J Biochem Technol (2026) 17(3): 83-91
https://doi.org/10.51847/z4LIiw9f0e
ISSN: 0974-2328

Beyond Total Protein Abundance: A Critical Review of Proteoform Diversity, Post-Translational Modification, and the Limits of Functional Interpretation in Mass Spectrometry-Based Proteomics

Thomas Berger*, Nadia Frei, Lukas Meier, Sara Brunner, Jonas Kaufmann

Received: 24 July 2026 / Received in revised form: 27 August 2026, Accepted: 30 August 2026, Published online: 20 October 2026

Abstract

Mass spectrometry-based proteomics routinely quantifies proteins through signals derived from peptides or intact molecular species, yet a protein name or total abundance value can encompass multiple proteoforms with different sequences, processing histories, post-translational modifications, conformational states, interaction partners, and cellular contexts. This critical review examines how those layers of molecular heterogeneity limit functional interpretation when proteomic measurements are collapsed to protein-level abundance. Evidence from 2017–2024 is evaluated across proteoform classification, post-translational modification analysis, bottom-up and top-down mass spectrometry, ion-mobility and orthogonal separation strategies, statistical treatment of modified peptides, and emerging state-sensitive proteomic measurements. The analysis distinguishes molecular detection from proteoform assignment, modification localization, biological-state characterization, and causal functional inference. Particular attention is given to information that may be lost during digestion, sample preparation, precursor reconstruction, database inference, or aggregation of heterogeneous molecular forms. Proteoform-resolved approaches preserve more molecular connectivity, although their interpretation remains constrained by sensitivity, fragmentation, sample preparation, computational deconvolution, and incomplete contextual validation. The central conclusion is that total protein abundance remains informative for many biological questions but cannot be assumed to represent a single molecular or functional state. Functionally informative proteomics therefore requires explicit accounting for proteoform identity, modification topology, measurement uncertainty, biological context, and orthogonal evidence whenever these dimensions materially affect the inference being made.

Keywords: Proteoforms, Top-down proteomics, Bottom-up proteomics, Post-translational modification, Mass spectrometry, Functional proteomics

Thomas Berger*
Department of Proteomics, University of Bern, Bern, Switzerland.

Nadia Frei
Department of Biochemistry, University of Zurich, Zurich, Switzerland.

Lukas Meier
Department of Molecular Biology, University of Basel, Basel, Switzerland.

Sara Brunner
Department of Proteomics, University of Geneva, Geneva, Switzerland.

Jonas Kaufmann
Department of Biochemistry, University of Lausanne, Lausanne, Switzerland.

***E-mail:** thomas.berger@unibe.ch

Introduction

Mass spectrometry-based proteomics has achieved remarkable depth, throughput, and quantitative precision, yet the biological entity summarized by a protein-level abundance value is often more heterogeneous than the reported label implies. A single gene product can exist as multiple molecular forms arising from sequence variation, alternative processing, proteolytic maturation, chemical modification, post-translational modification (PTM), and combinations of these events. The proteoform concept was introduced to make this molecular specificity explicit and to distinguish an experimentally relevant molecular species from a gene- or protein-name abstraction (Aebersold *et al.*, 2018; Smith & Kelleher, 2018). This distinction becomes important whenever two molecules assigned to the same canonical protein differ in catalytic competence, binding behavior, localization, turnover, structural state, or participation in molecular complexes.

The practical challenge is considerably larger than terminology. Comprehensive characterization of the human proteoform repertoire remains incomplete, motivating efforts such as the Human Proteoform Project to improve intact-protein measurement, identification, standardization, and reference resources (Smith *et al.*, 2021). Consequently, many current proteomic experiments operate at a level where proteins are quantified accurately while the molecular composition underlying that quantity is only partly known. An observed twofold or tenfold difference in a protein-level signal, for example, may represent altered abundance of most molecular forms, selective expansion of one modified form, redistribution among several forms with little change in total amount, or a mixture of these processes. The abundance measurement can be technically valid while remaining biologically underdetermined.

Bottom-up proteomics does not necessarily erase every indication of sub-protein heterogeneity. Peptides derived from different regions of a protein can exhibit divergent quantitative behavior,

© 2026 The Author(s). This is an **Open Access** article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0).
https://creativecommons.org/licenses/by/4.0/deed.en

and systematic covariance among peptide subsets has been used to infer putative functional proteoform groups from conventional bottom-up datasets (Bludau *et al.*, 2021). Such approaches demonstrate that digestion-based data can retain signatures of molecular heterogeneity. Their inferential status must nevertheless remain explicit: peptide covariance does not by itself reconstruct the intact combination of sequence features and PTMs that existed on any particular molecule before digestion.

The same need for calibrated interpretation applies to intact-protein measurements. Proteoform identification is not a binary state. A five-level classification system has been proposed to communicate how specifically an experimentally observed form is characterized, distinguishing fully specified proteoforms from assignments with unresolved sequence or modification features (Smith *et al.*, 2019). Standardized representation is equally important. ProForma 2.0 permits known and ambiguous modifications, terminal modifications, cross-links and related molecular information to be encoded in a machine-readable form (LeDuc *et al.*, 2022). These standards improve transparency because uncertainty can remain visible throughout analysis rather than being converted into an apparently definitive molecular label.

This review therefore treats total abundance, molecular identity, PTM topology, analytical reconstruction, and biological function as related but non-equivalent evidentiary levels. The central question is not whether protein abundance measurements are useful; they are indispensable across much of contemporary biology. The question is when the biological inference being requested depends on molecular distinctions that an abundance value does not preserve. The following sections first examine proteoform diversity as a source of hidden molecular heterogeneity, then analyze how PTM information and molecular connectivity can be lost, before evaluating the analytical architectures used to recover or infer that information and the additional evidence needed to connect proteoforms with biological function.

Proteoform Diversity and the Biological Limits of Protein-Level Quantification

A proteoform is meaningful analytically because intact molecular identity can integrate several features that are separated by protein-level summaries. Accurate mass may constrain composition, while fragmentation, sequence information, PTM localization, terminal processing, and other measurements progressively increase characterization specificity. Reviews of top-down proteomics emphasize that detection of an intact mass is only an initial step: rigorous characterization may require assignment of the underlying amino-acid sequence together with the identities and locations of relevant modifications (Po & Eyers, 2023). Large proteins, low-abundance forms, highly modified species, and coexisting near-isomers can therefore remain incompletely characterized even when an intact signal is confidently detected.

Biological heterogeneity becomes more difficult to resolve when sample quantity declines. Single-cell proteoform analysis is beginning to expose molecular states that are obscured by bulk averaging, but the same reduction in input material intensifies sensitivity, dynamic-range, throughput, and identification-depth constraints (Su *et al.*, 2024). An undetected proteoform in a single cell cannot automatically be interpreted as biologically absent. Sampling and measurement limitations remain alternative explanations, particularly for low-abundance or poorly ionizing species. Thus, increasing biological resolution may simultaneously increase the importance of explicit analytical uncertainty.

The measurable proteoform repertoire is also conditioned before ions reach the mass spectrometer. A multicenter community study of native top-down mass spectrometry demonstrated that proteoform and protein-complex characterization depends on instrumentation, acquisition conditions, and analysis workflows (Habeck *et al.*, 2024). Sample preparation introduces an earlier selection layer. Comparative work in top-down proteomics has shown that lysis, reduction and alkylation strategies, enrichment procedures, and fractionation influence which proteoforms are ultimately observed and can introduce or accentuate chemical modifications (Kaulich *et al.*, 2024). Consequently, differences between observed proteoform populations may reflect biology, analytical selection, or both.

Molecular forms can remain distinct even when conventional observables appear identical. Positional PTM isomers provide a particularly clear example because molecules can have the same elemental composition and nominal mass while carrying a modification at different residues. Top-down ion-mobility separation of histone H4 proteoforms has demonstrated that some such isomeric species can be resolved through gas-phase mobility characteristics (Berthias *et al.*, 2023). This result illustrates a broader principle: total protein abundance and intact mass alone may both fail to preserve the molecular information necessary to distinguish biologically relevant proteoform states.

Orthogonal measurements can reduce this ambiguity. Isotopic chemical labeling for cysteine counting has been combined with intact-mass measurements to constrain proteoform assignment directly in tissue, adding an independent molecular property to accurate mass (Pavek *et al.*, 2023). The approach does not make identification uncertainty disappear, because distinct sequences may still share compatible molecular constraints, but it demonstrates how convergent observables can narrow the set of plausible identities.

The analytical consequences of these distinctions are summarized in **Table 1**. They show why a protein-level quantity can remain reproducible while the molecular state represented by that quantity remains unresolved.

Table 1. Molecular sources of proteoform heterogeneity that can decouple protein abundance from functional state

Proteoform-defining source	Molecular distinction retained at proteoform level	What a total protein abundance value can collapse	Analytical route that may expose the distinction	Principal interpretive boundary

Sequence variation and alternative coding products	Distinct primary sequences assigned to a related protein identity	Relative contributions of individual sequence forms	Intact-mass analysis with sequence-resolving fragmentation	Sequence assignment may remain incomplete for closely related forms
Proteolytic processing and terminal maturation	Cleaved, truncated, or processed molecular species	Mature and precursor forms within one protein total	Top-down fragmentation and terminal characterization	Processing may be biological or introduced during handling
Post-translational modification	Modified and unmodified molecular populations	Occupancy and distribution of modification states	PTM-sensitive bottom-up analysis or intact-proteoform characterization	Identification of a modification does not necessarily localize it
Combinatorial PTM patterns	Co-occurring modifications on the same intact molecule	Whether modifications coexist on one proteoform or on separate molecules	Top-down or sufficiently long-range sequence analysis	Increasing combinatorial complexity reduces characterization depth
Positional modification isomerism	Identical mass with different modification sites	Site-specific molecular states	Fragmentation, ion mobility, or other orthogonal separation	Isomer separation can remain system dependent
Native association or structural state	Proteoform in a particular complex or conformational environment	State-specific abundance within pooled protein measurements	Native MS and complementary structural-state methods	Gas-phase or preparation effects can alter observed states
Cell- or tissue-restricted proteoform composition	Context-specific molecular distributions	Averaged forms across heterogeneous populations	Spatial, single-cell, or highly resolved sampling strategies	Low-input sensitivity and missing observations limit certainty

A useful representation of the measurement problem must therefore show both information preservation and information loss without implying that any analytical platform provides complete molecular truth. **Figure 1** compresses the principal transitions described above and distinguishes direct molecular observation from reconstruction and unresolved identity.

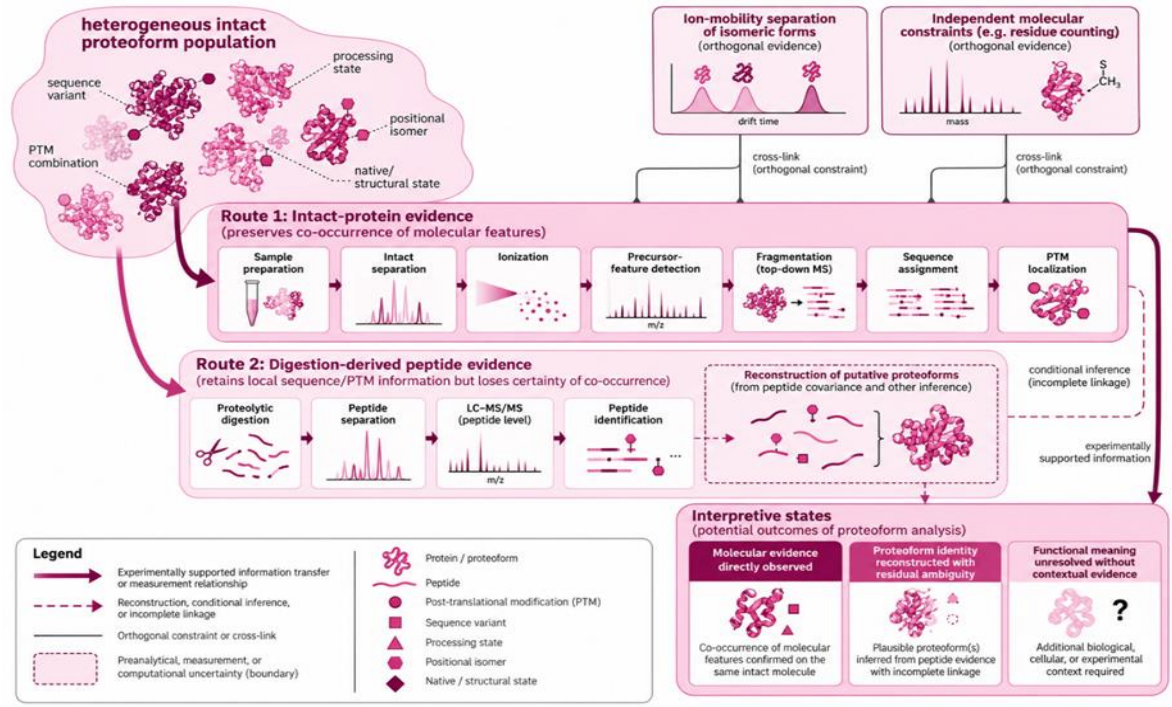


Figure 1. Proteoform information preserved, fragmented, and reconstructed across mass-spectrometric measurement routes

Post-Translational Modification and the Loss of Molecular Connectivity

PTMs create a second layer of interpretive difficulty because identifying a modified peptide is not equivalent to identifying the complete proteoform that carried it. Integrative analyses combining top-down and bottom-up mass spectrometry with curated annotations have shown that these evidence sources can complement one another when assigning proteoform PTMs (Chen *et al.*, 2023). Bottom-up measurements often provide high-quality

local sequence and modification evidence, whereas intact measurements preserve co-occurrence across a larger molecular span. Database annotations can add prior biological knowledge, but they remain distinct from a modification observed directly in the analyzed sample.

Fragmentation chemistry further determines which modification evidence survives into the spectrum. Diagnostic features in tandem mass spectra can support identification of specific modified peptide classes, while their absence may reflect fragmentation behavior as readily as biological absence (Geiszler *et al.*, 2023). This issue becomes particularly acute for labile PTMs. Search strategies designed to accommodate modification loss or characteristic fragments can recover evidence that conventional peptide-search assumptions would miss or misassign (Polasky *et al.*, 2023). The analytical workflow therefore participates in determining which PTM states become visible.

Quantitative interpretation requires an additional distinction between changes in protein abundance and changes in modification usage. Statistical frameworks developed for PTM-aware proteomics can model differential abundance at the modified-peptide or peptidoform level while accounting for the underlying protein signal (Demeulemeester *et al.*, 2024). This separation matters because an increase in a modified peptide can arise from increased abundance of the entire protein, altered fractional modification, or both. Protein-level normalization cannot resolve that ambiguity unless the data and model explicitly distinguish these contributions.

The biological relevance of preserving modification identity is evident in modification systems with extensive molecular heterogeneity. Clinical glyco(proteo)mics demonstrates that proteins carrying different glycan structures may have distinct biological or diagnostic associations even when they derive from the same canonical protein (van der Burgt & Wuhrer, 2023). Large-scale analysis of the human phosphoproteome likewise shows that functional evidence is distributed unevenly among phosphosites, with some sites strongly implicated in regulation and many others having limited or uncertain functional support (Ochoa *et al.*, 2020). The presence of a PTM should therefore not be converted automatically into a claim of functional control.

Localization uncertainty introduces another boundary. Methods that aggregate multiple evidence streams and explicitly control false-localization rates can materially alter the set of PTM sites considered confidently assigned (Camacho *et al.*, 2023). This means that downstream mechanistic interpretation depends not only on whether a modified peptide was detected but also on whether the modification was localized to the biologically relevant residue with adequate support.

These analytical differences are compared in **Table 2**. The central contrast concerns the molecular relationships each strategy can preserve or reconstruct, not a universal ranking of platforms.

Table 2. Molecular information preserved and inferred across bottom-up, middle-down, and top-down proteomic strategies

Analytical strategy	Principal measured unit	PTM information commonly accessible	Molecular connectivity retained	Major strength for proteoform interpretation	Dominant limitation for functional inference
Bottom-up proteomics	Proteolytic peptide	High local sensitivity for many peptide-associated PTMs and modification sites	Primarily within individual peptides; long-range co-occurrence is usually lost	Deep and scalable measurement of peptide-level abundance and localized modification evidence	Intact combinations of sequence features and distant PTMs must be inferred
Middle-down proteomics	Longer peptide or protein fragment	Multiple nearby PTMs and sequence features can remain linked	Intermediate-range connectivity across longer molecular segments	Extends local connectivity while retaining more tractable fragmentation than many intact proteins	Complete intact-proteoform composition may still remain unresolved
Top-down proteomics	Intact proteoform	Intact mass shifts, sequence variants, terminal processing and combinations of PTMs, subject to fragmentation quality	Co-occurrence is retained across the intact precursor before fragmentation	Direct interrogation of whole molecular forms and their combined modifications	Sensitivity, separation, deconvolution and complete sequence/PTM localization become increasingly demanding
Orthogonally constrained top-down analysis	Intact proteoform plus an independent molecular property	Same intact-level evidence supplemented by mobility, residue count, native	Intact connectivity retained with an additional discriminating dimension	Reduces specific identity or isomer ambiguities	Orthogonal constraints do not necessarily establish unique biological function

		association, or related constraints			
Integrative top-down/bottom-up analysis	Linked intact and peptide evidence	Localized peptide evidence can support intact mass-shift interpretation	Connectivity is preserved directly only where supported by intact evidence; other links remain reconstructed	Triangulates complementary strengths of intact and peptide measurements	Conflicting coverage, annotation, and confidence levels require explicit reconciliation

The distinction developed in this section is therefore narrower than the general observation that proteoforms are heterogeneous. PTM analysis introduces a specific molecular-connectivity problem: after digestion, a modification can be localized accurately on one peptide while its relationship to distant modifications, sequence variants, processing states, or other features on the original molecule remains unknown. Conversely, intact detection may preserve their co-occurrence while leaving individual sites incompletely localized. Functional interpretation becomes strongest when these forms of evidence are treated as complementary constraints rather than interchangeable substitutes.

Bottom-Up and Top-Down Mass Spectrometry: Analytical Capabilities and Interpretive Limits

Bottom-up proteomics achieves substantial analytical depth by converting proteins into peptides, but protein-level interpretation then depends on inference from those peptide observations. Shared peptides, indistinguishable isoforms, database composition, and protein-group construction can affect which protein entities are considered confidently identified (The *et al.*, 2022). Curated resources such as UniProt help connect peptide observations with sequence, PTM, and functional annotations, yet experimentally supported and predicted annotations remain distinct evidence classes (Bowler-Barnett *et al.*, 2023). Protein-group confidence therefore should not be interpreted automatically as intact-proteoform confidence.

Top-down proteomics addresses a different analytical objective by interrogating intact molecular species and thereby preserving relationships among sequence variants, terminal processing, and co-occurring modifications before fragmentation (Roberts *et al.*, 2024). This molecular connectivity is a major advantage when the scientific question concerns proteoform composition. It does not eliminate uncertainty. Intact-protein separations, ionization efficiency, precursor complexity, fragmentation coverage, molecular size, and PTM localization can all restrict characterization.

Computational reconstruction is itself an important uncertainty layer. Feature-detection algorithms such as TopFD determine which isotope envelopes and charge-state patterns are assembled into candidate proteoform features (Basharat *et al.*, 2023). Deconvolution approaches such as FLASHDeconv improve processing speed and spectral interpretation but remain dependent on signal quality and assumptions about charge and isotope structure (Jeong *et al.*, 2020). Integrated environments such as MASH Explorer support extensive top-down data processing, yet their outputs remain conditioned by the experimental data, database search space, scoring rules, and downstream validation (Wu *et al.*, 2020).

Physical separation similarly changes the observable proteoform population. Size-based fractionation combined with FAIMS can reduce sample complexity and increase accessible proteoform depth (Takemori *et al.*, 2022), while internal compensation-voltage stepping alters the subset of ions transmitted during FAIMS analysis (Kaulich *et al.*, 2022). Coupling capillary zone electrophoresis with FAIMS adds another separation dimension that can improve resolution of complex intact-protein mixtures (Xu *et al.*, 2023). These gains demonstrate that proteoform coverage is partly a property of the complete workflow rather than solely of the biological sample.

Alternative acquisition architectures can preserve different forms of precursor-fragment information. Two-dimensional mass spectrometry has been applied to intact-proteoform analysis through precursor-fragment correlations that differ from conventional sequential precursor isolation (Polák *et al.*, 2023). Deep capillary-zone-electrophoresis-based top-down studies have demonstrated thousands of identifiable proteoforms in bacterial systems (McCool *et al.*, 2018), while earlier single-shot workflows established the feasibility of recovering hundreds of intact forms from *Escherichia coli* (Lubeckyj *et al.*, 2017). These studies illustrate technical progress without establishing equivalent depth in more heterogeneous mammalian tissues.

Bottom-up workflows have advanced simultaneously. DIA-NN, diaPASEF, and chromatogram-library approaches substantially increase peptide detection, sensitivity, and quantitative consistency in data-independent acquisition (Searle *et al.*, 2018; Demichev *et al.*, 2020; Meier *et al.*, 2020). Their analytical depth should not be equated with preservation of intact molecular connectivity. The distinction is therefore not between a “good” and a “bad” platform, but between different observables and the biological questions those observables can support.

The main uncertainty sources linking measured signal to functional interpretation are consolidated in **Table 3**.

Table 3. Analytical uncertainties that intervene between proteomic measurement and biological-function inference

Uncertainty source	Measurement level affected	Information that remains directly observable	Information requiring inference	Main interpretive consequence

Sample preparation and fractionation	Proteoform repertoire entering MS	Detected molecular species under the applied preparation	Whether missing forms were biologically absent	Apparent proteoform composition is workflow-conditioned
Digestion and peptide generation	Protein-to-peptide mapping	Local peptide sequence and many local PTMs	Long-range co-occurrence on the original molecule	Intact molecular connectivity may be lost
Protein grouping	Peptide-to-protein assignment	Peptide evidence	Which specific protein or proteoform generated shared peptides	Protein-level abundance can aggregate non-equivalent species
Intact-feature detection	Spectrum-to-proteoform feature	Charge/isotope envelopes and intact mass	Exact molecular identity where evidence is incomplete	Detected feature does not always equal fully characterized proteoform
PTM localization	Peptide or intact-fragment spectrum	Evidence for modification-associated mass shifts/fragments	Exact residue occupancy when localization is ambiguous	Functional claims about specific sites may be premature
Database annotation	Sequence/function interpretation	Existing curated or predicted annotations	Sample-specific biological state	Prior knowledge must not replace direct measurement
Separation and ion transmission	Accessible ion population	Species transmitted under selected conditions	Forms suppressed or unresolved elsewhere in the mixture	Coverage depends on instrument and separation design
Functional interpretation	Molecule, cell, or system	Abundance, modification, or structural-state measurement	Causal role in phenotype or pathway activity	Additional contextual or perturbational evidence is required

The analytical distinctions in **Table 3** lead directly to a broader evidentiary problem. A confident proteoform detection is stronger than an unresolved protein-group assignment for questions of molecular identity, yet even a fully characterized proteoform does not by itself establish biological function.

Figure 2 summarizes the evidentiary transitions from proteoform detection to functional interpretation.

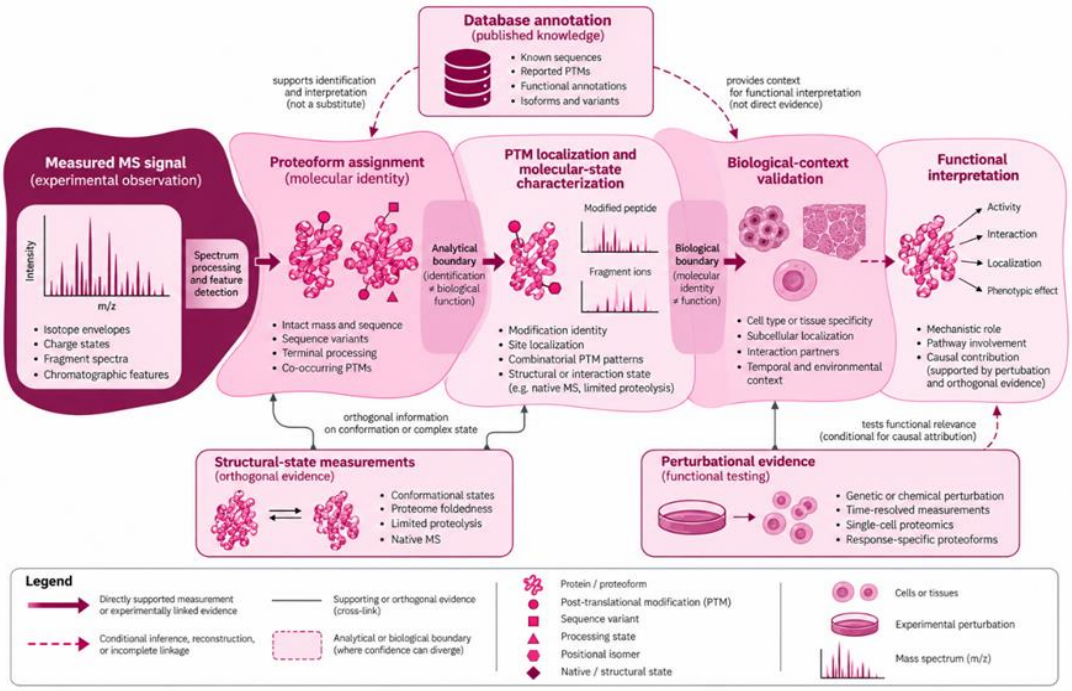


Figure 2. Evidentiary transitions from proteoform detection to functional interpretation

Critical Evaluation of Proteoform-to-Function Inference

The strongest limitation in proteoform interpretation appears after analytical identification has succeeded. Biological function is conditional on cellular state, localization, interaction partners,

substrate availability, regulatory networks, and temporal context. Consequently, detecting a proteoform—or even demonstrating that it changes in abundance—does not establish why that change occurred or whether the proteoform caused the associated phenotype.

Perturbational single-cell proteomics strengthens interpretation by relating molecular changes to defined cellular interventions while retaining cell-level heterogeneity (Brunner *et al.*, 2022). Such experiments move beyond descriptive covariance, although direct and indirect responses can remain difficult to separate. Structural-state measurements provide another orthogonal dimension. Limited-proteolysis mass spectrometry can report proteome-wide changes in accessibility associated with conformational or interaction-state changes, while proteome foldedness measurements reveal functional information not captured by abundance alone (Cox *et al.*, 2022; Malinovska *et al.*, 2023). Neither measurement is a universal proxy for activity, but both illustrate how molecular state can constrain interpretation.

Proteoform-to-function inference is therefore strongest when several independent dimensions converge: molecular identity, modification topology, structural or interaction state, biological context, and a perturbation capable of discriminating among competing explanations. The evidentiary value lies in convergence rather than in assigning a numerical confidence hierarchy. A discordant result should remain informative because it can identify which assumed link—from proteoform detection to functional consequence—requires re-examination.

Future Directions for Functionally Informative Proteomics

Future progress depends less on maximizing protein counts in isolation than on matching analytical resolution to the biological claim. Standardized proteoform representation and explicit uncertainty reporting should accompany improved intact-protein coverage so that unresolved modifications or sequence ambiguity remain visible rather than disappearing inside definitive labels. Single-cell, native-state, and orthogonal separation approaches will be particularly valuable when biological heterogeneity is itself the object of study.

Functional interpretation also requires closer integration between proteoform measurements and assays of molecular state. Thermal proteome profiling of the cell surface demonstrates how perturbation-sensitive stability measurements can connect abundance-independent protein behavior with drug or biological context (Kalxdorf *et al.*, 2021). Comparable integration with structural proteomics, interaction measurements, temporal perturbations, and spatial information could test whether a detected proteoform is merely associated with a state or participates mechanistically in producing it.

Benchmarking should accordingly assess more than identification depth. Useful benchmarks should examine whether competing proteoform assignments can be discriminated, whether PTM localization survives changes in acquisition and analysis, whether independently measured molecular properties support the same interpretation, and whether conclusions transfer across biological contexts. Functionally informative proteomics will advance most when uncertainty is treated as an analyzable property of the evidence rather than an inconvenience removed during data reduction.

Conclusion

Total protein abundance remains a powerful and often appropriate proteomic measurement, but its biological meaning depends on what molecular diversity is hidden beneath the protein label. Proteoforms can differ in sequence, processing, PTM topology, structural state, and biological context while contributing to the same aggregate abundance estimate. Bottom-up and top-down mass spectrometry preserve different portions of this information, and neither analytical architecture independently converts molecular detection into functional knowledge. The most defensible functional interpretations arise when proteoform identity and modification localization are integrated with contextual, structural, and perturbational evidence. Proteomics therefore becomes more functionally informative not by abandoning protein abundance, but by recognizing when abundance is insufficient for the biological question being asked.

Acknowledgments: None

Conflict of interest: None

Financial support: None

Ethics statement: None

References

Aebersold, R., Agar, J. N., Amster, I. J., Baker, M. S., Bertozzi, C. R., Boja, E. S., Costello, C. E., Cravatt, B. F., Fenselau, C., Garcia, B. A., et al. (2018). How many human proteoforms are there? *Nature Chemical Biology*, *14*(3), 206-214. doi:10.1038/nchembio.2576

Basharat, A. R., Zang, Y., Sun, L., & Liu, X. (2023). TopFD: A proteoform feature detection tool for top-down proteomics. *Analytical Chemistry*, *95*(21), 8189-8196. doi:10.1021/acs.analchem.2c05244

Berthias, F., Thurman, H. A., Wijegunawardena, G., Wu, H., Shvartsburg, A. A., & Jensen, O. N. (2023). Top-down ion mobility separations of isomeric proteoforms. *Analytical Chemistry*, *95*(2), 784-791. doi:10.1021/acs.analchem.2c02948

Bludau, I., Frank, M., Dörig, C., Cai, Y., Heusel, M., Rosenberger, G., Picotti, P., Collins, B. C., Röst, H., & Aebersold, R. (2021). Systematic detection of functional proteoform groups from bottom-up proteomic datasets. *Nature Communications*, *12*(1), 3810. doi:10.1038/s41467-021-24030-x

Bowler-Barnett, E. H., Fan, J., Luo, J., Magrane, M., Martin, M. J., Orchard, S., & UniProt Consortium. (2023). UniProt and mass spectrometry-based proteomics—a 2-way working relationship. *Molecular & Cellular Proteomics*, *22*(8), 100591. doi:10.1016/j.mcpro.2023.100591

Brunner, A.-D., Thielert, M., Vasilopoulou, C., Ammar, C., Coscia, F., Mund, A., Hoerning, O. B., Bache, N.,

Apalategui, A., Lubeck, M., et al. (2022). Ultra-high sensitivity mass spectrometry quantifies single-cell proteome changes upon perturbation. *Molecular Systems Biology, 18*(3), e10798. doi:10.15252/msb.202110798

Camacho, O. M., Ramsbottom, K. A., Collins, A., & Jones, A. R. (2023). Assessing multiple evidence streams to decide on confidence for identification of post-translational modifications, within and across data sets. *Journal of Proteome Research, 22*(6), 1828-1842. doi:10.1021/acs.jproteome.2c00823

Chen, W., Ding, Z., Zang, Y., & Liu, X. (2023). Characterization of proteoform post-translational modifications by top-down and bottom-up mass spectrometry in conjunction with annotations. *Journal of Proteome Research, 22*(10), 3178-3189. doi:10.1021/acs.jproteome.3c00207

Cox, D., Ang, C.-S., Nillegoda, N. B., Reid, G. E., & Hatters, D. M. (2022). Hidden information on protein function in censuses of proteome foldedness. *Nature Communications, 13*(1), 1992. doi:10.1038/s41467-022-29661-2

Demeulemeester, N., Gébelin, M., Caldi Gomes, L., Lingor, P., Carapito, C., Martens, L., & Clement, L. (2024). msqrob2PTM: Differential abundance and differential usage analysis of MS-based proteomics data at the posttranslational modification and peptidoform level. *Molecular & Cellular Proteomics, 23*(2), 100708. doi:10.1016/j.mcpro.2023.100708

Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S., & Ralser, M. (2020). DIA-NN: Neural networks and interference correction enable deep proteome coverage in high throughput. *Nature Methods, 17*(1), 41-44. doi:10.1038/s41592-019-0638-x

Geiszler, D. J., Polasky, D. A., Yu, F., & Nesvizhskii, A. I. (2023). Detecting diagnostic features in MS/MS spectra of post-translationally modified peptides. *Nature Communications, 14*(1), 4132. doi:10.1038/s41467-023-39828-0

Habeck, T., Brown, K. A., Des Soye, B., Lantz, C., Zhou, M., Alam, N., Hossain, M. A., Jung, W., Keener, J. E., Volny, M., et al. (2024). Top-down mass spectrometry of native proteoforms and their complexes: A community study. *Nature Methods, 21*(12), 2388-2396. doi:10.1038/s41592-024-02279-6

Jeong, K., Kim, J., Gaikwad, M., Hidayah, S. N., Heikaus, L., Schlüter, H., & Kohlbacher, O. (2020). FLASHDeconv: Ultrafast, high-quality feature deconvolution for top-down proteomics. *Cell Systems, 10*(2), 213-218.e6. doi:10.1016/j.cels.2020.01.003

Kalxdorf, M., Günthner, I., Becher, I., Kurzawa, N., Knecht, S., Savitski, M. M., Eberl, H. C., & Bantscheff, M. (2021). Cell surface thermal proteome profiling tracks perturbations and drug targets on the plasma membrane. *Nature Methods, 18*(1), 84-91. doi:10.1038/s41592-020-01022-1

Kaulich, P. T., Cassidy, L., Winkels, K., & Tholey, A. (2022). Improved identification of proteoforms in top-down proteomics using FAIMS with internal CV stepping. *Analytical Chemistry, 94*(8), 3600-3607. doi:10.1021/acs.analchem.1c05123

Kaulich, P. T., Jeong, K., Kohlbacher, O., & Tholey, A. (2024). Influence of different sample preparation approaches on proteoform identification by top-down proteomics. *Nature Methods, 21*(12), 2397-2407. doi:10.1038/s41592-024-02481-6

LeDuc, R. D., Deutsch, E. W., Binz, P.-A., Fellers, R. T., Cesnik, A. J., Klein, J. A., Van Den Bossche, T., Gabriels, R., Yalavarthi, A., Perez-Riverol, Y., et al. (2022). Proteomics Standards Initiative's ProForma 2.0: Unifying the encoding of proteoforms and peptidoforms. *Journal of Proteome Research, 21*(4), 1189-1195. doi:10.1021/acs.jproteome.1c00771

Lubeckyj, R. A., McCool, E. N., Shen, X., Kou, Q., Liu, X., & Sun, L. (2017). Single-shot top-down proteomics with capillary zone electrophoresis-electrospray ionization-tandem mass spectrometry for identification of nearly 600 Escherichia coli proteoforms. *Analytical Chemistry, 89*(22), 12059-12067. doi:10.1021/acs.analchem.7b02532

Malinovska, L., Cappelletti, V., Kohler, D., Piazza, I., Tsai, T.-H., Pepelnjak, M., Stalder, P., Dörig, C., Sesterhenn, F., Elsässer, F., et al. (2023). Proteome-wide structural changes measured with limited proteolysis-mass spectrometry: An advanced protocol for high-throughput applications. *Nature Protocols, 18*(3), 659-682. doi:10.1038/s41596-022-00771-x

McCool, E. N., Lubeckyj, R. A., Shen, X., Chen, D., Kou, Q., Liu, X., & Sun, L. (2018). Deep top-down proteomics using capillary zone electrophoresis-tandem mass spectrometry: Identification of 5700 proteoforms from the Escherichia coli proteome. *Analytical Chemistry, 90*(9), 5529-5533. doi:10.1021/acs.analchem.8b00693

Meier, F., Brunner, A.-D., Frank, M., Ha, A., Bludau, I., Voytik, E., Kaspar-Schoenefeld, S., Lubeck, M., Raether, O., Bache, N., et al. (2020). diaPASEF: Parallel accumulation–serial fragmentation combined with data-independent acquisition. *Nature Methods, 17*(12), 1229-1236. doi:10.1038/s41592-020-00998-0

Ochoa, D., Jarnuczak, A. F., Viéitez, C., Gehre, M., Soucheray, M., Mateus, A., Kleefeldt, A. A., Hill, A., Garcia-Alonso, L., Stein, F., et al. (2020). The functional landscape of the human phosphoproteome. *Nature Biotechnology, 38*(3), 365-373. doi:10.1038/s41587-019-0344-3

Pavek, J. G., Frey, B. L., Frost, D. C., Gu, T.-J., Li, L., & Smith, L. M. (2023). Cysteine counting via isotopic chemical labeling for intact mass proteoform identifications in tissue. *Analytical Chemistry, 95*(41), 15245-15253. doi:10.1021/acs.analchem.3c02473

Po, A., & Eyers, C. E. (2023). Top-down proteomics and the challenges of true proteoform characterization. *Journal of Proteome Research, 22*(12), 3663-3675. doi:10.1021/acs.jproteome.3c00416

Polák, M., Palasser, M., Kádek, A., Kavan, D., Wootton, C. A., Delsuc, M.-A., Breuker, K., Novák, P., & van Agthoven, M. A. (2023). Top-down proteoform analysis by 2D MS with quadrupolar detection. *Analytical Chemistry, 95*(44), 16123-16130. doi:10.1021/acs.analchem.3c02225

Polasky, D. A., Geiszler, D. J., Yu, F., Li, K., Teo, G. C., & Nesvizhskii, A. I. (2023). MSFragger-Labile: A flexible method to improve labile PTM analysis in proteomics. *Molecular & Cellular Proteomics, 22*(5), 100538.

doi:10.1016/j.mcpro.2023.100538

Roberts, D. S., Loo, J. A., Tsybin, Y. O., Liu, X., Wu, S., Chamot-Rooke, J., Agar, J. N., Paša-Tolić, L., Smith, L. M., & Ge, Y. (2024). Top-down proteomics. *Nature Reviews Methods Primers, 4*(1), 38. doi:10.1038/s43586-024-00318-2

Searle, B. C., Pino, L. K., Egertson, J. D., Ting, Y. S., Lawrence, R. T., MacLean, B. X., Villén, J., & MacCoss, M. J. (2018). Chromatogram libraries improve peptide detection and quantification by data independent acquisition mass spectrometry. *Nature Communications, 9*(1), 5128. doi:10.1038/s41467-018-07454-w

Smith, L. M., Agar, J. N., Chamot-Rooke, J., Danis, P. O., Ge, Y., Loo, J. A., Paša-Tolić, L., Tsybin, Y. O., Kelleher, N. L., & Consortium for Top-Down Proteomics. (2021). The Human Proteoform Project: Defining the human proteome. *Science Advances, 7*(46), eabk0734. doi:10.1126/sciadv.abk0734

Smith, L. M., & Kelleher, N. L. (2018). Proteoforms as the next proteomics currency. *Science, 359*(6380), 1106-1107. doi:10.1126/science.aat1884

Smith, L. M., Thomas, P. M., Shortreed, M. R., Schaffer, L. V., Fellers, R. T., LeDuc, R. D., Tucholski, T., Ge, Y., Agar, J. N., Anderson, L. C., et al. (2019). A five-level classification system for proteoform identifications. *Nature Methods, 16*(10), 939-940. doi:10.1038/s41592-019-0573-x

Su, P., Hollas, M. A. R., Butun, F. A., Kanchustambham, V. L., Rubakhin, S., Ramani, N., Greer, J. B., Early, B. P., Fellers, R. T., Caldwell, M. A., et al. (2024). Single cell analysis of proteoforms. *Journal of Proteome Research, 23*(6), 1883-1893. doi:10.1021/acs.jproteome.4c00075

Takemori, A., Kaulich, P. T., Cassidy, L., Takemori, N., & Tholey, A. (2022). Size-based proteome fractionation through polyacrylamide gel electrophoresis combined with LC-FAIMS-MS for in-depth top-down proteomics. *Analytical Chemistry, 94*(37), 12815-12821. doi:10.1021/acs.analchem.2c02777

The, M., Samaras, P., Kuster, B., & Wilhelm, M. (2022). Reanalysis of ProteomicsDB using an accurate, sensitive, and scalable false discovery rate estimation approach for protein groups. *Molecular & Cellular Proteomics, 21*(12), 100437. doi:10.1016/j.mcpro.2022.100437

van der Burgt, Y., & Wuhrer, M. (2023). The role of clinical glyco(proteo)mics in precision medicine. *Molecular & Cellular Proteomics, 22*(6), 100565. doi:10.1016/j.mcpro.2023.100565

Wu, Z., Roberts, D. S., Melby, J. A., Wenger, K., Wetzel, M., Gu, Y., Ramanathan, S. G., Bayne, E. F., Liu, X., Sun, R., et al. (2020). MASH Explorer: A universal software environment for top-down proteomics. *Journal of Proteome Research, 19*(9), 3867-3876. doi:10.1021/acs.jproteome.0c00469

Xu, T., Wang, Q., Wang, Q., & Sun, L. (2023). Coupling high-field asymmetric waveform ion mobility spectrometry with capillary zone electrophoresis-tandem mass spectrometry for top-down proteomics. *Analytical Chemistry, 95*(25), 9497-9504. doi:10.1021/acs.analchem.3c00551