

Sample preparation and cleanup methods for clinical top-down proteomics

Amal Mohamed Kamal, Amy Carfagno, Cynthia Nagy & Luca Fornelli

To cite this article: Amal Mohamed Kamal, Amy Carfagno, Cynthia Nagy & Luca Fornelli (2025) Sample preparation and cleanup methods for clinical top-down proteomics, Expert Review of Proteomics, 22:9, 329-343, DOI: [10.1080/14789450.2025.2539913](https://doi.org/10.1080/14789450.2025.2539913)

To link to this article: <https://doi.org/10.1080/14789450.2025.2539913>



© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 29 Jul 2025.



Submit your article to this journal [↗](#)



Article views: 1522



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 5 View citing articles [↗](#)

REVIEW



Sample preparation and cleanup methods for clinical top-down proteomics

Amal Mohamed Kamal^{a,b}, Amy Carfagno^c, Cynthia Nagy^c and Luca Fornelli^{a,c}

^aDepartment of Chemistry and Biochemistry, University of Oklahoma, Norman, OK, USA; ^bFaculty of Pharmacy, Cairo University, Cairo, Egypt;

^cSchool of Biological Sciences, University of Oklahoma, Norman, OK, USA

ABSTRACT

Introduction: The investigation of different proteoforms in clinical samples is a promising approach to elucidate the molecular mechanisms of diseases. Furthermore, proteoform analysis holds great potential for identifying disease-specific biomarkers and targets for personalized medicine. Despite advances in top-down proteomics (TDP) instrumentation, sample preparation and cleanup remain challenging. Work in this area has focused on developing rapid, cost-effective, and less-labor-intensive protocols aimed at minimizing the introduction of artefactual modifications to endogenous proteoforms or bias in proteoform recovery during sample processing.

Area covered: To inform the selection of sample processing approaches in clinical TDP, this review summarizes state-of-the-art targeted (i.e. affinity and non-affinity-based enrichment) and untargeted (i.e. gel-based fractionation) sample preparation protocols. In addition, currently available offline and online sample cleanup procedures (e.g. dialysis, solid-phase extraction, filter-aided sample preparation, precipitation, and solid-phase protein preparation) are reviewed, highlighting their effectiveness for desalting and/or detergent removal.

Expert opinion: TDP demonstrates great potential in the clinical setting due to its ability to capture disease-specific proteoforms commonly overlooked in traditional diagnostic assays. The establishment of standardized guidelines for reproducible clinical TDP workflows is essential to leverage advances in sample preparation techniques and analytical instrumentation to facilitate wider adoption of TDP for clinical applications.

ARTICLE HISTORY

Received 22 May 2025

Accepted 7 July 2025

KEYWORDS

Clinical proteomics; Filter-aided sample preparation; Nanoparticle; Precipitation; serial size exclusion chromatography; single-pot solid phase-enhanced sample preparation; Solid-phase extraction; Top-down proteomics

1. Introduction

Analysis of clinical samples ranging from tissue biopsy to body fluids have identified biomarkers associated with disease progression, a number of which have been FDA-approved for clinical practice [1–3]. Examples include carcinoembryonic antigen, alpha-fetoprotein for testicular cancer, and breast cancer susceptibility gene 1 in breast cancer, identified via immunological assays [4]. For protein biomarker identification in particular, aptamer- and antibody-based arrays have been commercialized as the SomaLogic and Olink platforms, enabling high-throughput, targeted protein identification and quantification. These platforms rely on the specificity of aptamer or antibody interaction with a target protein epitope. However, the potential for protein cross-reactivity and inability of the assays to distinguish between protein structural variants have raised questions regarding their reproducibility and sensitivity for clinical practice [5,6].

The ability to distinguish between different protein forms, or proteoforms, is important in a clinical context, since these may be more closely associated with differences in disease phenotype than generic protein families. Proteoforms arise from different combinations of structural modifications at the nucleic acid level, such as genetic mutation of the protein-coding gene or alternative mRNA splicing, as well as from co- or post-translational modifications (PTMs) at the amino acid level. Different proteoforms, even if corresponding to the

same protein family, may exhibit structural differences that are in turn associated with different functional roles in disease mechanism [7–10]. The clinical relevance of proteoform-level analysis is supported by studies indicating that different proteoforms of the same canonical sequence exhibit different disease-associated patterns, as detailed below.

As with aptamer- and antibody-based approaches, mass spectrometry (MS) supports high-throughput protein analysis, yet offers the distinct advantage of increased measurement specificity. Importantly, MS allows for distinguishing subtle differences between proteoforms based on multiple dimensions of measurement, namely the mass-to-charge ratio (m/z) of the proteoform ion and its associated fragment ions and retention time, if chromatography is employed [10,11]. To facilitate MS – based biomarker discovery, repositories resulting from the human plasma proteome project (<https://www.hupo.org/plasma-proteome-project>) and the human proteoform project (<https://proteomics.northwestern.edu/services/human-proteoform-project/>) have been made publicly available and updated with information regarding variation in protein structure [12,13].

MS-based approaches for protein characterization include bottom-up (BUP) and top-down (TDP) proteomics [11]. BUP relies on the chemical or enzymatic cleavage of proteoforms into smaller peptides prior to MS analysis. Due to the relative ease of peptide analysis in terms of separation, ionization and

Article highlights

- Proteoform characterization via TDP is expected to support more selective and specific biomarker discovery.
- Advances in fractionation and cleanup techniques have made it possible to routinely characterize 0–30 kDa proteoforms in clinical samples, though choice of technique may introduce bias in identified proteoforms, and approaches optimized for higher molecular weight proteoforms remain under development.
- A number of different sample preparation and cleanup approaches have been applied in recent clinical TDP studies for diagnostic and prognostic biomarker identification from samples such as blood, urine, and tissue biopsies.
- Large cohort application of clinical TDP is still limited by the lack of standardized sample preparation and analysis workflows.

gas-phase manipulation, BUP is more widely used but is inherently limited in its ability to unambiguously infer proteoforms from peptide fragments [14,15]. Alternatively, TDP analyzes intact proteoforms without digestion, utilizing both accurate intact mass analysis and tandem mass spectrometry (MS^n) to allow for accurate proteoform identification [16,17]. TDP remained difficult to implement until the introduction of soft ionization techniques, matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI) [18–20]. The higher protein charge states yielded by ESI are ideal for TDP, enabling intact protein analysis within the upper m/z limit of many high-resolution mass analyzers. However, as protein molecular weight (Mw) increases, signal dilution across multiple charge states and isotopologues becomes more pronounced, lowering signal-to-noise ratio (S/N) and limiting large-scale TDP analysis to proteins <30 kDa [21]. As such, current clinical applications of TDP have focused on the analysis of proteoforms in the 0–30 kDa range.

After providing an overview of top-down application to clinical proteomics, this review first summarizes strategies for sample preparation. Next, building on earlier guidelines [22], a discussion of clinical sample cleanup techniques for top-down workflows is presented, followed by a summary of practical considerations.

2. Top-down application in clinical proteomics

The case for applying TDP approaches to clinical studies, despite the analytical challenges, is based on the idea that knowledge of proteoforms, not simply protein identity, is relevant for understanding and monitoring disease progression. For example, for cardiovascular diseases (CVD), Sinari et al. reported the association of the plasma-enriched apolipoprotein CIII (apoC III, 8.8 kDa) proteoforms, truncated ($CIII_{0a}$), non-glycosylated ($CIII_{0b}$), O-glycosylated without sialic acid ($CIII_1$), and O-glycosylated with sialic acid ($CIII_2$) with disease demographic and clinical risk factors, highlighting the inverse association of $CIII_{0b}/CIII_1$ with CVD [9]. Wilkins et al. analyzed 15 apolipoprotein A-I (apoA-I, 28 kDa) and 9 apolipoprotein A-II (apoA-II, 17 kDa) proteoforms enriched from serum, reporting that unmodified apoA-I is positively associated with obesity, while the glycosylated form is positively associated with diabetes mellitus. On the other hand, the truncated and

dimerized proteoforms of apoA-II were inversely associated with obesity [23]. Furthermore, the analysis of plasma samples of the three stages of liver cirrhosis identified more than 2800 proteoforms in total, with 209 proteoforms having discriminatory power, including 11 apoA-I proteoforms differentially expressed between stage III and stage I [24]. In an analysis of urine samples, Moran et al. characterized different proteoforms of prostate-specific antigen (PSA), an FDA approved protein biomarker for prostate cancer, observing tri-, di-, mono-, and non-sialylated type N-glycans of the untruncated PSA in addition to six proteolytically cleaved variants [25]. Additionally, a TDP analysis of human saliva from patients with multiple sclerosis and healthy controls identified 23 proteoforms exhibiting differential abundance between groups, including oxidized and truncated forms of cystatin, a cysteine protease inhibitor that has been previously associated with neurodegeneration [26].

3. Effects of sample preparation approaches on proteoform identification

Challenges in clinical sample preparation include, especially in the case of biofluids, the high dynamic range of proteoform concentration. The signal dilution effect inherent to TDP, mentioned above, and proteoform complexity further complicate the detection of low abundant proteoforms from clinical samples. To address these challenges and improve MS detection of low-abundance proteoforms from biofluids, several immune-based and nonimmune-based depletion methods are commonly used in BUP to reduce levels of the top 2–14 most abundant proteins [27] and are now applied in TDP studies. Preparation strategies ranging from the targeted enrichment of individual proteins – using both immunoaffinity or non-antibody-based approaches – to untargeted methods, including gel-based fractionation, size exclusion chromatography (SEC), solid-phase extraction (SPE), molecular weight cut-off (MWCO) filter, and nanoparticles (NP) have further improved the characterization of lowly-abundant proteoforms [28,29].

Regardless of sample source and processing strategy, selection of the protein extraction and cleanup conditions is highly critical as such choices may impart artefactual proteoform modifications [22,30,31]. Cai et al. studied the effect of increasing sample storage temperature from 4 to 22°C on protein identifications from cardiac tissue. Interestingly, changes in the degree of phosphorylation were protein dependent, increasing for cardiac troponin I and decreasing for cardiac troponin T [32]. Kaulich et al. observed increased proteoform cleavage C-terminal to Asp for a workflow involving sample heating in unbuffered guanidinium chloride, potentially reflecting artefactual truncation at low pH and high temperature. Sample heating in loading buffer prior to gel-based fractionation may also introduce artefactual proteoform truncations [33]. On the other hand, use of acidic conditions during acetone precipitation may minimize introduction of a +98 Da modification [34].

In addition to generating artefactual proteoforms, extraction buffers can hinder MS analysis due to signal suppression

or adduct formation [30,35]. For example, signal suppression of 50% has been observed for sodium dodecyl sulfate (SDS) concentrations at low as 1.7 μ M [22]. Different sample cleanup procedures have therefore been developed to improve the efficiency of protein recovery based on the nature and level of buffer contaminants. However, it is important to note that technique performance in terms of applicable concentration ranges, proteoform recovery, and sensitivity influence the observed proteoform distribution. For example, many intact proteoforms are insoluble or poorly soluble, especially in the absence of detergent, thus the detergent removal step may introduce bias in observed proteoforms based on solubility differences [36–38]. For plasma and serum analysis, even selection of the type of abundant proteoform depletion kit, if such a step is included, can affect the observed proteoform distribution; comparison of five commercial kits indicated variable IgG removal, ranging from 70–93% [27].

Total proteoform recovery is affected also by the amount of sample available for LC-MS/MS analysis, with low injection amounts potentially resulting in a substantial decrease in proteoform identifications. Though not based on analysis of a clinical sample, the decrease in proteoform identifications from 738 ± 8 to 197 ± 10 observed by Kaulich and colleagues upon reducing proteoform injection amount from 1200 ng to 30 ng is also expected to be relevant for clinical studies [33]. Beyond details of the specific techniques, basic sample processing considerations also affect the observed proteoform distribution. Detection of Enigma Homolog isoform 2 from heart tissue was compromised after just 1 hour of sample storage at room temperature [32]. The next two sections provide further details on sample enrichment and cleanup approaches applied to TDP analysis (summarized in Figure 1, Table 1), with an emphasis on the implications for low molecular weight (0–30 kDa) proteoform identification from clinical samples.

4. Intact protein enrichment

The following section summarizes different strategies for sample preparation and introduces a number of techniques also applicable to sample clean-up, discussed in Section 5. Given the emphasis on clean-up approaches, additional details regarding each technique are included in Section 5.

4.1. Single/Targeted protein enrichment

4.1.1. Non-affinity-based targeted enrichment

Non-immunoaffinity targeted enrichment can be applied to study of highly abundant proteins such as plasma albumin, immunoglobulin, apoA-I, apoA-II, apoC I, apoC II, and apoC III. Starting from circulating proteins, plasma and/or serum are characterized by a complex proteome where 66 kDa albumin represents more than 50% of the total protein mass. MS analysis of human albumin can be carried out after sample dilution and disulfide bond reduction or after sequential precipitation with increasing concentration of ethanol [39,40]. With sequential ethanol precipitation and by applying intact mass analysis, Leblanc et al. reported the identification of nine different albumin proteoforms including oxidation,

glycosylation, deamidation, and N/C-terminal truncation [41]. A recent study demonstrated substantial improvement in the gas-phase sequencing of albumin, suggesting that in the near future advanced ion activation technologies will be available also for the characterization of endogenous serum albumin proteoforms [42].

Seckler et al. proposed a method for enriching 29 kDa apoA-I from human serum based on the extraction of the lipophilic part of serum followed by gel-based fractionation and collection of the 0–30 kDa fractions. This enabled the characterization of 18 apoA-I proteoforms including full-length, unmodified apoA-I along with carboxymethylated and acylated proteoforms [43]. ApoC III is another member of the apolipoprotein family, which was enriched using reversed-phase SPE by Jian et al. who characterized the ratio of three apoC III proteoforms (one canonical and two glycosylated forms) in individuals with diabetes mellitus [44]. Hemoglobin (Hb) proteoforms have also been interrogated via TDP due to their diagnostic potential. The main component of red blood cells (RBCs), Hb is enriched through a multi-step process encompassing the separation of RBCs from the whole blood, cell lysis followed by centrifugation for debris removal, and the collection of solubilized hemoglobin from the supernatant. After optimizing the chromatographic separation of Hb subunits, Chen et al. characterized seven clinically relevant proteoforms (Hb C with α -thalassemia trait, Hb E, Hb D-Punjab, Hb G-Accra, Hb G-Siriraj, Hb Tarrant, and Hb G-Waimanalo) via LC-MS [45].

Overall, targeted non-affinity techniques are simple and highly cost-effective, but they are limited in their application mostly to highly abundant proteins. The enrichment of proteoforms occurring in lower abundance is often confounded by the presence of residual higher-abundance proteoforms.

4.1.2. Affinity-based targeted enrichment

Affinity enrichment is the selective extraction of target proteins from complex mixtures based on the interaction of the protein of interest with a chemical moiety (e.g., aptamer) or antibody (immunoaffinity (IA) enrichment). The selectivity arises from combinations of non-covalent interactions between the target protein and the ligand such as hydrophobic interaction, hydrogen bonding, van der Waals or electrostatic forces [28]. The efficacy of protein recovery depends on three major steps: (i) binding, (ii) washing, and (iii) elution. An optimal ligand-to-target ratio is essential during the binding step to ensure the efficient capture of the target protein. The washing buffer must be carefully selected to effectively remove excess unbound proteins while preserving the integrity of the target-partner binding. The recovery of the target protein during the elution step relies on disrupting the non-covalent interaction with low or high pH or with denaturing agents [46].

Antibodies continue to be the workhorse of affinity enrichment due to their high specificity and selectivity. Antibodies used for immunoprecipitation can be coupled with different solid supports including pre-coated pipette tips, spin trap columns, and functionalized beads (either agarose or magnetic). Utilizing IA columns, Sinari et al. captured and analyzed the various proteoforms of the high-abundance protein apoC

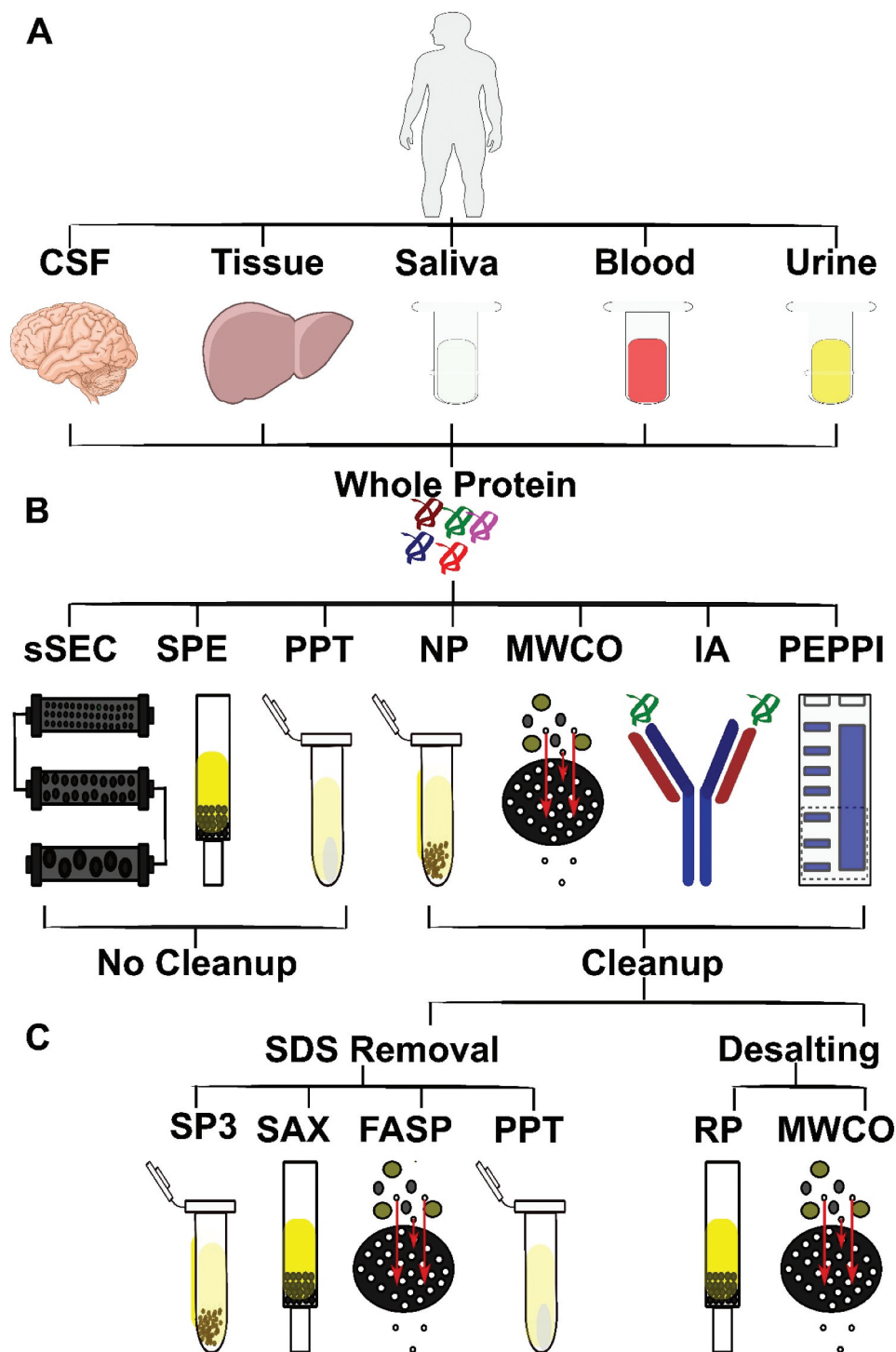


Figure 1. Overview of TDP sample preparation and cleanup protocols. (A) Different types of human samples used in clinical proteomics (blood, urine, saliva, cerebrospinal fluids (CSF), and tissue biopsy). The protein content from each clinical specimen is extracted for further manipulation. (B) Approaches used in TDP for enrichment/fractionation. (C) Classification of cleanup procedures according to the nature of contamination. Certain images used for creating the figure were taken from NIH bioart (<https://bioart.niaid.nih.gov/>). sSEC, serial size exclusion chromatography; SPE, solid-phase extraction; PPT, precipitation; NP, nanoparticle; MWCO, molecular weight cut-off; IA, immunoaffinity; PEPPI, passively eluting proteins from polyacrylamide gels as intact species; SP3, single-pot, solid phase-enhanced sample preparation; SAX, strong anion exchange; FASP, filter-aided sample preparation; RP, reversed phase.

III [9]. To evaluate the performance of different solid supports, Mazzarino et al. compared various options in terms of recovery and reproducibility for immunoenrichment of synthetic analogs of insulin-like growth factor 1 (IGF-I) from different biological matrices, urine and serum/plasma. Techniques relying on

the conjugation of antibodies to magnetic beads and pipette tips as support proved to be the most efficient [47]. With its great reproducibility and ability to capture proteins in the ng/mL range, a magnetic bead-based approach was used for the enrichment of lowly abundant protein biomarkers from serum,

Table 1. Top-down proteomics sample preparation protocols.

Preparation Approach	Principle	Practical Considerations	References
Non-affinity targeted enrichment	Abundance	Limited to highly abundant proteoforms	[39–41,45]
Lipoprotein enrichment	Lipophilic extraction and size-based fractionation	Unable to differentiate between bound and free apolipoproteins	[43,44]
Affinity (IA)	Selective binding to antibody or aptamer	Multiprotein enrichment is not feasible; may be integrated into online, automated workflow	[9,47–51]
1D-PAGE	Size based	Need for protein extraction from gel matrix may limit recovery	[54,56,81]
2D-PAGE	Isoelectric point and size based	Need for protein extraction from gel matrix may limit recovery	[57–59]
GELFrEE	Size based	Unit needs to be built in-house	[60–63]
MWCO filter	Size based	Membrane efficiency is vendor dependent	[65,66,82]
SPE enrichment	Reversed-phase; hydrophilic-lipophilic balanced; affinity (depending on sorbent)	Nonspecific; may yield lower protein recovery	[27,67,69]
SEC	Size based	Limited resolving power	[73,74]
Nanoparticle	Selective binding to the functionalized nanoparticle	Nanoparticle production may require specialized synthesis	[78–80,83]

such as neuron-specific enolase gamma (NSE), the acetylated proteoform being associated with lung cancer [48]. An additional advantage of magnetic bead-based sample enrichment is its suitability for automation using liquid handling robots. Des Soye et al. implemented an entirely automated workflow for isolating immunoglobulins G from serum samples obtained from SARS-CoV-2 patients by coupling IA enrichment with the KingFisher robot and sample cleanup with SampleStream technology followed by MS analysis [49].

Coupling affinity enrichment with SPE in fritted or fritless micro cartridges has enabled the online integration of sample preparation and cleanup with capillary electrophoresis (CE)-MS. The Benavente lab developed two platforms, antibody- and aptamer-based SPE-CE-MS for the targeted analysis of transthyretin and α -synuclein proteoforms extracted from plasma and serum, respectively. Automating sample preparation and cleanup steps reduced manual workload and enhanced the reproducibility and sensitivity of downstream analysis [50,51].

In summary, antibody-based immunoaffinity is the most commonly applied strategy for affinity-based sample preparation for TDP clinical samples, despite its shortcomings [16,52]. A major limitation of this approach is its inability to support multiplexed protein analysis, as a specific antibody is required for each protein target. Although a wide range of antibodies is commercially available, not all are suitable for affinity enrichment and often require extensive validation, especially to confirm specificity. Additionally, the cost of antibodies may

limit their adoption for large-scale TDP studies. Finally, it is unclear whether antibody-based enrichment introduces proteoform-specific bias.

4.2. Untargeted approaches

4.2.1. Fractionation based on gel electrophoresis

Gel-based fractionation utilizes the cost-effective polyacrylamide gel electrophoresis (PAGE) for the separation of proteoforms based on size and charge (Native-PAGE), size (SDS-PAGE), or isoelectric point (isoelectric focusing, IEF). PAGE is one of the most common offline fractionation approaches for TDP due to its orthogonality to separation techniques like reversed-phase liquid chromatography (RPLC) and CE [29]. Despite its efficiency in providing high-resolution protein separation, the bottleneck of gel-based fractionation is intact protein recovery. Several strategies such as continuous elution electrophoresis, electroelution, and passive extraction with aqueous solvents have been tested, with minimal protein loss observed from passive elution. Passive elution was first implemented to study tomato extract and was later optimized for TDP analysis of biological samples by Takemori and coworkers as a protocol known as passively eluting proteins from polyacrylamide gels as intact species for mass spectrometry (PEPPI-MS, Figure 2) [53–55]. Recovery after PEPPI extraction of up to 68% was observed. However, studies have demonstrated an extraction bias favoring low Mw proteoforms and acidic and neutral proteins, suggesting avenues for further optimization.

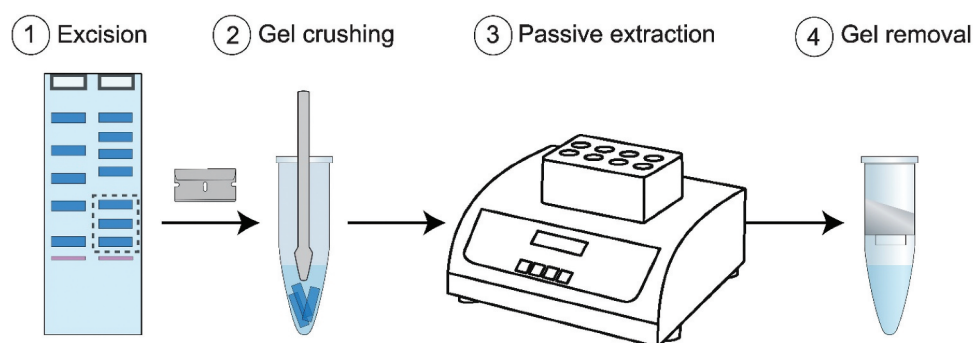


Figure 2. Workflow for sample fractionation using PEPPI: (1) Bands of interest are excised from the SDS-PAGE gel, (2) Gel pieces are crushed with a pestle in ammonium bicarbonate (ABC) buffer containing 0.5% SDS (pH 8–8.5), (3) Agitate tubes to facilitate extraction from gel. (4) Gel pieces are separated from the recovered protein solution using a spin filter. After this step, the protein solution is ready for cleanup. Certain images used for creating the figure were taken from opencipart (<https://opencipart.org/>) and Labicons (<https://www.labicons.net/>).

One-dimensional electrophoresis (1D-PAGE) and two-dimensional electrophoresis (2D-PAGE) are the two widely available gel-based protein/proteoform separation techniques applied in TDP. In 1D-PAGE, proteins are separated under denaturing conditions using a 10 or 12% SDS-PAGE gel. SDS disrupts the native protein structure by inducing denaturation and uniformly imparts negative charges to the proteins, allowing their migration based on size. If the preservation of native structure is desired, native PAGE can be performed. Forte et al. used PEPPi-MS to isolate the 0–30 kDa proteoform fraction for the liver cirrhosis study referenced above [24]. Takemori et al. applied PEPPi under native-like conditions to achieve the purification of tetrameric hemoglobin complex, highlighting the promise of 1D-Native-PAGE in purifying protein ensembles for subsequent native mass spectrometry analysis [56].

2D-PAGE combines two orthogonal separation methods. In the first dimension, a gel strip containing a pH gradient is used for separation based on isoelectric point – proteoforms migrate until reaching the point in the pH gradient where their net charge is zero. Proteoforms showing differences as minimal as an extra PTM, such as phosphorylation, can be resolved. In the second dimension, standard SDS-PAGE is applied, allowing size-based separation [57,58]. Almuslehi et al. revealed after analysis of tissue biopsy extract the differential abundance of 34 protein spots in a 2D-PAGE gel that can be associated with visual disturbances preceding multiple sclerosis. Subsequent TDP analysis identified 75 characteristic proteoforms among which 24 were unique [59].

Tran and Doucette introduced the gel-eluted liquid fraction entrapment electrophoresis (GELFrEE) separation device [60,61]. Later, Skinner et al. described a native-like version of this method for the separation of protein complexes [62]. The device consists of a tube gel (divided into stacking & resolving gels) and a collection chamber. The major advantages of GELFrEE over other gel-based fractionation methods are negligible sample loss with short trapping time and the elimination of the passive elution step, though sample clean-up is still mandatory when SDS is used. To enable the simultaneous fractionation of multiple samples, Tran and Doucette developed a multiplexed GELFrEE unit [63] that has been widely used for TDP in 2D (coupled with LC or CE separation) or 3D formats (coupled with sIEF and LC or CE) [7,61,64]. Unfortunately, the commercial GELFrEE device, which allowed for multiplexed sample fractionation is no longer available, thus the implementation of GELFrEE now requires in-house built devices.

4.2.2. Other fractionation strategies applied to whole proteomes

Gel-based fractionation is the most widely employed approach for reducing sample complexity prior to TDP analysis; however, alternative methods have also demonstrated significant potential in this context. These include MWCO filters, SPE, SEC, and the use of NP. While most of these methods (namely: MWCO, SPE, SEC) have their primary utility in sample cleanup – and hence more details on their mechanism will be provided in [section 5 Intact protein cleanup](#)—, notable examples are provided below for their application in the context of extracting proteoforms from biological samples and proteome fractionation.

MWCO filters contain polymeric membranes and utilize centrifugal force to separate molecules based on their size. Different MWCO filters are available with molecular weight cutoffs of 3, 5, 10, 50, 100, 300, and 1000 kDa. A desired Mw range can be enriched by the sequential application of filters units with different cutoff values. Fulcher et al. utilized a series of MWCO filtration procedures to enrich proteins from 0–100 kDa from postmortem brain tissue of individuals with Alzheimer's disease. After tissue homogenization and the removal of insoluble material, 100 kDa filters were used to remove large proteins, followed by 3 kDa filters to remove small molecular weight contaminants. TDP data suggested metal (mainly iron) binding to β -synuclein and γ -synuclein as well as succinylation of cysteine residue 106 of PARK7 in a patient with Alzheimer's [65]. Though use of MWCO filtration alone is not adequately selective for targeted work, combining filtration with targeted methods is a strategy to improve sample purity and thus sensitivity for the targeted proteoforms. Pont et al. combined immunoaffinity capture and two stages of MWCO filtering for the intact mass analysis of transthyretin (TTR) proteoforms enriched from the plasma of a patient with familial amyloidotic polyneuropathy type I (FAP-I). Following immunoprecipitation, the sample was first desalted with a 10 kDa filter, followed by antibody removal with a 100 kDa filter. The authors reported that the oxidized forms of TTR, derived from either the wildtype or mutant alleles, are associated with FAP-I incidence [66].

SPE allows for extraction of subproteomes of interest based on selection of stationary phase, encompassing options ranging from reversed-phase and ion exchange sorbents to affinity sorbents for depletion of abundant plasma proteins [27,67]. Initial attempts to utilize SPE for intact protein enrichment focused on single protein extraction from human serum (i.e., targeted approach). Following denaturation and the precipitation of high molecular weight proteins, IGF-I (~7.6 kDa) was enriched using an anion exchange resin [68]. Expanding on this study, Marakova et al. further optimized the protocol for the enrichment of low Mw proteins spiked into human plasma and urine. The authors expected that larger, hydrophobic proteins would precipitate during the elution step using high acetonitrile content (75% (v/v)), therefore they targeted hydrophilic proteins of Mw < 30 kDa. Multiple factors were tested to maximize protein recovery, including sample pretreatment with acid/detergent (to disrupt protein-protein interactions), selection of both sorbent and solution chemistries for the proper binding, washing and subsequent elution of the proteins. The highest protein recoveries (> 65% and > 50% from urine and plasma, respectively) were reported using the following conditions: sample pretreatment with 1% (v/v) trifluoroacetic acid (plasma) and 0.2% triton (urine); hydrophilic-lipophilic balanced resin; elution with 1% (v/v) trifluoroacetic acid in 75% (v/v) acetonitrile [69]. Affinity-based stationary phases have shown utility in the enrichment of human glycoproteins from both urine and plasma samples, utilizing either single- or multi-lectin columns [70,71]. Although these works reported on subsequent BUP processing, they highlight the potential of lectin-based strategies for the purification/enrichment of intact glycoproteins. Another significant achievement in this field was demonstrated by Wu et al. Here, two standards – haptoglobin and α 1-acid glycoprotein – were used to highlight the power of combining lectin-affinity

purification and native MS to elucidate the microheterogeneities of their glycoproteoforms [72]. These findings support the exploration of agarose bead – based lectin enrichment protocols in TDP for the extraction of the glycoproteome from complex biological samples. The studies mentioned above are very promising and will hopefully spur further improvements that allow SPE fractionation of a wider variety of subproteomes.

SEC is a separation method based on the entrapment of species within a porous matrix. Pore sizes range from 100 to 4000 Å, and for the purposes of fractionation one must choose the resin that accommodates the fractionation range of the protein(s) of interest. The separation is based mostly on hydrodynamic volume, where high Mw components elute first, and the elution (generally) follows the order of decreasing Mw. Although SEC has limited resolving power, it proves very efficient for sample cleanup [28] and the enrichment of extracellular vesicles (EVs). SEC columns specifically designed for isolating EVs are commercially available and have been utilized by Lattmann et al. to compare the protein cargo of EVs enriched from the plasma of healthy vs melanoma patients. Although this study applied BUP, such SEC columns are expected to be applied in TDP workflows as well, due to the great potential EVs hold for biomarker discovery [73]. In an effort to improve the resolution of SEC, the Ge lab developed a serial size exclusion chromatography (sSEC) strategy [74]. In sSEC, a series of SEC columns are aligned in order of decreasing pore size (1000 Å followed by 500 Å) and proteins are eluted isocratically. While the use of a single SEC column allowed for the detection of more than 2000 proteoforms in cardiac tissue extract (with the majority of these being in the 10–25 kDa range); with sSEC, the number of intact proteoform masses identified exceeded 5000, with a 15-fold increase in the number of individual proteoforms in the 25–50 kDa range. The solvent compatibility of the sSEC with MS eliminates the need for sample cleanup. Additionally, the lack of proteoform bias highlights sSEC as a promising method for further optimization for the enrichment of proteoforms from other biological matrixes.

NP are fine particles with diameters ranging from 1 to 500 nm. They can be synthesized from a variety of materials including lipids, carbon-based compounds and polymers and were originally utilized in nanomedicine for delivering drugs to the desired site of action. Upon exposure to biological fluids, a protein corona (a dynamic coating of biomolecules, usually proteins) forms on the surface of NPs, potentially altering their intended behavior. The composition of the protein corona is dictated by the physicochemical properties of NPs, a phenomenon which was leveraged by Seer Inc. to develop engineered magnetic NPs that allow the capture and interrogation of different subsets of the proteome. The main motivation was to compress the large dynamic range of proteomes by utilizing a panel of multiple NPs that became the foundation of the Proteograph Platform [75–77]. While these studies included a protein digestion step, the concept has subsequently been applied to TDP.

The first application of NPs in TDP was presented by the Sun lab, using polystyrene NPs prepared in-house. The NPs were incubated with human plasma and the protein corona formed was eluted off with SDS. To optimize recovery, SDS

concentration was varied in the range of 0.4–2%, whereby 1% SDS was deemed optimal. The authors also pointed out that using 1–2% SDS improved the recovery of large proteoforms, specifically; even so, transmission electron microscopy images suggested incomplete elution. The protocol enabled the characterization of nearly 900 proteoforms (by combining CE-MS/MS and LC-MS/MS results) in the Mw range 2–70 kDa. The major drawback of using non-engineered NPs is that the protein corona is mainly comprised of highly abundant plasma proteins, such as apolipoproteins [78]. In 2024, the Kelleher lab demonstrated the application of Seer-engineered nanoparticles in a TDP study of human plasma. They modified the original Proteograph workflow and eluted the protein corona with SDS. The application of Seer NPs enabled the identification of a record number (2841) of unique plasma proteoforms from 114 proteins with Mw < 50 kDa. Furthermore, their approach allowed the detection of proteoforms in concentrations as low as 10 pg/mL, expanding the proteomic depth 10⁵ fold, compared to unprocessed plasma. With that said, further testing of engineered NPs is needed to evaluate the extent of bias arising from the interaction of differently modified proteoforms with NPs [79]. Chemically modified magnetic NPs have been developed for the enrichment of the phosphoproteome. Chen et al. synthesized cobalt ferrite NPs, which were functionalized with dinuclear Zn (II)-dipiclylamine (Zn-DPA) ligands coupled to glutaric acid, providing the specific binding of phosphate ions. These NPs showed great utility in enriching multiple cardiac phosphoproteins from (swine) heart tissue lysates, allowing the localization of potential phosphorylation sites for a previously uncharacterized phosphoprotein, hepatoma-derived growth factor [80]. Currently, the lack of commercially available phosphoproteome enrichment kits/beads amenable for TDP limits the expansion of the field into the clinical setting. Finally, it is important to underline that NP enrichment does not substitute Mw-based fractionation: in the above reported studies, low Mw proteoforms were isolated by PEPPI-MS or similar methods.

5. Intact protein cleanup

Salts and detergents are the major contaminants influencing MS analysis of proteins. Salts are naturally present in the tissue and body fluids but may also be introduced during sample preparation, for example, when implementing ion-exchange. Further, detergents may also be added during sample preparation to facilitate solubilization and/or elution of membrane, hydrophobic, and extracellular matrix proteins. Detergents can suppress protein ionization in ESI-MS, in addition to forming protein adducts that lead to further signal dilution and decreased sensitivity [22]. Multiple cleanup approaches for salt and detergent removal have been developed and made commercially available, including dialysis, ultrafiltration, SPE, SEC, and precipitation.

5.1. Dialysis

Dialysis separates low Mw compounds such as salts from macromolecules like proteins via selective diffusion across semi-permeable membranes, typically composed of

derivatized cellulose, with selectivity based on choice of MWCO, and may also be used for buffer exchange. Commercially available kits include devices for sample volumes ranging from 10 μ L to 250 mL. Offline dialysis before TDP mass spectrometry has a limited application due to the time required, and the need for a large volume of buffers (at least 200 times sample volume). In practice, dialysis may simply result in sample dilution and even sample loss due to binding to the dialysis membrane, rather than effective salt removal [28]. Further, dialysis alone is unsuitable for SDS cleanup, as only unbound SDS is removed [31].

These limitations have inspired a focus on using the dialysis membrane as an integral part of online sample cleanup. Kim et al. developed a matrix removal device based on asymmetrical flow field-flow fractionation (AF4) for online cleanup of GELFrEE enriched proteins, enabling direct MS analysis of samples containing up to 0.1% SDS and salts. In an initial step, cross-flow perpendicular to a 10 kDa MWCO membrane allows for protein retention and contaminant removal, prior to redirecting flow to elute sample directly to an emitter for MS analysis [81]. Additionally, Seckler et al., employed SampleStream for desalting and buffer exchange of immunoenriched apoA-I and apoC III for characterization of proteoforms associated with CVD [82]. SampleStream functions with the same principle as AF4, with membrane MWCO controlling the desalting procedure [83]. SampleStream was tested by direct coupling with MS with no orthogonal sample separation applied before MS analysis. Though offering the advantage of sample cleanup online with MS analysis, SampleStream provides limited capacity for protein/proteoform separation, thereby limiting its application to targeted proteomics. Further improvements to on-line coupled separation technologies such as LC and CE would enhance its utilization for online sample cleanup.

5.2. Solid phase extraction (SPE)

SPE relies on physicochemical interaction between the analyte of interest and the resin, which can be RP, normal phase, ion-exchange or affinity based. The process consists of three main steps: (i) binding the analyte of interest to solid support, (ii) elimination of contaminants, and (iii) elution of target analytes. The main criteria to be considered when selecting SPE beads include the nature of the analyte to be cleaned, quantity of analyte in the sample, nature of contaminants, and the level of contamination [67].

For TDP, C4 resin (RP) is most commonly used – the short alkyl groups efficiently bind proteins and large peptides, while polar compounds like salts remain unretained. Due to its hydrophobic nature, the method is suitable only for desalting and sample concentration, not detergent removal. Palmblad et al. reported desalting efficiency reaching 98% and protein recovery generally of 70–80% for C4 tips [84]. The major drawback of the RP micropipette tip cleanup is the loss of hydrophilic protein targets, inconsistencies in desalting efficiency due to the variability in the packing material, and the laborious and time-consuming pipetting procedure, which can introduce errors due to manual handling.

The removal of detergents can be carried out using ion exchange resins. Ion exchange is based on the interaction of solutes with a stationary phase of opposite charge. Four classes of ion exchange resins are available, namely strong anion exchange (SAX), weak anion exchange (WAX), strong cation exchange (SCX), and weak cation exchange (WCX). SAX StageTips was developed by the Takemori lab for on-tip digestion of protein samples. The Takemori group implemented the same StageTip technique as a complementary cleaning procedure for TDP sample enrichment via the PEPPI-AnExSP protocol. The StageTip enables binding of both the negatively charged proteins and SDS. Proteins are eluted using 50% ethanol/0.5% FA for cleanup of samples containing 0.05% SDS [55]. The main disadvantage of using SAX on sample cleanup is the loss of cationic proteins since they do not bind to the cationic resin. A combination of cation and anion exchange (mixed-bed) has been implemented to improve protein recovery in BUP, though this approach has not been tested for TDP [85].

5.3. Molecular weight cut-off (MWCO) filters

In addition to proteoform concentration and relatively imprecise separation by Mw, filtration may be used for buffer exchange and desalting. Filters with MWCO approximately half the molecular weight of proteoforms of interest provide a balance between contaminant removal and sample recovery, since higher MWCO filters may lead to sample loss into the filtrate. A typical workflow starts with conditioning the membrane with appropriate solutions; extremes of pH may damage the membrane filter. After conditioning, the sample is added, and contaminants are filtered after a series of centrifugation steps using MS-compatible solvents. The rate of centrifugation depends on the level of contamination. The retentate containing the sample is then collected for further analysis [28].

For SDS removal, Takemori et al. adapted the filter-aided sample preparation (FASP) protocol, originally introduced by Nagaraj et al. for BUP, as part of the PEPPI-AnExSP workflow mentioned above [86]. In the TDP adaptation, proteoforms extracted from PEPPI fractions are loaded onto a 3 kDa MCWO filter, and a series of centrifugation steps with 8 M urea are performed to remove both unbound and bound SDS, followed by buffer exchange into ammonium bicarbonate (ABC) buffer prior to the final anion exchange cleanup step [55]. Kachuk et al. developed a transmembrane electrophoresis (TME) device incorporating MWCO filters to retain proteins while applying an electric field to remove bound and unbound SDS. Tests with bovine serum albumin indicated that protein recovery of greater than 90% may be achievable while decreasing SDS concentrations from as high as 2% to less than 10 ppm. Membrane proteoform solutions containing 40% methanol may also be cleaned via TME. In this case, more rapid SDS removal was observed, possibly due to an increased SDS critical micelle concentration in methanol solutions. However, TME is not commercially available [87,88].

5.4. Precipitation

Protein precipitation is a cost-effective, readily available sample cleanup procedure and employs solvents that interfere

with proteoform non-covalent interactions or protein hydration resulting in aggregation and precipitation. The degree of precipitation is dependent on pH, ionic strength, type and percentage of the precipitating agent, temperature, and protein concentration [28,89].

Organic solvents are widely used for sample cleanup procedures to precipitate large molecular proteins while keeping low Mw contaminants in solution. While ethanol precipitation may be useful for removal of certain lower Mw contaminants, the limited solubility of salts in ethanol necessitates an additional desalting step [90]. For detergent removal, a widely used approach in TDP is methanol-chloroform-water (MCW) precipitation [35–37,91] via a protocol that has not substantially changed from the original one developed by Wessel and Flügge [92]. Proteins are initially precipitated at the aqueous-organic phase interface using a solution of 4 parts methanol, 3 parts water, and 1 part of chloroform per 1 part sample, followed by a series of methanol-washing steps. Although the process is straightforward, protein recovery is variable, with sample loss potentially resulting from the extensive pipetting procedure, especially for dilute samples. In addition, MCW is not ideal for the recovery of low Mw proteins [38]. Cold acetone at a ratio of 4 parts acetone: 1 part sample may be used for overnight precipitation at low temperature (–20 °C). The cold temperature minimizes the conformational elasticity of the protein, facilitating preservation of native structure assuming no denaturing agent is added to the sample. However, recovery is variable depending on the nature of the protein [93]. TCA on the other hand, induces protein aggregation and precipitation by lowering solution pH, which affects hydrogen bonding interactions and alters protein surface charge. Development of a precipitation protocol for platelet samples suggested that intermediate TCA concentrations of 15 to 40% were necessary for protein precipitation [90].

5.5. Magnetic beads

Automated sample preparation and cleanup offer the advantages of less labor-intensive procedures and high reproducibility. Magnetic bead-based sample cleanup may be more readily automated compared to other approaches while also allowing for improved sample recovery compared to SampleStream and TME. Beads are based on iron oxide or cobalt ferrite nanoparticle cores and do not aggregate except in the presence of a magnetic field. Modification of particle surface chemistry allows for relatively straightforward adaptation of the same base protocol to a range of samples. Single-pot, solid phase-enhanced sample preparation (SP3) is the first magnetic bead-based sample cleanup procedure originally developed for salt and detergent removal in BUP [94,95]. A modified protocol was introduced by Dagley et al. for both TDP and BUP analysis utilizing carboxylated magnetic beads, referred to as universal solid-phase protein preparation (USP3). With USP3, the protein sample is acidified prior to aggregation and captured by the magnetic beads. Both acetonitrile and ethanol are used for this step, and intact proteins are eluted from the beads using cold 80% formic acid. However, samples must be analyzed immediately and may not be stored [96]. To avoid the harsh formic acid elution step, Takemori et al. modified the protocol for the cleanup of PEPPI fractions in a protocol referred to as PEPPI-SP3 (Figure 3), with the elution of the intact protein from the carboxylated beads performed with 100 mM ABC containing 0.05–0.1% SDS. SDS is then removed via AnExSP. SP3-AnExSP clean-up of PEPPI fractions yielded a higher number of total proteoform identifications compared to MCW or FASP-AnExSP cleanup [38]. Although SP3-AnExSP improved < 30 kDa proteoform recovery compared to FASP, the efficient elution of larger proteoforms from beads may become challenging due to increased hydrophobicity and will require careful evaluation and optimization.

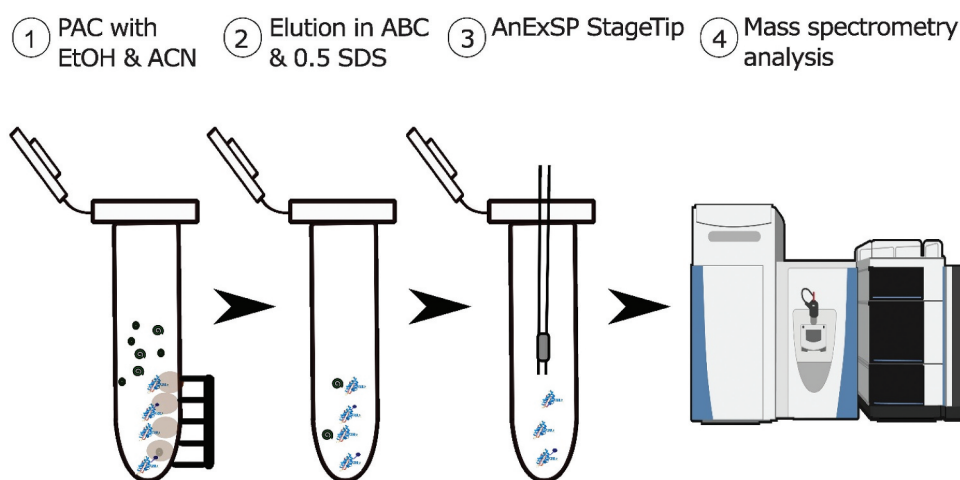


Figure 3. Workflow for SP3-AnExSP sample clean-up for TDP-MS analysis. (1) To remove contaminants from PEPPI fractions, the sample is subjected to protein aggregation capture (PAC) using carboxylated magnetic beads, with sequential mixing in ethanol (EtOH) and acetonitrile (ACN). (2) Proteoforms are then eluted from the beads using ABC buffer containing 0.5% SDS. (3) The SDS introduced is then removed using an in-house – built stage-tip containing an anion-exchange disk, prior to (4) LC-MS/MS analysis. Certain images used for creating the figure were taken from NIH bioart (<https://bioart.niaid.nih.gov/>).

6. Comparative evaluation of sample preparation and cleanup procedures

Though not specifically focused on clinical samples, studies comparing performance characteristics of different cleanup methods for TDP workflows provide general insights for clinical proteomics. In terms of protein recovery, Yang et al. compared 30 kDa MWCO filtration, MCW, and USP3 protocols for cleanup of cell lysate samples containing 1% SDS. High protein recovery (~60%) was observed for MWCO filtration and USP3. However, MCW exhibited a bias for high Mw proteoforms (> 60 kDa), whereas USP3 exhibited a bias for low Mw proteoforms (< 30 kDa). MWCO filtration did not show a proteoform bias based on Mw. This is consistent with the observations of Takemori et al., who performed a comparative TDP analysis of the three cleanup protocols after gel-based fractionation [36,38]. In this optimization of SP3, use of 500 µg beads allowed for recovery of 1.88 µg and 0.94 µg protein for 48–20 kDa and < 20 kDa PEPPI fractions of 10 µg human cell protein extract, respectively. Total ion current observed for top-down LC-MS/MS analyses was ~50% lower for MCW samples compared to SP3 samples, even though injection amount was ~75% higher for MCW samples, providing semi-quantitative confirmation of relatively higher proteoform recovery for SP3.

Further, for studies with myoglobin and cytochrome C, Puchades and colleagues reported ~50% recovery with MCW, with higher recovery (~80%) observed for precipitation with 80% acetone [91]. However, proteoform recovery for solvent precipitation methods is variable, with Wessel and Flügge observing recovery after MCW cleanup of $92 \pm 2.5\%$ for 40 µg BSA in 5% SDS solution and $94 \pm 1.5\%$ for 50 µg rabbit serum proteins in 5% SDS [92]. A series of thorough studies from Doucette and colleagues identified ionic strength as an important variable for proteoform recovery during precipitation. Specifically, for acetone precipitation, they observed that 1–100 mM NaCl (or other ionic species) was needed for high recovery of soluble proteoforms, with presence of salt allowing $98 \pm 1\%$ recovery at room temperature [93,97]. However, while recovery remains important, removal of salts and detergents is paramount during sample cleanup. Botelho et al. observed a reduction in SDS concentration from 2% to less than 0.006% after MCW precipitation, whereas residual SDS as high as 0.05% was observed after acetone precipitation, though this was successfully mitigated by inclusion of additional wash steps [35]. Despite the lower proteoform recovery observed for MCW compared to acetone precipitation, Puchades et al. observed a reduction in SDS from 0.5% to 0.0005% after MCW precipitation, whereas residual SDS was 4 times higher for 80% acetone precipitation [91].

Along with global measures of total proteoform recovery and contaminant reduction, number of unique proteoform identifications and lack of bias in proteoform recovery also represent important benchmarks. For the top-down investigation of clinical samples in particular, a previous study compared 3 protocols for serum sample cleanup: MCW, AnExSP, and a commercially available kit, HiPPR Detergent Removal Column. Both HiPPR and AnExSP yielded greater enrichment of low Mw proteoforms compared to MCW, with the highest

number of unique proteoform identifications observed for AnExSP [37]. A comprehensive summary of the quantitative recovery of the different sample preparation and cleanup procedures is shown in Table 2.

Practical considerations in terms of time and cost are also relevant, especially when optimizing a workflow for clinical applications. AnExSP and FASP are both labor-intensive and time-consuming. While commercial kits support rapid detergent cleanup, they are more costly solutions compared to MCW. Protein precipitation with either MCW or acetone are inexpensive, but, as noted, exhibit more limited recovery of low molecular weight proteoforms and may result in sample loss during pipetting. In contrast, SPE favors recovery of low molecular weight proteoforms, yet may result in loss of hydrophilic and cationic proteoforms for hydrophobic SPE and anion exchange SPE, respectively. MWCO filtration may provide less substantial bias in terms of proteoform MW, but is time-intensive and may yield reduced recovery of proteoforms with certain modifications [38]. Online sample cleanup with either SampleStream or TME are rapid but relatively expensive. In clinical applications where high throughput is desired, automated sample cleanup with magnetic beads is a promising option.

When optimizing a TDP clinical proteomics workflow, the effectiveness of the sample cleanup step may be monitored by adapting the ContamSPOT method, originally developed for BUP. The assay relies on ion pairing between o-toluidine blue and SDS to detect detergent concentrations as low as 0.0004% [98]. Residual SDS may also be quantified colorimetrically using Stains All, a carbocyanine dye, as described by Rusconi et al. [99].

7. Conclusion

Despite the versatile approaches for sample preparation and clean-up available for TDP, each protocol has drawbacks, including potential introduction of bias in proteoform identifications and artefactual modifications. As such, method selection is not trivial. For small sample volumes typical of clinical proteomics studies, gel-based fractionation is a reasonable choice for sample preparation but necessitates sample clean-up for detergent removal. A key advantage of the sSEC protocol is its use of MS-compatible elution solvent, yet the requirement for large sample volumes limits its feasibility for routine clinical assays. Work toward improving and standardizing TDP clean-up techniques is nonetheless valuable, as the shift from protein to proteoform analysis in clinical proteomics is expected to allow for identification of more selective disease-biomarkers and targets for personalized medicine.

8. Expert opinion

The main limitations in existing TDP clean-up techniques relevant for clinical proteomics include protocols that are difficult to scale up for high-throughput studies and that are difficult to reproduce by non-experts. To address this, promising approaches include adaption of protocols to formats compatible with automated sample handling systems and protocol commercialization. Continued improvements that minimize

Table 2. Reported quantitative measures of technique performance in top-down clinical proteomics studies¹.

Technique	Material/Format	Sample Input	Performance	Reference
Targeted affinity enrichment	Magnetic Protein G-labeled bead immunoprecipitation	1 mL	Concentration: quantification of 6.25–100 ng/mL neuron-specific enolase γ in serum Sensitivity: LOD 3.5 or 4.2 ng/mL (depending on peak integration method); LOQ: 10.6 or 12.8 ng/mL (depending on peak integration method)	[48]
	SARS-CoV-2 receptor binding domain conjugated beads	100 μ L plasma	Concentration: ion titer (signal intensity attributed to immunoprecipitated antibody light chains and heavy chains compared to intensity of standard) greater than 95% upper confidence level of signal from nonspecific binding Relative quantification: 19 of 27 patients testing positive for COVID-19 also identified as positive via automated protocol for individual ion MS, based on ion titer Relative quantification: low apo CIII _{0B} /III ₁ associated with increased cardiovascular disease risk	[49]
Untargeted nanoparticle enrichment	Anti-apo CIII antibody immunoaffinity columns	3 μ L plasma		[9]
	Aptamer affinity/solid phase extraction	Thermo-enriched red blood cell lysates	Concentration: linear range, 0.5–10 μ g/mL (CE-MS) Sensitivity: LOD 0.2 μ g/mL (CE-MS), alpha-synuclein standard; detected at 0.5 μ g/mL spiked into thermo-enriched red blood cell lysate	[51]
	Immunoaffinity	5 mL urine 100 μ L plasma or serum	Sensitivity: LOD 0.05–0.15 ng/mL, urine; 0.5–1.5 ng/mL, blood Proteome recovery: 50–71% (magnetic beads), 49–77% (pipette tips), IGF1 analogues in urine; 51–73% (magnetic beads), 51–75% (pipette tips), in serum/plasma	[47]
	anti-ApoB immunoprecipitation resin	10 μ L serum	Relative quantification: quantification of proteoforms with relative abundance < 1%	[23]
Electrophoresis	Immunoprecipitation/centrifugal filtration	50 μ L serum	Concentration: detection of estimated 50–130 μ g/mL transthyretin in serum Sensitivity: LOD < 50 μ g/mL (each monomer form, transthyretin standard solution)	[66]
	Seer nanoparticles/PEPPI	100 μ L plasma	Sensitivity: detection of ~1 μ g/mL to 10 μ g/mL proteoforms in plasma	[79]
	Nanoparticle enrichment coupled to online capillary electrophoresis	550 μ L plasma (note: nanoparticle-protein complexes split, with 2/3 portion used for top-down analysis)	Concentration: Top-down approach, slope of BSA calibration curve (0.5–2 μ g/ μ L) 3 orders of magnitude lower than that for bottom-up approach (< 0.1 μ g/ μ L–1 μ g/ μ L)	[78]
	PEPPI	500 μ g plasma protein	Proteome recovery: 144 μ g from 200 μ g input, after buffer exchange for SDS removal	[24]
Solvent precipitation	Abundant protein depletion/PEPPI	60 μ L serum (depletion); 300 μ g protein (PEPPI)	Proteome recovery: based on report for non-clinical samples median 68%, proteoforms < 100 kDa; median 57%, proteoforms > 100 kDa (Takemori 2020)	[37]
	GELF-EE	200–500 μ g protein	Relative quantification: differential expression detected for 198 proteoforms in 0–30 kDa fraction, PBMC proteoforms, in liver transplant rejection study based on post-transplant status	[7]
	GELF-EE	5–8 mL whole blood collected; 100 μ g PBMC protein (GELF-EE)	Concentration: based on report for non-clinical samples, maximum input up to 1 mg for proteome mixture (Tran 2008) Sensitivity: based on report for non-clinical samples, LOD ~ 8 ng, single protein (for silver stain detection after GELF-EE) (Tran 2008)	[64]
Solid phase extraction (SPE)	precipitation (TCA or ethanol)	1500 μ L platelet suspension (TCA precipitation); 900 μ L platelet suspension (ethanol precipitation)	Proteome recovery: 124 $\pm$ 10 μ g per 100 $\times$ 10 ⁶ platelets without precipitation, versus ~120 μ g for either precipitation method; based on 2D-electrophoresis analysis, ~75% spots similar fluorescence intensity for ethanol precipitation (versus no precipitation), and ~79% spots similar for TCA precipitation (versus no precipitation)	[93]
	Precipitation (1% TFA or 0.2% Triton)/microSPE	100 μ L urine, plasma, or serum	Sensitivity: LOQ 2–1200 ng/mL, depending on protein standard and matrix Proteome recovery: 65–91%, for standards spiked into urine; 0.6–77% and 0.6–73% for standards spiked into plasma and serum, respectively	[69]
Filtration	Precipitation (5% acetic acid in acetonitrile, 0.6% SDS, excess IGF-II)/SPE	100 μ L serum	Concentration: 5–1000 ng/mL Sensitivity: LOQ 5 ng/mL	[68]
	10% formic acid addition/SPE (96 well plate)	100 μ L plasma	Proteome recovery: ~90% Proteome recovery: Peak area ratios of ApoC III-1/ApoC III-0 and ApoC III-2/ApoC III-0 varied < 15% for the following conditions: 24 storage at room temperature; 3 freeze-thaw cycles; 131 days storage in freezer	[44]
Filtration	On-line immunoaffinity SPE	100 μ L serum	Sensitivity: LOD 1 μ g/mL (transthyretin standards)	[50]
	Tissue homogenization/MwCO filtration, low and high Mw	33 mg medial frontal cortex tissue	Proteome recovery: 0.22 mg	[65]

LOD: limit of detection; LOQ: limit of quantitation; CE-MS: capillary electrophoresis – mass spectrometry; PEPPI: Passively eluting proteins from polyacrylamide gels as intact species; GELF-EE: Gel-eluted liquid fraction entrapment electrophoresis; PBMC: peripheral blood mononuclear cells; TCA: trichloroacetic acid; MWCO: molecular weight cutoff; Mw: molecular weight.

¹Quantitative measures reported where available based on the indicated references; depending on the study, not all measures were relevant, and, to the best of our knowledge, not all measures were indicated in the published results.

sample loss, allow for online coupling to MS, and accommodate low sample volumes are also important priorities.

Though this review focuses on sample preparation and clean-up, downstream sample processing cannot mitigate irreproducibility in sample collection. Practical implementation requires rigorous intra- and inter-laboratory standardization of sample collection and storage. Further, establishing standardized procedures as well as independent validation of protocols across different laboratories is necessary to combat irreproducibility stemming from incomplete methodological reporting and publishing bias toward positive results. Inter-laboratory validations would not only enhance confidence in protocol robustness but could also facilitate method commercialization, thereby supporting broader adoption in large-scale studies. There is also a need for proteoform-specific standards to evaluate the reproducibility of techniques across laboratories and benchmark new techniques against existing methods. However, production of proteoform-specific standards is not possible using approaches such as solid-phase peptide synthesis. Currently, TDP standards are produced via recombinantly synthesized proteins with induced chemical modifications, though such modifications may not accurately reflect endogenous proteoforms.

Toward the goal of standardization, the TDP community has established a consortium to harmonize methodologies across diverse research settings. The consortium has published guidelines for sample cleanup, and current efforts are focused on standardizing data acquisition protocols across different instruments [22]. However, critical gaps remain – most notably, the absence of validated guidelines for sample preparation. These should include: clear recommendations on when to apply specific protocols, acceptable proteoform recovery thresholds, procedures for assessing proteoform loss, standardized data acquisition parameters depending on the type of mass analyzer, and defined criteria supporting the use of recombinant proteins as internal standards. Since the ultimate goal is to advance TDP toward routine clinical application, we believe initiating an inter-laboratory study using commercially available plasma (one specific product that each participating group would need to purchase) would be a valuable step toward standardization and ensuring reproducibility.

Another factor to consider is cost per sample. TDP entails significantly higher costs due to expensive instrumentation, the need for regular maintenance, and stringent operational requirements necessary to ensure accurate results. Operating such instruments also demands specialized expertise, unlike more routine approaches such as ELISA. Furthermore, the sample volume required for TDP is variable and application dependent. When analyzing human plasma, volumes can range from 3–7 μ L up to 1 mL, depending on the target protein's abundance or, for untargeted studies, the desired proteome coverage [9,24,48,82]. This poses a significant limitation in clinical proteomics, where sample volumes are often restricted.

Despite these challenges, we contend that investment toward these goals is justified by the unique value of TDP in monitoring disease progression and elucidating disease mechanisms through the characterization of proteoform-level changes.

Funding

This work was supported by the National Institute of General Medical Sciences of the National Institutes of Health under [grant R35GM147397] to L.F. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Declaration of interest

The authors report there are no competing interests to declare.

Reviewer disclosures

Peer reviewers on this manuscript have no relevant financial or other relationships to disclose.

Authors contributions

Amal Mohamed Kamal wrote the initial draft of the review; Amy Carfagno and Cynthia Nagy contributed to the writing and reviewed the draft; Luca Fornelli supervised and conceptualized the work, reviewed the draft, and approved the final review.

Abbreviations

1D-PAGE	One-dimensional electrophoresis
2D-PAGE	Two-dimensional electrophoresis
ABC	Ammonium bicarbonate
CE	Capillary electrophoresis
ESI	Electrospray ionization
FASP	Filter-aided sample preparation
GELFREE	Gel-eluted liquid fraction entrapment electrophoresis
IA	Immunoaffinity
IEF	Isoelectric focusing
LC	Liquid chromatography
m/z	Mass-to-charge ratio
MALDI	Matrix-assisted laser desorption-ionization
MCW	Methanol chloroform water
MS	Mass spectrometry
Mw	Molecular weight
MWCO	Molecular weight cutoff
NP	Nanoparticle
PAGE	polyacrylamide gel electrophoresis
PEPPI-MS	Passively eluting proteins from polyacrylamide gels as intact species for mass spectrometry
RP-LC	Reversed-phase liquid chromatography
S/N	Signal-to-noise ratio
SAX	Strong anion exchange
SCX	Strong cation exchange
SDS	Sodium dodecyl sulfate
SP3	Single-pot, solid phase-enhanced sample preparation
SPE	Solid-phase extraction
sSEC	Serial size exclusion chromatography
TDP	Top-down proteomics
TME	Transmembrane electrophoresis
USP3	Universal solid-phase protein preparation
WAX	Weak anion exchange
WCX	Weak cation exchange

ORCID

Luca Fornelli  <http://orcid.org/0000-0001-6334-1046>

References

- Birhanu AG. Mass spectrometry-based proteomics as an emerging tool in clinical laboratories. *Clin Proteomics*. 2023;20(1):32. doi: [10.1186/s12014-023-09424-x](#)
- Verrills NM. Clinical proteomics: present and future prospects. *Clin Biochem Rev*. 2006;27(2):99–116.
- Ganepola GA, Nizin J, Rutledge JR, et al. Use of blood-based biomarkers for early diagnosis and surveillance of colorectal cancer. *World J Gastrointest Oncol*. 2014;6(4):83–97. doi: [10.4251/wjgo.v6.i4.83](#)
- Füzéry AK, Levin J, Chan MM, et al. Translation of proteomic biomarkers into FDA approved cancer diagnostics: issues and challenges. *Clin Proteomics*. 2013;10(1):13. doi: [10.1186/1559-0275-10-13](#)
- Cui M, Cheng C, Zhang L. High-throughput proteomics: a methodological mini-review. *Lab Invest*. 2022;102(11):1170–1181. doi: [10.1038/s41374-022-00830-7](#)
- Hoofnagle AN, MacCoss MJ. The need for better validation: evaluating aptamer and proximity extension assays for large-scale clinical proteomics studies. *Clin Chem*. 2025;71(6):618–620. doi: [10.1093/clinchem/hvaf046](#)
- Melani RD, Gerbasi VR, Anderson LC, et al. The blood proteoform atlas: a reference map of proteoforms in human hematopoietic cells. *Science*. 2022;375(6579):411–418. doi: [10.1126/science.aaz5284](#)
- Bogaert A, Fernandez E, Gevaert K. N-Terminal proteoforms in human disease. *Trends Biochem Sci*. 2020;45(4):308–320. doi: [10.1016/j.tibs.2019.12.009](#)
- Sinari S, Koska J, Hu Y, et al. Apo CIII Proteoforms, Plasma Lipids, and Cardiovascular Risk in MESA. *Arterioscler Thromb Vasc Biol*. 2023;43(8):1560–1571. doi: [10.1161/ATVBAHA.123.319035](#)
- Smith LM, Kelleher NL. Consortium for top down P. Proteoform: a single term describing protein complexity. *Nat Methods*. 2013;10(3):186–187. doi: [10.1038/nmeth.2369](#)
- Chait BT. Mass spectrometry: bottom-up or top-down? *Science*. 2006;314(5796):65–66. doi: [10.1126/science.1133987](#)
- Smith LM, Agar JN, Chamot-Rooke J, et al. The human proteoform project: defining the human proteome. *Sci Adv*. 2021;7(46):eabk0734. doi: [10.1126/sciadv.abk0734](#)
- Muthusamy B, Hanumanth G, Suresh S, et al. Plasma proteome database as a resource for proteomics research. *Proteomics*. 2005;5(13):3531–3536. doi: [10.1002/pmic.200401335](#)
- Nesvizhskii AI, Aebersold R. Interpretation of shotgun proteomic data: the protein inference problem. *Mol Cell Proteomics*. 2005;4(10):1419–1440. doi: [10.1074/mcp.R500012-MCP200](#)
- Plubell DL, Kall L, Webb-Robertson BJ, et al. Putting humpty dumpty back together again: what does protein quantification mean in bottom-up proteomics? *J Proteome Res*. 2022;21(4):891–898. doi: [10.1021/acs.jproteome.1c00894](#)
- Roberts DS, Loo JA, Tsybin YO, et al. Top-down proteomics. *Nat Rev Methods Primers*. 2024;4(1). doi: [10.1038/s43586-024-00318-2](#)
- Toby TK, Fornelli L, Kelleher NL. Progress in top-down proteomics and the analysis of proteoforms. *Annu Rev Anal Chem (palo Alto Calif)*. 2016;9(1):499–519. doi: [10.1146/annurev-anchem-071015-041550](#)
- Fenn JB, Mann M, Meng CK, et al. Electrospray ionization for mass spectrometry of large biomolecules. *Science*. 1989;246(4926):64–71. doi: [10.1126/science.2675315](#)
- Loo JA, Edmonds CG, Smith RD. Primary sequence information from intact proteins by electrospray ionization tandem mass spectrometry. *Science*. 1990;248(4952):201–204. doi: [10.1126/science.2326633](#)
- Meier-Credo J, Preiss L, Wüllenweber I, et al. Top-down identification and sequence analysis of small membrane proteins using MALDI-MS/MS. *J Am Soc Mass Spectrom*. 2022;33(7):1293–1302. doi: [10.1021/jasms.2c00102](#)
- Compton PD, Zamdborg L, Thomas PM, et al. On the scalability and requirements of whole protein mass spectrometry. *Anal Chem*. 2011;83(17):6868–6874. doi: [10.1021/ac2010795](#)
- Donnelly DP, Rawlins CM, DeHart CJ, et al. Best practices and benchmarks for intact protein analysis for top-down mass spectrometry. *Nat Methods*. 2019;16(7):587–594. doi: [10.1038/s41592-019-0457-0](#)
- Wilkins JT, Seckler HS, Rink J, et al. Spectrum of apolipoprotein AI and Apolipoprotein AII proteoforms and their associations with indices of cardiometabolic Health: the CARDIA study. *J Am Heart Assoc*. 2021;10(17):e019890. doi: [10.1161/JAHA.120.019890](#)
- Forte E, Sanders JM, Pla I, et al. Top-down proteomics identifies plasma proteoform signatures of liver cirrhosis progression. *Mol Cell Proteomics*. 2024;23(12):100876. doi: [10.1016/j.mcpro.2024.100876](#)
- Moran AB, Dominguez-Vega E, Nouta J, et al. Profiling the proteoforms of urinary prostate-specific antigen by capillary electrophoresis – mass spectrometry. *J Proteomics*. 2021;238:104148. doi: [10.1016/j.jpro.2021.104148](#)
- Manconi B, Liori B, Cabras T, et al. Top-down proteomic profiling of human saliva in multiple sclerosis patients. *J Proteomics*. 2018;187:212–222. doi: [10.1016/j.jpro.2018.07.019](#)
- Ahsan N, Fornelli L, Najjar FZ, et al. Proteomics evaluation of five economical commercial abundant protein depletion kits for enrichment of diseases-specific biomarkers from blood serum. *Proteomics*. 2023;23(20):e2300150. doi: [10.1002/pmic.202300150](#)
- Thomas SL, Thacker JB, Schug KA, et al. Sample preparation and fractionation techniques for intact proteins for mass spectrometric analysis. *J Sep Sci*. 2021;44(1):211–246. doi: [10.1002/jssc.202000936](#)
- Nickerson JL, Baghalabadi V, Rajendran S, et al. Recent advances in top-down proteome sample processing ahead of MS analysis. *Mass Spectrom Rev*. 2023;42(2):457–495. doi: [10.1002/mas.21706](#)
- Karki S, Shi F, Archer JJ, et al. Direct analysis of proteins from solutions with high salt concentration using laser electrospray mass spectrometry. *J Am Soc Mass Spectrom*. 2018;29(5):1002–1011. doi: [10.1007/s13361-018-1893-2](#)
- Kachuk C, Doucette AA. The benefits (and misfortunes) of SDS in top-down proteomics. *J Proteomics*. 2018;175:75–86. doi: [10.1016/j.jpro.2017.03.002](#)
- Cai W, Hite ZL, Lyu B, et al. Temperature-sensitive sarcomeric protein post-translational modifications revealed by top-down proteomics. *J Mol Cell Cardiol*. 2018;122:11–22. doi: [10.1016/j.yjmcc.2018.07.247](#)
- Kaulich PT, Jeong K, Kohlbacher O, et al. Influence of different sample preparation approaches on proteoform identification by top-down proteomics. *Nat Methods*. 2024;21(12):2397–2407. doi: [10.1038/s41592-024-02481-6](#)
- Guray MZ, Zheng S, Doucette AA. Mass spectrometry of intact proteins reveals +98 u chemical artifacts following precipitation in acetone. *J Proteome Res*. 2017;16(2):889–897. doi: [10.1021/acs.jproteome.6b00841](#)
- Botelho D, Wall MJ, Vieira DB, et al. Top-down and bottom-up proteomics of SDS-containing solutions following mass-based separation. *J Proteome Res*. 2010;9(6):2863–2870. doi: [10.1021/pr900949p](#)
- Yang Z, Shen X, Chen D, et al. Toward a universal sample preparation method for denaturing top-down proteomics of complex proteomes. *J Proteome Res*. 2020;19(8):3315–3325. doi: [10.1021/acs.jproteome.0c00226](#)
- Kline JT, Belford MW, Boeser CL, et al. Orbitrap mass spectrometry and high-field asymmetric waveform ion mobility spectrometry (FAIMS) enable the in-depth analysis of human serum proteoforms. *J Proteome Res*. 2023;22(11):3418–3426. doi: [10.1021/acs.jproteome.3c00488](#)
- Takemori A, Sugiyama N, Kline JT, et al. Gel-based sample fractionation with SP3-purification for top-down proteomics. *J Proteome Res*. 2025;24(2):850–860. doi: [10.1021/acs.jproteome.4c00941](#)
- Naldi M, Baldassarre M, Nati M, et al. Mass spectrometric characterization of human serum albumin dimer: a new potential biomarker in chronic liver diseases. *J Pharm Biomed Anal*. 2015;112:169–175. doi: [10.1016/j.jpba.2014.12.001](#)
- Cohn EJ, Strong LE, Hughes WL, et al. Preparation and properties of serum and plasma proteins; a system for the separation into

- fractions of the protein and lipoprotein components of biological tissues and fluids. *J Am Chem Soc.* 1946;68(3):459–475. doi: 10.1021/ja01207a034
41. Leblanc Y, Bihoreau N, Chevreux G. Characterization of human serum albumin isoforms by ion exchange chromatography coupled on-line to native mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci.* 2018;1095:87–93. doi: 10.1016/j.jchromb.2018.07.014
 42. Lieu LB, Hinkle JD, Syka JEP, et al. Leveraging ion-ion and ion-photon activation to improve the sequencing of proteins carrying multiple disulfide bonds: the human serum albumin case study. *J Am Soc Mass Spectrom.* 2024;35(12):3265–3273. doi: 10.1021/jasms.4c00391
 43. Seckler HDS, Fornelli L, Mutharasan RK, et al. A targeted, differential top-down proteomic methodology for comparison of ApoA-I proteoforms in individuals with high and low HDL efflux capacity. *J Proteome Res.* 2018;17(6):2156–2164. doi: 10.1021/acs.jproteome.8b00100
 44. Jian W, Edom RW, Wang D, et al. Relative quantitation of glycoisoforms of intact apolipoprotein C3 in human plasma by liquid chromatography-high-resolution mass spectrometry. *Anal Chem.* 2013;85(5):2867–2874. doi: 10.1021/ac3034757
 45. Chen A, Aquino RM, Vidal HA, et al. A liquid chromatography-high-resolution mass spectrometry method for separation and identification of hemoglobin variant subunits with mass shifts less than 1 Da. *J Mass Spectrom Adv Clin Lab.* 2025;35:1–7. doi: 10.1016/j.jmsacl.2025.01.002
 46. Thermo Scientific Tech tip #64: immunoprecipitation (IP) technical guide and protocols 2009. Available from: <https://assets.thermo-fisher.com/TFS-Assets/LSG/Application-Notes/TR0064-Immunoprecipitation-guide.pdf>
 47. Mazzarino M, Melis I, Quaresima E, et al. Detection of synthetic analogues of insulin-like growth factor 1 in different biological fluids by liquid chromatography quadrupole time-of-flight mass spectrometry: comparison of different immunoaffinity protocols. *Anal Bioanal Chem.* 2023;415(24):6117–6131. doi: 10.1007/s00216-023-04885-3
 48. van den Wildenberg SAH, Genet S, Broeren MAC, et al. Immunoaffinity intact top-down mass spectrometry for quantification of neuron-specific enolase Gamma, a low-abundance protein biomarker. *Anal Chem.* 2025;97(1):516–525. doi: 10.1021/acs.analchem.4c04677
 49. Des Soye BJ, McGee JP, Hollas MAR, et al. Automated immunoprecipitation, sample preparation, and individual ion mass spectrometry platform for proteoforms. *Anal Chem.* 2024. doi: 10.1021/acs.analchem.4c01962
 50. Pero-Gascon R, Pont L, Sanz-Nebot V, et al. On-line immunoaffinity solid-phase extraction capillary electrophoresis-mass spectrometry for the analysis of serum transthyretin. Vol. 1972. *Clinical Applications of Capillary Electrophoresis: Methods Mol Biol*; 2019. p. 57–76. doi: 10.1007/978-1-4939-9213-3_5
 51. Salim H, Pont L, Gimenez E, et al. On-line aptamer affinity solid-phase extraction direct mass spectrometry for the rapid analysis of α -synuclein in blood. *Anal Chim Acta.* 2023;1256:341149. doi: 10.1016/j.aca.2023.341149
 52. Kwon D. The antibodies don't work! The race to rid labs of molecules that ruin experiments. *Nature.* 2024;635(8037):26–28. doi: 10.1038/d41586-024-03590-0
 53. Scheer JM, Ryan CA. A method for the quantitative recovery of proteins from polyacrylamide gels. *Anal Biochem.* 2001;298(1):130–132. doi: 10.1006/abio.2001.5384
 54. Meng F, Cargile BJ, Patrie SM, et al. Processing complex mixtures of intact proteins for direct analysis by mass spectrometry. *Anal Chem.* 2002;74(13):2923–2929. doi: 10.1021/ac020049i
 55. Takemori A, Kaulich PT, Cassidy L, et al. Size-based proteome fractionation through polyacrylamide gel electrophoresis combined with LC-FAIMS-MS for in-depth top-down proteomics. *Anal Chem.* 2022;94(37):12815–12821. doi: 10.1021/acs.analchem.2c02777
 56. Takemori A, Butcher DS, Harman VM, et al. 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. doi: 10.1021/acs.jproteome.0c00303
 57. Wright EP, Partridge MA, Padula MP, et al. Top-down proteomics: enhancing 2D gel electrophoresis from tissue processing to high-sensitivity protein detection. *Proteomics.* 2014;14(7–8):872–889. doi: 10.1002/pmic.201300424
 58. Wilkins MR, Pasquali C, Appel RD, et al. From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Nat Biotechnol.* 1996;14(1):61–65. doi: 10.1038/nbt0196-61
 59. Almuslehi MSM, Sen MK, Shortland PJ. Histological and top-down proteomic analyses of the visual pathway in the Cuprizone demyelination Model. *J Mol Neurosci.* 2022;72(6):1374–1401. doi: 10.1007/s12031-022-01997-w
 60. Tran JC, Doucette A. Gel-eluted liquid fraction entrapment electrophoresis: an electrophoretic method for broad molecular weight range proteome separation. *Anal Chem.* 2008;80(5):1568–1573. doi: 10.1021/ac702197w
 61. Tran JC, Zamdborg L, Ahlf DR, et al. Mapping intact protein isoforms in discovery mode using top-down proteomics. *Nature.* 2011;480(7376):254–258. doi: 10.1038/nature10575
 62. Skinner OS, Do Vale LH, Catherman AD, et al. Native GELFrEE: a new separation technique for biomolecular assemblies. *Anal Chem.* 2015;87(5):3032–3038. doi: 10.1021/acs04678d
 63. Tran JC, Doucette A. Multiplexed size separation of intact proteins in solution phase for mass spectrometry. *Anal Chem.* 2009;81(15):6201–6209. doi: 10.1021/ac900729r
 64. Toby TK, Fornelli L, Szrentic K, et al. A comprehensive pipeline for translational top-down proteomics from a single blood draw. *Nat Protoc.* 2019;14(1):119–152. doi: 10.1038/s41596-018-0085-7
 65. Fulcher JM, Makaju A, Moore RJ, et al. Enhancing top-down proteomics of brain tissue with FAIMS. *J Proteome Res.* 2021;20(5):2780–2795. doi: 10.1021/acs.jproteome.1c00049
 66. Pont L, Benavente F, Barbosa J, et al. Analysis of transthyretin in human serum by capillary zone electrophoresis electrospray ionization time-of-flight mass spectrometry. Application to familial amyloidotic polyneuropathy type I. *Electrophoresis.* 2015;36(11–12):1265–1273. doi: 10.1002/elps.201400590
 67. Bladergroen MR, van der Burgt YE. Solid-phase extraction strategies to surmount body fluid sample complexity in high-throughput mass spectrometry-based proteomics. *J Anal Methods Chem.* 2015;2015:250131. doi: 10.1155/2015/250131
 68. Tanna NN, Lame ME, Wrona M. Development of an UPLC/MS-MS method for quantification of intact IGF-I from human serum. *Bioanalysis.* 2020;12(1):53–65. doi: 10.1555/bio-2019-0234
 69. Maráková K, Renner BJ, Thomas SL, et al. Solid phase extraction as sample pretreatment method for top-down quantitative analysis of low molecular weight proteins from biological samples using liquid chromatography – triple quadrupole mass spectrometry. *Anal Chim Acta.* 2023;1243:340801. doi: 10.1016/j.aca.2023.340801
 70. Feng S, Yang N, Pennathur S, et al. Enrichment of glycoproteins using Nanoscale chelating concanavalin a monolithic capillary chromatography. *Anal Chem.* 2009;81(10):3776–3783. doi: 10.1021/ac900085k
 71. Zeng Z, Hincapie M, Pitteri SJ, et al. A proteomics platform combining depletion, multi-lectin affinity chromatography (M-LAC), and isoelectric focusing to study the breast cancer proteome. *Anal Chem.* 2011;83(12):4845–4854. doi: 10.1021/ac2002802
 72. Wu D, Li J, Struwe WB, et al. Probing N-glycoprotein microheterogeneity by lectin affinity purification-mass spectrometry analysis. *Chem Sci.* 2019;10(19):5146–5155. doi: 10.1039/C9SC00360F
 73. Lattmann E, Rass L, Tognetti M, et al. Size-exclusion chromatography combined with DIA-MS enables deep proteome profiling of extracellular vesicles from melanoma plasma and serum. *Cell Mol Life Sci.* 2024;81(1):90. doi: 10.1007/s00018-024-05137-y

74. Cai W, Tucholski T, Chen B, et al. Top-down proteomics of large proteins up to 223 kDa enabled by serial size exclusion chromatography strategy. *Anal Chem.* 2017;89(10):5467–5475. doi: [10.1021/acs.analchem.7b00380](https://doi.org/10.1021/acs.analchem.7b00380)
75. Blume JE, Manning WC, Troiano G, et al. Rapid, deep and precise profiling of the plasma proteome with multi-nanoparticle protein corona. *Nat Commun.* 2020;11(1):3662. doi: [10.1038/s41467-020-17033-7](https://doi.org/10.1038/s41467-020-17033-7)
76. Ferdosi S, Tangeysh B, Brown TR, et al. Engineered nanoparticles enable deep proteomics studies at scale by leveraging tunable nano–bio interactions. *Proc Natl Acad Sci U S A.* 2022;119(11):e2106053119. doi: [10.1073/pnas.2106053119](https://doi.org/10.1073/pnas.2106053119)
77. Bertrand N, Grenier P, Mahmoudi M. Mechanistic understanding of in vivo protein corona formation on polymeric nanoparticles and impact on pharmacokinetics. *Nat Commun.* 2017;8(1):777. doi: [10.1038/s41467-017-00600-w](https://doi.org/10.1038/s41467-017-00600-w)
78. Sadeghi SA, Ashkarran AA, Wang Q, et al. Mass spectrometry-based top-down proteomics in nanomedicine: Proteoform-specific measurement of protein corona. *ACS Nano.* 2024;18(38):26024–26036. doi: [10.1021/acsnano.4c04675](https://doi.org/10.1021/acsnano.4c04675)
79. Huang CF, Hollas MA, Sanchez A, et al. Deep profiling of plasma proteoforms with engineered nanoparticles for top-down proteomics. *J Proteome Res.* 2024;23(10):4694–4703. doi: [10.1021/acs.jproteome.4c00621](https://doi.org/10.1021/acs.jproteome.4c00621)
80. Chen B, Hwang L, Ochowicz W, et al. Coupling functionalized cobalt ferrite nanoparticle enrichment with online LC/MS/MS for top-down phosphoproteomics. *Chem Sci.* 2017;8(6):4306–4311. doi: [10.1039/C6SC05435H](https://doi.org/10.1039/C6SC05435H)
81. Kim KH, Compton PD, Tran JC, et al. Online matrix removal platform for coupling gel-based separations to whole protein electrospray ionization mass spectrometry. *J Proteome Res.* 2015;14(5):2199–2206. doi: [10.1021/pr501331q](https://doi.org/10.1021/pr501331q)
82. Santos Seckler HD, Park HM, Lloyd-Jones CM, et al. New interface for faster proteoform analysis: immunoprecipitation coupled with SampleStream-Mass Spectrometry. *J Am Soc Mass Spectrom.* 2021;32(7):1659–1670. doi: [10.1021/jasms.1c00026](https://doi.org/10.1021/jasms.1c00026)
83. Park HM, Winton VJ, Drader JJ, et al. Novel interface for high-throughput analysis of Biotherapeutics by electrospray mass spectrometry. *Anal Chem.* 2020;92(2):2186–2193. doi: [10.1021/acs.analchem.9b04826](https://doi.org/10.1021/acs.analchem.9b04826)
84. Palmblad M, Vogel JS. Quantitation of binding, recovery and desalting efficiency of peptides and proteins in solid phase extraction micropipette tips. *J Chromatogr B Analyt Technol Biomed Life Sci.* 2005;814(2):309–313. doi: [10.1016/j.jchromb.2004.10.052](https://doi.org/10.1016/j.jchromb.2004.10.052)
85. Mommen GP, Meiring HD, Heck AJ, et al. Mixed-bed ion exchange chromatography employing a salt-free pH gradient for improved sensitivity and compatibility in MudPIT. *Anal Chem.* 2013;85(14):6608–6616. doi: [10.1021/ac400995e](https://doi.org/10.1021/ac400995e)
86. Nagaraj N, Lu A, Mann M, et al. Detergent-based but gel-free method allows identification of several hundred membrane proteins in single LC-MS runs. *J Proteome Res.* 2008;7(11):5028–5032. doi: [10.1021/pr800412j](https://doi.org/10.1021/pr800412j)
87. Kachuk C, Faulkner M, Liu F, et al. Automated SDS depletion for mass spectrometry of intact membrane proteins through transmembrane electrophoresis. *J Proteome Res.* 2016;15(8):2634–2642. doi: [10.1021/acs.jproteome.6b00199](https://doi.org/10.1021/acs.jproteome.6b00199)
88. Said HH, Doucette AA. Enhanced electrophoretic depletion of sodium dodecyl sulfate with methanol for membrane proteome analysis by mass spectrometry. *Proteomes.* 2024;12(1):5. doi: [10.3390/proteomes12010005](https://doi.org/10.3390/proteomes12010005)
89. Doucette A, Crowell A, Doucette A, et al. Precipitation of detergent-containing samples for top-down and bottom-up proteomics. *Proteomics Technol Appl: Intech.* 2019:1–17. doi: [10.5772/intechopen.73445](https://doi.org/10.5772/intechopen.73445)
90. Zellner M, Winkler W, Hayden H, et al. Quantitative validation of different protein precipitation methods in proteome analysis of blood platelets. *Electrophoresis.* 2005;26(12):2481–2489. doi: [10.1002/elps.200410262](https://doi.org/10.1002/elps.200410262)
91. Puchades M, Westman A, Blennow K, et al. Removal of sodium dodecyl sulfate from protein samples prior to matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun Mass Spectrom.* 1999;13(5):344–349. doi: [10.1002/\(SICI\)1097-0231\(19990315\)13:5<344::AID-RCM489>3.0.CO;2-V](https://doi.org/10.1002/(SICI)1097-0231(19990315)13:5<344::AID-RCM489>3.0.CO;2-V)
92. Wessel D, Flügge UI. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal Biochem.* 1984;138(1):141–143. doi: [10.1016/0003-2697\(84\)90782-6](https://doi.org/10.1016/0003-2697(84)90782-6)
93. Crowell AM, Wall MJ, Doucette AA. Maximizing recovery of water-soluble proteins through acetone precipitation. *Anal Chim Acta.* 2013;796:48–54. doi: [10.1016/j.aca.2013.08.005](https://doi.org/10.1016/j.aca.2013.08.005)
94. Hughes CS, Moggridge S, Muller T, et al. Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat Protoc.* 2019;14(1):68–85. doi: [10.1038/s41596-018-0082-x](https://doi.org/10.1038/s41596-018-0082-x)
95. Batth TS, Tollenaere MX, Ruther P, et al. Protein aggregation capture on microparticles enables multipurpose proteomics sample preparation. *Mol Cell Proteomics.* 2019;18(5):1027–1035. doi: [10.1074/mcp.TIR118.001270](https://doi.org/10.1074/mcp.TIR118.001270)
96. Dagley LF, Infusini G, Larsen RH, et al. Universal solid-phase protein preparation (USP(3)) for bottom-up and top-down proteomics. *J Proteome Res.* 2019;18(7):2915–2924. doi: [10.1021/acs.jproteome.9b00217](https://doi.org/10.1021/acs.jproteome.9b00217)
97. Nickerson JL, Doucette AA. Rapid and quantitative protein precipitation for proteome analysis by mass spectrometry. *J Proteome Res.* 2020;19(5):2035–2042. doi: [10.1021/acs.jproteome.9b00867](https://doi.org/10.1021/acs.jproteome.9b00867)
98. Smeriglio N, Li H, Mazli W, et al. Contaminant spot check and removal assay (ContamSPOT) for mass spectrometry analysis. *Anal Chem.* 2024;96(6):2574–2581. doi: [10.1021/acs.analchem.3c05020](https://doi.org/10.1021/acs.analchem.3c05020)
99. Rusconi F, Valton É, Nguyen R, et al. Quantification of sodium dodecyl sulfate in microliter-volume biochemical samples by visible light spectroscopy. *Anal Biochem.* 2001;295(1):31–37. doi: [10.1006/abio.2001.5164](https://doi.org/10.1006/abio.2001.5164)