L; Mesri M; Rodriguez H; Rudnick P; Sadowski P; Sedransk N; Shaddox K; Skates SJ; Kuhn E; Smith D; Whiteaker JR; Whitwell C; Zhang S; Borchers CH; Fisher SJ; Gibson BW; Liebler DC; MacCoss MJ; Neubert TA; Paulovich AG; Regnier FE; Tempst P; Carr SA Large-Scale Interlaboratory Study to Develop, Analytically Validate and Apply Highly Multiplexed, Quantitative Peptide Assays to Measure Cancer-Relevant Proteins in Plasma. Mol. Cell. Proteomics 2015, 14 (9), 2357–2374. [PubMed: 25693799]

(188). Nedelkov D; Niederkofler EE; Oran PE; Peterman S; Nelson RW Top-down Mass Spectrometric Immunoassay for Human Insulin and Its Therapeutic Analogs. J. Proteomics 2018, 175, 27–33. [PubMed: 28780057]

(189). Brunner A-D; Thielert M; Vasilopoulou C; Ammar C; Coscia F; Mund A; Horning OB; Bache N; Apalategui A; Lubeck M; Raether O; Park MA; Richter S; Fischer DS; Theis FJ; Meier F; Mann M Ultra-High Sensitivity Mass Spectrometry Quantifies Single-Cell Proteome Changes upon Perturbation. 2020, 2020.12.22.423933. biorxiv.org e-Print archive. https://www.biorxiv.org/content/10.1101/2020.12.22.423933v1.

(190). Niu L; Thiele M; Geyer PE; Rasmussen DN; Webel HE; Santos A; Gupta R; Meier F; Strauss M; Kjaergaard M; Lindvig K; Jacobsen S; Rasmussen S; Hansen T; Krag A; Mann M A Paired Liver Biopsy and Plasma Proteomics Study Reveals Circulating Biomarkers for Alcohol-Related Liver Disease. 2020, 2020.10.16.337592. biorxiv.org e-Print archive. https://www.biorxiv.org/content/10.1101/2020.10.16.337592v1.

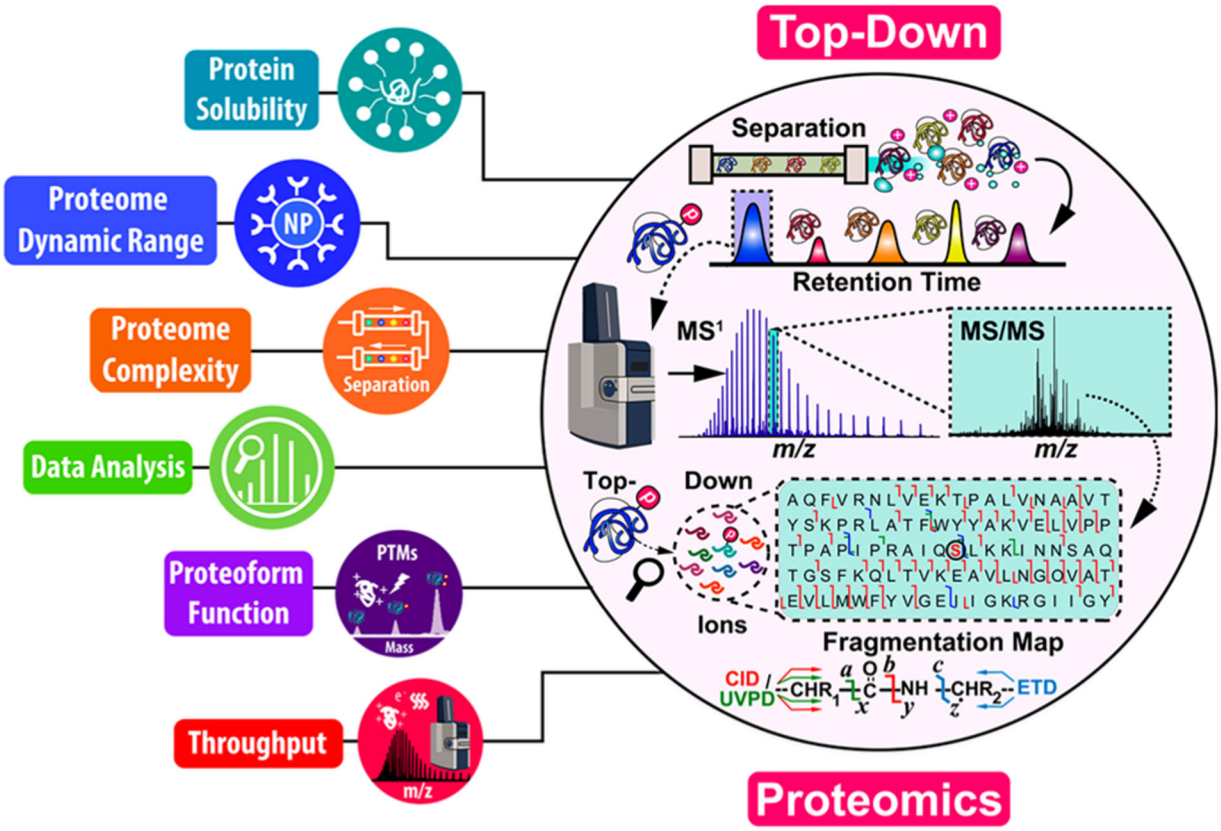


Figure 1.

Addressing the challenges in top-down proteomics through novel strategies. Illustration of the current major challenges in top-down proteomics. These challenges include protein solubility, proteome dynamic range, proteome complexity, intact protein data analysis, proteoform–function relationships, and analytical throughput.

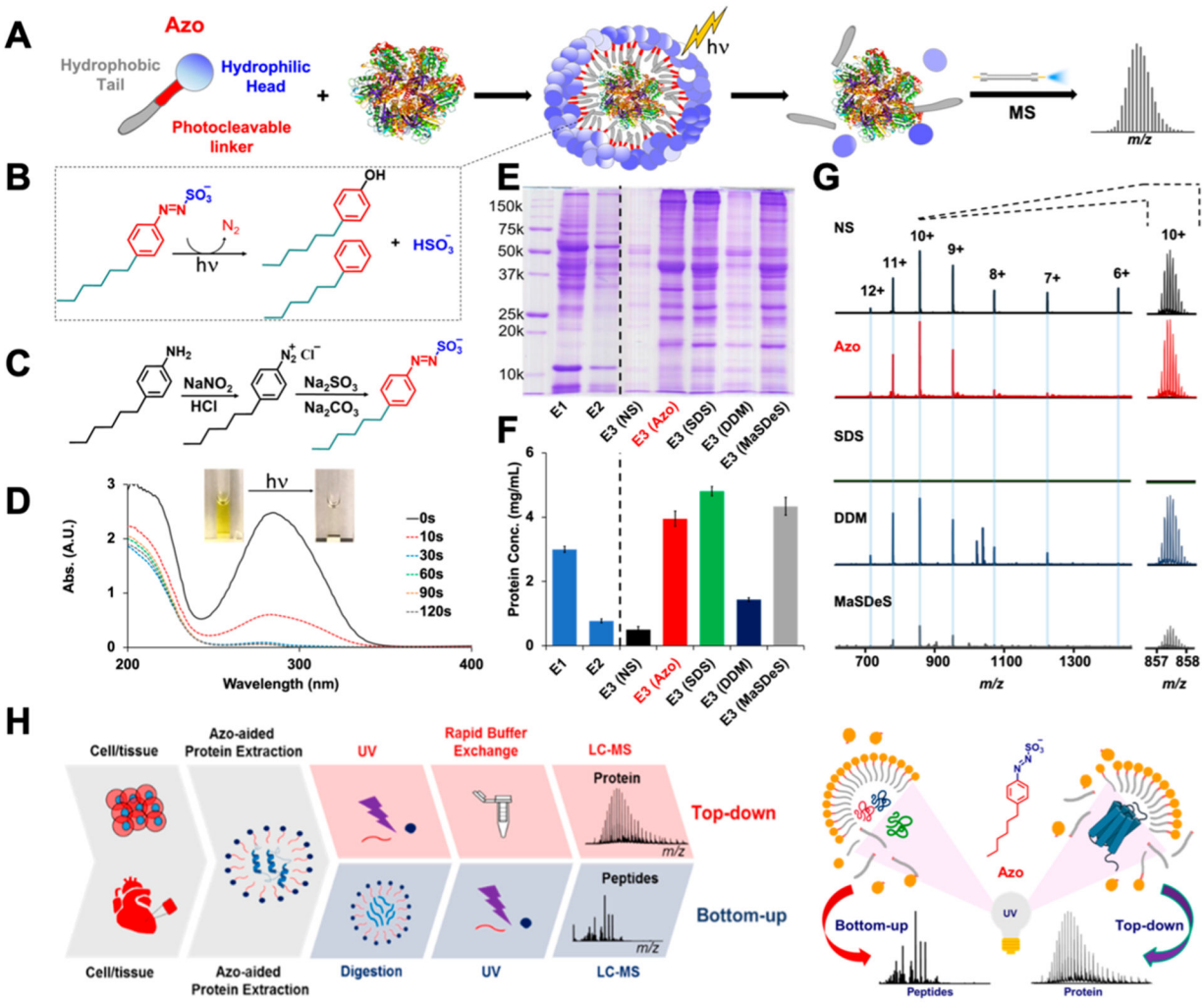


Figure 2.

A novel photocleavable surfactant to address protein solubility challenge. (A) Scheme illustrating the use of Azo in solubilizing proteins, followed by rapid degradation with UV irradiation, and MS analysis of the intact proteins. Note that the molecules are not drawn to scale. (B) Degradation of Azo into 4-hexylphenol, 4-hexylbenzene, nitrogen, and hydrogen sulfate under UV irradiation. (C) Synthetic scheme for Azo. (D) UV–Vis spectra of Azo (0.1%) degradation as a function of time, showing that Azo can be rapidly degraded by UV irradiation at ambient temperature. SDS–PAGE analysis (E) and protein assay (F) for the evaluation of effectiveness of surfactant-aided protein extractions (E3) following the initial HEPES buffer extractions (E1 and E2) to deplete the cytosolic proteins from the cardiac tissue. NS, no surfactant (serving as a control). Error bars represent the standard error of measurement for protein assay experiments ($n = 3$). (G) ESI–MS analysis of ubiquitin with 0.1% surfactant to investigate MS compatibility. The mass spectra were normalized to an intensity of 1.7×10^6. Data are representative of three independent experiments. (H) Scheme for Azo-enabled high-throughput top-down and bottom-up proteomics. Figure adapted from refs 59 and 73. Copyright 2019 Springer Nature (A–G). Copyright 2020 John Wiley & Sons, Inc. (H).

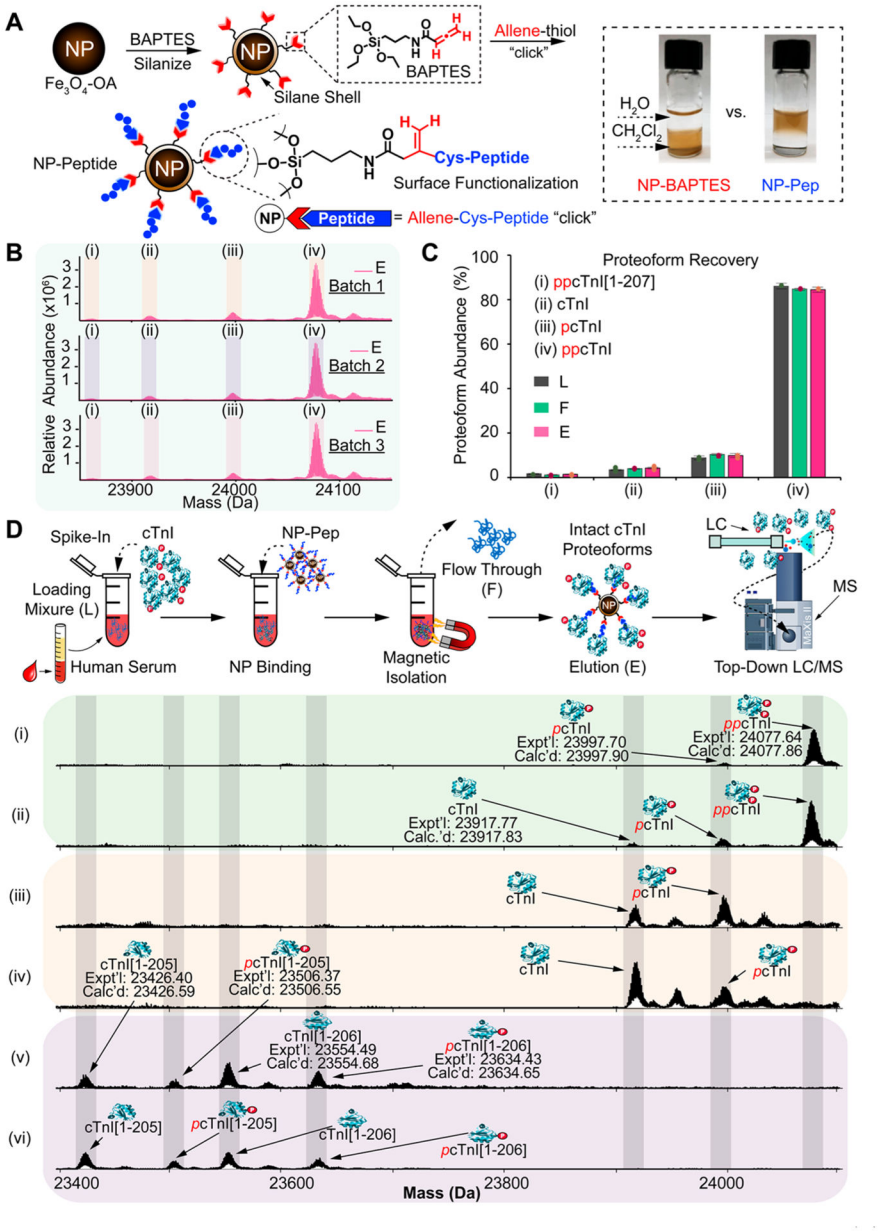


Figure 3.

Nanoproteomics enables proteoform-resolved analysis of low-abundance cardiac troponin I in human serum. (A) Silanization of Fe_3O_4 NPs using an allene carboxamide-based organosilane monomer (BAPTES) for cysteine thiol-specific bioconjugation. The rationally designed NPs are surface functionalized with a 13-mer peptide that has a high affinity for cTnI (NP-Pep) for cTnI enrichment. (B) Top-down MS-based evaluation of cTnI enrichment using three different synthetic batches of NP-Pep showing the reproducible enrichment performance. (C) Evaluation of total cTnI proteoform recovery. The deconvoluted top-down mass spectra corresponding to cTnI proteoforms were used to calculate the relative abundance of each cTnI proteoform when normalizing for total protein amount injected. Proteoform abundance data are representative of $n = 6$ independent experiments with error

bars indicating the standard error of the mean. cTnI relative abundance data. Roman numerals correspond to N-terminally acetylated cTnI proteoforms following Met exclusion: (i) ppcTnI[1–207]; (ii) cTnI; (iii) pcTnI; (iv) ppcTnI; (v) cTnI[1–205]; (vi) pcTnI[1–205]; (vii) cTnI[1–206]; (viii) pcTnI[1–206]. (D) Nanoproteomics assay utilizing NP-Pep for specific enrichment of cTnI from serum and subsequent top-down MS analysis of cTnI proteoforms. cTnI is first spiked into human serum to prepare the loading mixture (L). The NPs are then incubated with the serum loading mixture, the cTnI-bound NPs are magnetically isolated, and the unwanted and nonspecific proteins are removed as flow-through (F). The captured cTnI is then eluted, and the final elution fraction after enrichment is analyzed by top-down LC-MS/MS. The cTnI (~10–20 ng/mL) spiked in the human serum (10 mg) were extracted from various human hearts: (i) and (ii), donor hearts; (iii) and (iv), diseased hearts with dilated cardiomyopathy, (v) and (vi), post-mortem hearts. p, phosphorylation. pp, bisphosphorylation. Figure adapted from ref 30. Copyright 2020 Springer Nature.

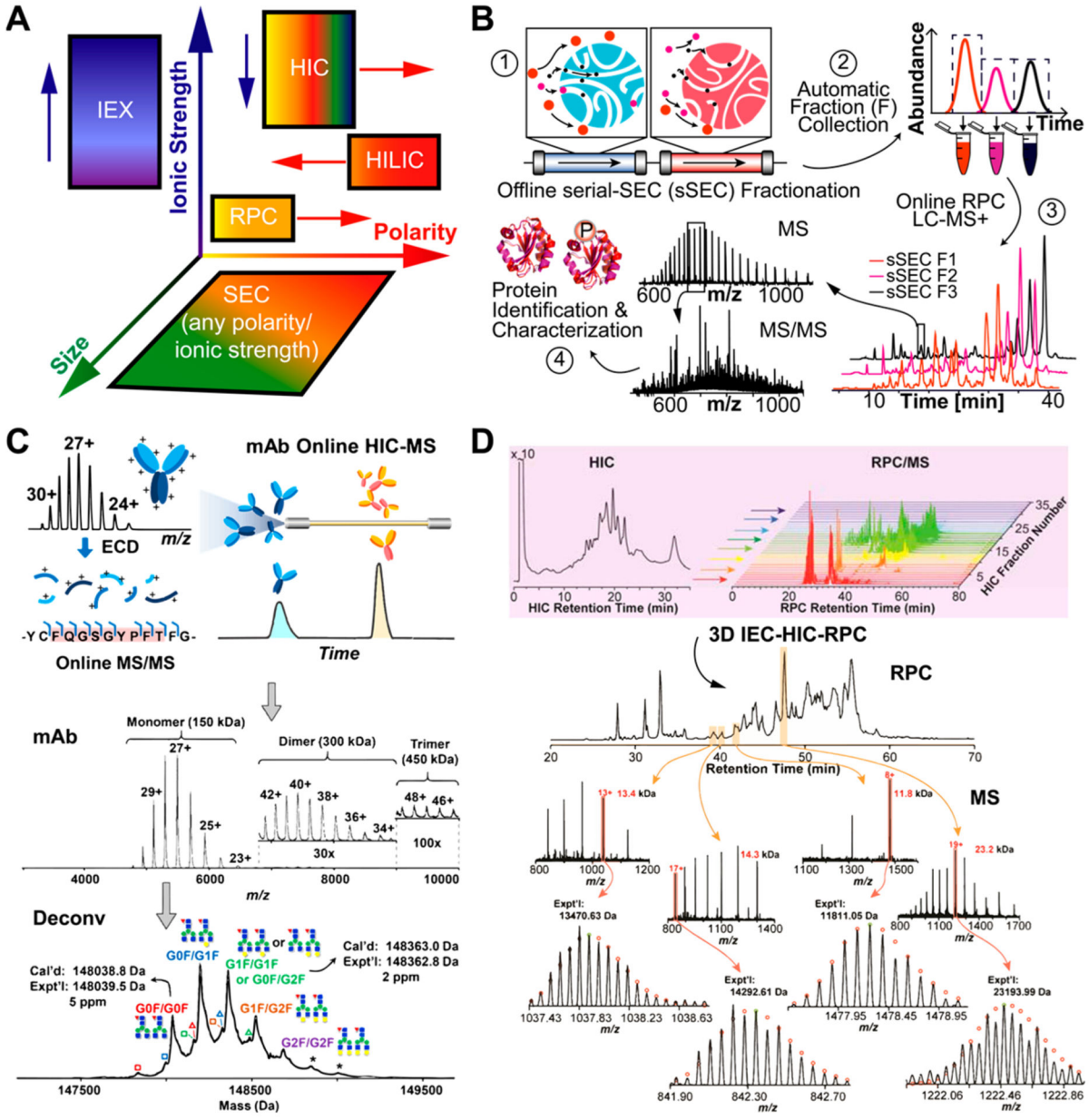


Figure 4.

Novel separation strategies to address proteome complexity. (A) Comparison of the various protein separation techniques. Ion-exchange chromatography (IEX or IEC), hydrophobic interaction chromatography (HIC), reverse phase chromatography (RPC), size-exclusion chromatography (SEC), and hydrophilic interaction chromatography (HILIC) are shown as representative separation techniques. (B) Overview of the serial size-exclusion chromatography (sSEC) strategy for complex protein mixture separation. (1) Protein mixtures pass through two sSEC columns (1000–500 Å) to separate the proteins by high (1000 Å) and low (500 Å) molecular weights, (2) specific protein fractions are collected, (3) sample fractions are analyzed using online RPC-MS, (4) individual proteins are characterized. Figure adapted from ref 29. Copyright 2017 American Chemical Society. (C)

Online HIC-MS of mAb mixtures on a maXis II Q-TOF mass spectrometer. Mass spectrum of a specific mAb (mAb2) showed the detection of monomers, dimers (30× zoom-in), and trimers (100× zoom-in). A deconvoluted mass spectrum of the mAb2 monomer is shown with annotated glycosylation forms (red triangle, fucose; blue square, GlcNAc; green circle, mannose; yellow circle, galactose); a hollow square represents the loss of one GlcNAc (−203 Da), a hollow triangle represents the preservation of C-terminal Lys on the heavy chain (+128 Da), and an asterisk represents the addition of a hexose (+162 Da). GxF indicates Fc-oligosaccharides terminated by *x* number of galactoses. Figure adapted from ref 23. Copyright 2018 American Chemical Society. (D) Illustration of the 3D IEC (or IEX) /HIC/RPC-MS/MS separation strategy. HIC was employed as a second dimension of separation (following IEX) prior to top-down RPC-MS/MS analysis. Online RPC-MS analysis of intact proteins from a HEK293 cell lysate following IEC-HIC fractionations (3DLC approach) with representative RPC/MS results of 3DLC (IEC-HIC-RPC) is shown. Figure adapted from ref 148. Copyright 2015 American Chemical Society.

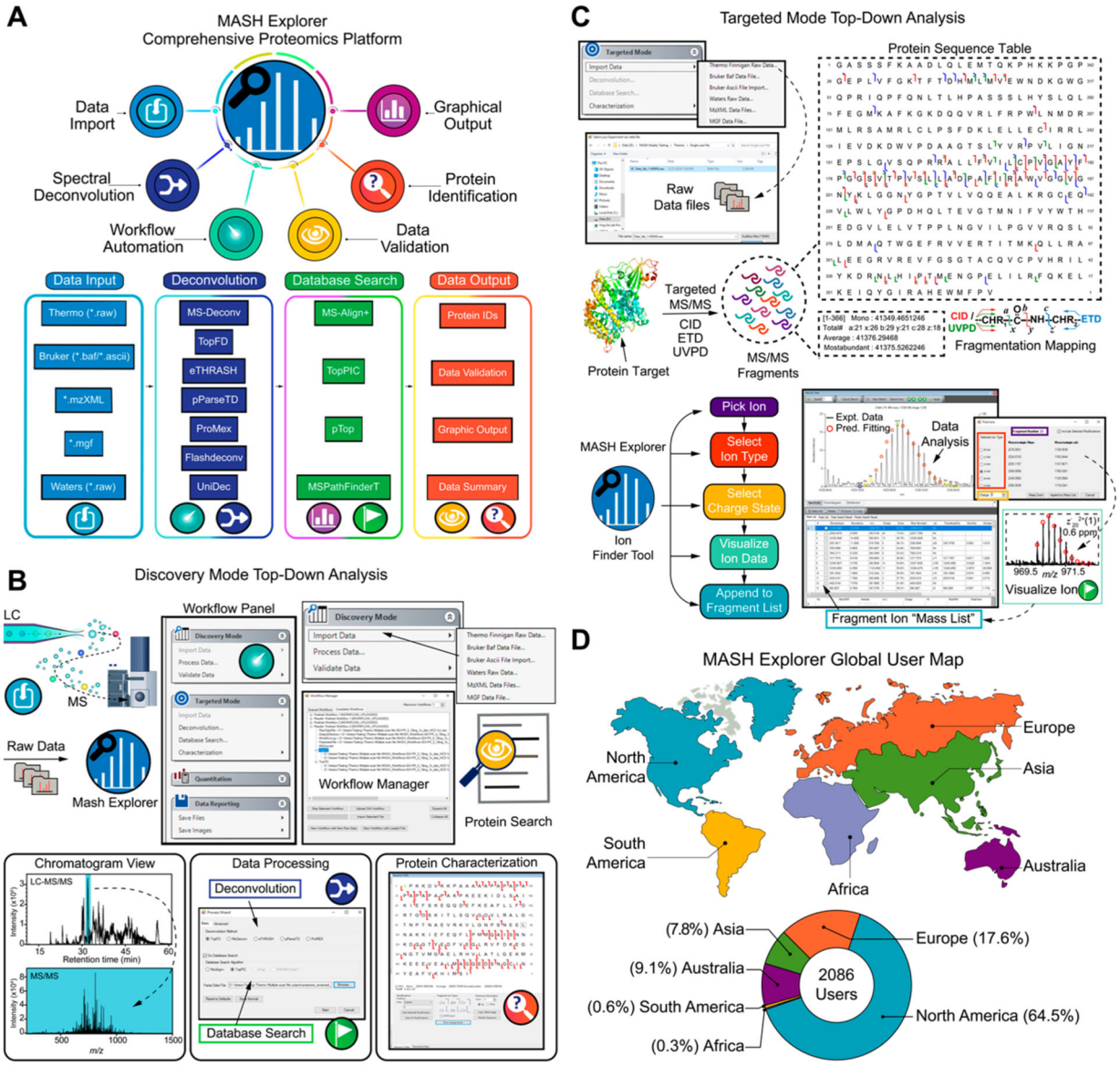


Figure 5.

MASH Explorer as a universal software environment for top-down proteomics. (A) Schematic of the various MASH Explorer functions for proteomics data processing. The main functions of MASH Explorer include data import, spectral deconvolution, workflow automation, data validation, protein identification, and graphical output. MASH Explorer utilizes a new data processing module based on the ProteoWizard Library to accept various data input file formats from major instrument vendors (e.g., Thermo, Bruker, and Waters). Raw MS and MS/MS data files are then processed by deconvolution algorithms (i.e., MS-Deconv, TopFD, eTHRASH, pParseTD, Flashdeconv, and UniDec) and database search algorithms (i.e., MS-Align+, TopPIC, pTop, and MSPathFinderT). (B) Illustration of "Discovery mode" for LC-MS/MS data processing. "Discovery mode" can handle batch LC-MS/MS raw data files and includes features such as data import, data processing (deconvolution and database search), and data validation for protein identification. (C)

Illustration of the “Targeted Mode” workflow for MASH Explorer. The “Targeted Mode” workflow includes data import, spectral deconvolution to identify and verify isotopic distributions, database search based on identified isotopic distributions, and proteoform characterization by matching identified isotopic distributions to the target proteoform sequence. (D) Cartoon schematic of a “world map” featuring the location distribution of MASH users across the globe. There are currently 2086 active users (March 1, 2021) with ~65% of users from North America, ~18% from Europe, and ~8% from Asia. Figure updated and adapted from ref 31. Copyright 2020 American Chemical Society.

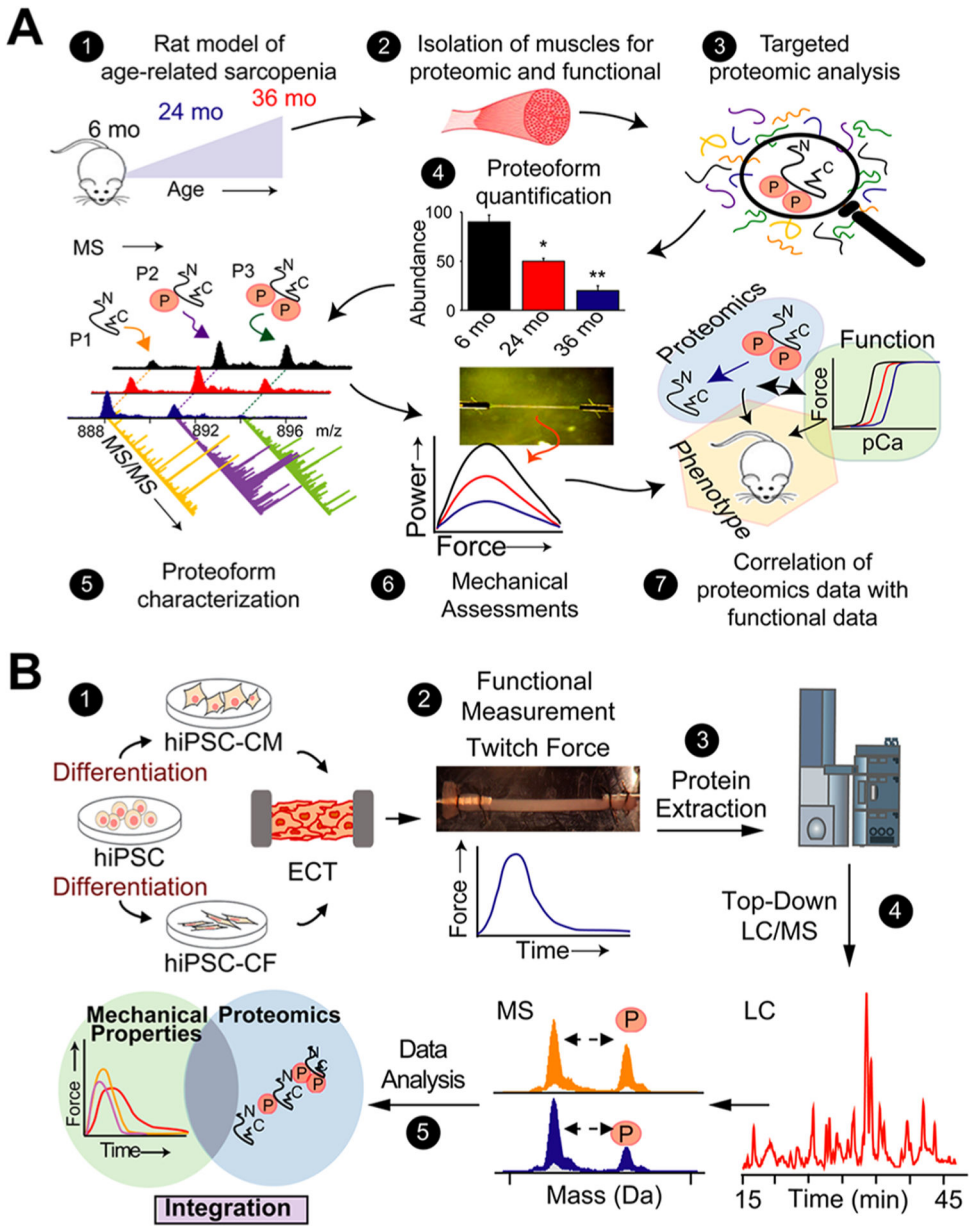


Figure 6.

Linking protoeoforms with biological function using top-down proteomics. (A) Integrated approach combining top-down targeted proteomics with mechanical measurements to elucidate the molecular mechanism(s) underlying age-related sarcopenia. This approach includes the following: (1) use of a rat model of age-related sarcopenia; (2) isolation of skeletal muscle for proteomic and mechanical analyses; (3) top-down targeted proteomics for RLC proteoform analysis; (4) MS-based proteoform quantification; (5) MS/MS analysis for the comprehensive characterization of RLC proteoform sequences and PTMs; (6) mechanical measurements on single fibers; and (7) correlation of the targeted proteomics data with functional data to explain the sarcopenic phenotype. mo, month. Figure adapted from ref 39. Copyright 2016 American Chemical Society. (B) Schematic of integrated

functional assessments and the top-down proteomics workflow for the same hiPSC-ECT. (1) hiPSCs are differentiated into CMs and CFs which are used to generate hiPSC-ECTs. (2) Functional assessments are performed on the hiPSC-ECTs to measure the isometric twitch force. (3) Sarcomeric proteins are extracted via a dual extraction method from the functionally tested hiPSC-ECTs. (4) Top-down proteomics is performed on the functionally tested hiPSC-ECTs. (5) Integrated assessment of hiPSC-ECT constructs. Figure adapted from ref 45. Copyright 2021 American Chemical Society.