Journal of Medicinal Chemistry

pubs.acs.org/jmc

Open Access

This article is licensed under CC-BY 4.0

Perspective

Mass Spectrometry Proteomics: A Key to Faster Drug Discovery

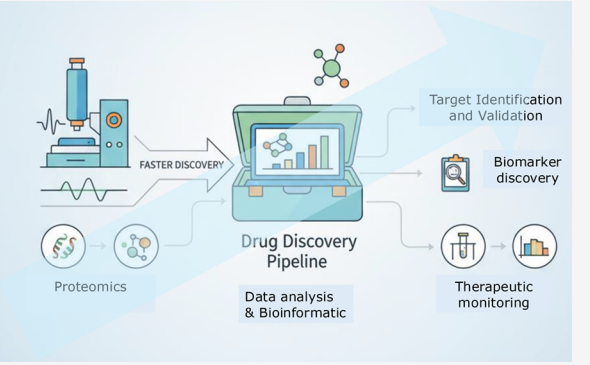
Lorenzo Tagliazucchi* and Maria Paola Costi*

Cite This: https://doi.org/10.1021/acs.jmedchem.5c01986

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Mass spectrometry (MS)-based proteomics is a disruptive platform in drug discovery that offers an exhaustive view of the proteome's complexity. Focusing on bottom-up MS proteomics, this technology enables high-throughput analysis of protein expression, interactions, and modifications, far surpassing the capabilities of traditional single-protein methods. The MS proteomics toolbox is essential in both early- and late-stage development of new drugs. The techniques discussed here, such as unlabeled and labeled proteomics and chemoproteomic approaches (e.g., thermal proteome profiling and photoaffinity labeling), facilitate target binding site exploration and the identification of putative off-targets. By accelerating the identification of new druggable proteins and supporting early biomarker discovery, MS proteomics significantly accelerates the preclinical-to-clinical transition. Ongoing progress in data acquisition, new computational tools, and artificial intelligence further enhances the high-throughput properties of these approaches, marking a significant step toward personalized medicine.



Downloaded via 101.56.107.211 on January 5, 2026 at 19:33:28 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

1. INTRODUCTION

Drug discovery remains a time-consuming and failure-prone process, which often requires over a decade and a consistent financial investment to bring a new therapeutic agent to the market.[1] As a result, there is a growing need for more robust and integrative technologies to update drug discovery workflows and improve success rates.[2,3] Mass spectrometry (MS)-based proteomics has emerged as an innovative approach in this context, as it offers a complete qualitative and quantitative insight into protein expression, post-translational modifications, and their interaction with other proteins or molecules.[4,5] Unlike genomic or transcriptomic tools, which provide an indirect description of cellular function, MS proteomics directly measures the functional biomolecules that regulate nearly all biological processes.[6] Despite its innovation potential, MS proteomics remains underexploited in medicinal chemistry, partly due to its technological complexity, data interpretation challenges, and a historical disconnection between proteomics and the drug design process.[7] Thus, this review aims to provide a structured overview of how MS proteomics can be exploited across the drug discovery pipeline from early target identification to translational applications in precision medicine.

The journey begins with target discovery, where comparative MS-based proteomics between healthy and diseased states can reveal differentially expressed proteins and dysregulated pathways. Untargeted and targeted chemoproteomics strategies, such as thermal proteome profiling (TPP) and other chemical proteomics techniques, allow the identification of druggable proteins, even in complex samples[8,9] (Figure 1, from left to right, drug discovery process and MS technologies progression). These techniques represent a critical starting point to identify the modifiable targets. Next, target engagement studies determine whether and how drug candidates interact with their targets in biological systems. Activity-based protein profiling (ABPP), photoaffinity labeling (PAL), limited proteolysis coupled to MS (LiP-MS), and cross-linking MS (XL-MS) are used to evaluate binding specificity, stoichiometry, and cellular accessibility.[5] When combined with intact-cell approaches, real-time, in situ feedback on compound efficacy and selectivity information can be achieved. Once validated, the targets are subjected to biochemical mechanistic investigation. MS proteomic readouts can reveal changes in protein expression levels, protein complex formation, and pathway activation, all of which contribute to the drug's mechanism of action (MoA).[10] Covalent chemical proteomic methods, such as hydrogen-deuterium exchange (HDX-MS), help to map ligand-binding sites and identify conformational changes, providing medicinal chemists with data to guide structure-activity relationship (SAR) optimization.[11] Fragment-based drug discovery (FBDD) can

Received: July 16, 2025
Revised: November 28, 2025
Accepted: December 8, 2025

ACS Publications

© XXXX The Authors. Published by American Chemical Society

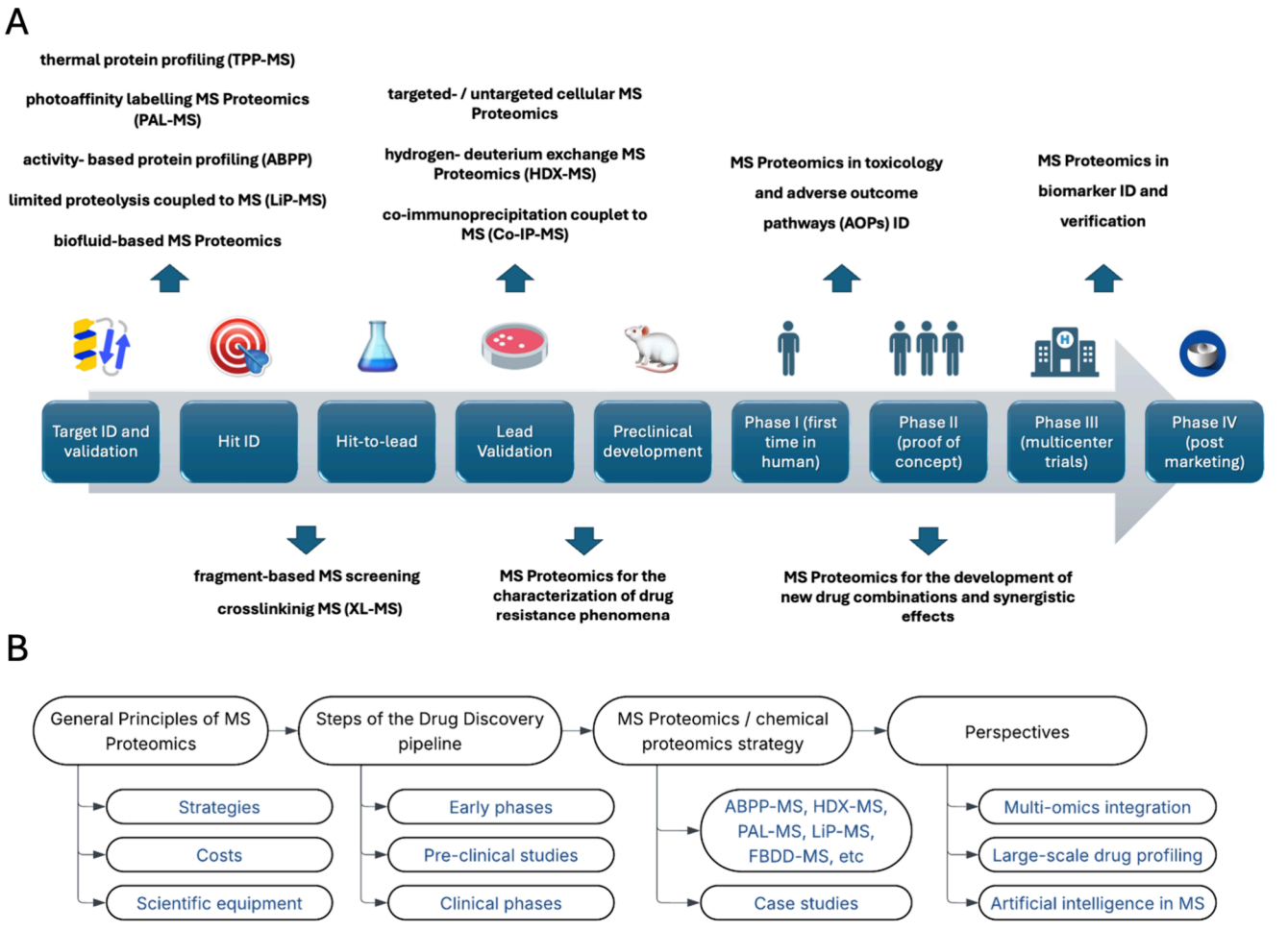


Figure 1. (A) Applications of MS proteomic and chemoproteomic techniques at different stages of the drug discovery process, from the target identification to clinical and post-marketing surveillance. (B) Workflow of the review content.

also be pursued by MS chemoproteomics, where weak but specific fragment-protein interactions are characterized and refined toward high-affinity leads.[12] MS proteomics has the possibility to play a central role in understanding off-target effects and resistance mechanisms, which are key drivers of drug failure.[13] Proteome-wide analysis can detect alterations in signaling pathways and protein abundance associated with adverse outcomes or compensatory responses.[14] Techniques such as co-immunoprecipitation followed by MS (Co-IP-MS) allow for the identification of toxicity markers or resistance biomarkers early in development.[15] This strategy supports the design of safer and more effective drugs and facilitates combination therapy strategies to prevent drug resistance onset, e.g., the synthetic lethality paradigm.[16] In translational and clinical research, MS proteomics bridges preclinical insights with patient data, contributing to precision medicine.[17,18] Clinical proteomics enables patient stratification by identifying disease-specific signatures, predictive biomarkers, and pharmacodynamic markers from biopsies, plasma, or other biofluids. Data-independent acquisition (DIA) methods and isobaric labeling (e.g., TMT or iTRAQ) have improved the reproducibility and throughput of proteomics in complex clinical matrices.[19] These advances, coupled with the availability of large clinical metadata sets, support early detection of therapy resistance and real-time monitoring of therapeutic response.[20] Recent innovations in instrumentation (e.g., hybrid MS with different fragmentations) and computational platforms (e.g., Skyline, Proteome Discoverer) have improved data depth and quality.[21] Artificial intelligence (AI) now supports pattern recognition, predictive modeling, and automated hypothesis generation, which are crucial for navigating complex MS proteomic data.[22]

This perspective provides an integrated overview of the role of bottom-up MS proteomics strategies as integrated in the modern drug discovery pipeline, illustrated in Figure 1, as applied to each stage of the research. We also address recent advancements in MS instrumentation and software tools, which have improved the accuracy and reproducibility of MS proteomics analyses. Finally, we highlight emerging trends and future directions in MS proteomics and chemical proteomics, including their integration with AI and machine learning algorithms, which are expected to further accelerate drug discovery efforts and enhance data interpretation. By bridging the gap between molecular biology and medicinal chemistry, MS proteomics continues to advance our understanding of complex diseases and put the bases for next-generation therapeutics. All these strategies will be discussed in detail with examples throughout the present perspective.

2. PRINCIPLES OF MS PROTEOMICS ANALYSIS

Overall, MS proteomics pipelines can be broadly classified into bottom-up and top-down approaches. In bottom-up MS proteomics, proteins are enzymatically digested into peptides, which are then analyzed by MS to reconstruct protein identities

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

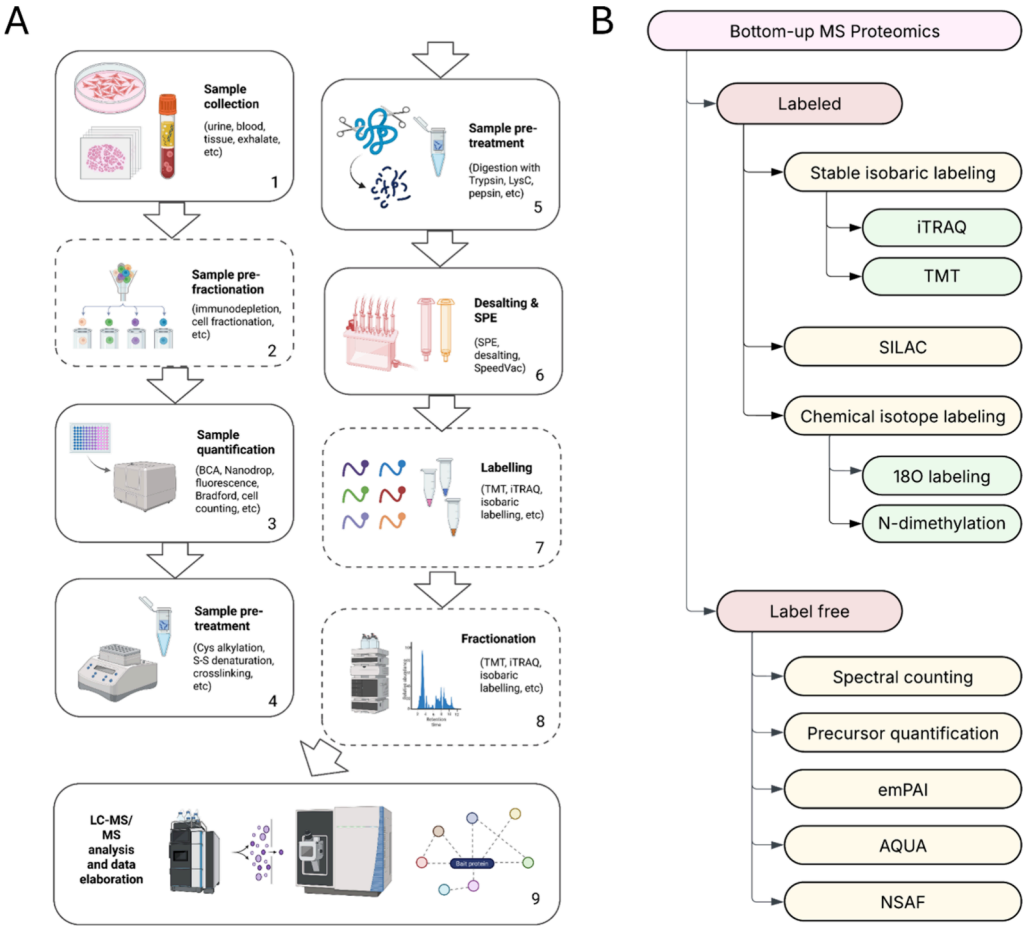


Figure 2. (A) General sample preparation pipeline for a bottom-up proteomics experiment. Steps out of brackets are usually compulsory for any kind of experimental setup, whereas steps in brackets represent optional steps (i.e., for labeling experiments and fractionation experiments). General workflow for bottom-up MS proteomics includes quantification of protein content in each sample specimen to prepare equal volumes of initial proteome, sample preprocessing and digestion into peptides, peptides desalting from organic reactants and salts, and the proper LC−MS/MS analysis. Dashed boxes represent optional steps. (B) General workflows and data acquisition methods for bottom-up MS proteomics experiments. Acronyms are defined as follows: Normalized Spectral Abundance Factor (NSAF), emPAI, AQUA (Absolute QUAntification), TMT Tandem Mass Tag (TMT), isobaric tags for relative and absolute quantitation (iTRAQ), stable isotope-labeled amino acids (SILAC).[38−46]

and abundances with respect to a control sample.[23] On the other hand, MS top-down proteomics is used to analyze intact proteins without prior digestion, preserving information about isoforms and modifications but requiring high-resolution instruments due to the complexity of intact protein spectra.[24] This perspective will discuss only bottom-up MS proteomics applications. Indeed, they are the most often used applications to inform about protein/peptide expression and the most exploited for chemical proteomics purposes these days.

Depending on the purpose of the study, an MS proteomic approach can be defined as either targeted or untargeted. Targeted MS proteomics focuses on the quantitation of specific proteins or peptides of interest.[25] This approach is often implemented through techniques such as selected reaction monitoring (SRM), multiple reaction monitoring (MRM), and parallel reaction monitoring (PRM).[25,26] The workflow typically begins with the selection of unique peptide sequences, termed 'proteotypic peptides', which represent the target proteins.[26,29] Following digestion of the protein, these peptides are analyzed through MS. Precursor ions are fragmented, the selected dissociation transitions are monitored according to the peptides of interest, and quantitation is based on peak areas.[27,30] One of the significant advantages of targeted proteomics is its high sensitivity and reproducibility. Also, by employing stable isotope-labeled internal standards, researchers can achieve absolute quantification of target proteins across various samples.[28] In contrast, untargeted MS proteomics aims to provide a global overview of the entire proteome without prior knowledge of which proteins may be present or their abundance; therefore, it is an unsupervised approach.[29] Untargeted proteomics is particularly useful in exploratory or preliminary studies aimed at biomarker discovery, understanding disease mechanisms, or the MoA and off-target activity of investigational drugs.[30,31] The ability to identify novel biomarkers can lead to significant advancements in personalized medicine by informing treatment decisions based on individual patient profiles.[30,32] The volume of data generated necessitates robust and multiple statistical methods to ensure that identified proteins are confirmed in the biological samples.[33] Additionally, issues such as ion suppression and variability in peptide ionization can complicate quantitative analyses.[34]

In the field of untargeted proteomics, two prominent methodologies are "*whole sample*" (e.g., whole plasma, whole cell, and whole tissue) proteomics, also known as "total" proteomics, and fractionated experiments[35] (Figure 2A). One of the primary advantages of whole sample proteomics is its ability

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

to analyze thousands of proteins simultaneously, making it suitable for large-scale studies.[36] Whole sample MS proteomics is effective for handling complex mixtures that are difficult to fractionate or separate according to physical or biological characteristics in a reproducible manner without sample loss, such as cell lysates or tissue extracts.[37] On the other hand, a fractionated MS proteomics experiment includes, at different stages of the preprocessing, a separation of the samples into different subfractions that can be analyzed one by one by the MS to improve instrumental sensitivity and thus protein coverage.[37]

Examples of fractionation pipelines include serum or plasma immunodepletion to remove the most abundant proteins like albumins or immunoglobulins, which can make up to 80% of the protein content, through the use of immunoaffinity chromatography or cell fractionation of the subcellular compartments (e.g., cytosol from cell nucleus from membrane fraction).[38] Otherwise, fractionation can occur at the bottom of the digestion protocol, directly on the peptide mixture.[39,40] In this case, it is usually performed with anionic exchange chromatography (AEX), which allows for peptide separation according to their isoelectric point, resulting in an off-line combined LCxLC for pH and lipophilicity.[41] Phosphoproteomics is another example of fractionated MS proteomics in which phosphopeptides are purposely isolated and concentrated through ion metal affinity chromatography (IMAC) or TiO_2 beads, which retain specifically phosphorylated residues and are often combined further with AEX fractionation.[42,43] Generally speaking, the authors suggest that, unless specific findings are expected from other previous observations, phopshoproteomics is performed only after total proteomics on the same samples. The preparation of samples for bottom-up MS proteomics, either for labeled or label-free pipelines, is a crucial step that ensures the accurate identification and quantification of proteins. Given the diverse objectives of MS proteomic experiments, no universal protocol exists, but a general workflow can be outlined to standardize sample processing, as illustrated in Figure 2A.

2.1. Quantification, Peptide Labeling, and Data Acquisition Modes. Quantification approaches are broadly divided into two main categories: label-free quantification (LFQ) and label-based quantification.[44,45] Both approaches offer distinct advantages and challenges, depending on the experimental goals and sample complexity.[45] LFQ relies on comparing the intensity of MS signals between different samples to estimate peptide, and thus protein, abundances. This method is straightforward and cost-effective, as it does not require chemical or metabolic labeling of the samples.[46] There are several types of LFQ pipelines and include spectral counting (SC), peptide ion intensity (PII), normalized spectral abundance factor (NSAF), emPAI, absolute quantification (AQUA).[47] SC counts the number of MS/MS spectra matched to a particular peptide or protein.[48] Higher spectral counts are assumed to correlate with a higher protein abundance. Although simple, this method is semiquantitative and may not capture small changes in protein expression.[48] PII, also referred to as precursor quantification, relies on extracted ion chromatograms (XIC) and measures the intensity of peptide ions (based on their MS^1 signal) across multiple samples.[45] The signal intensity of a peptide is proportional to its concentration, allowing for quantitative comparisons, either on the chromatographic peak areas (e.g., PEAKS from Bioinformatics Solutions Inc. or Progenesis QIP from Waters Corp) or to their relative intensity signals (e.g., Proteome Discoverer).[49] Therefore, this is the most common LFQ method used. The NSAF method normalizes spectral counts by protein length and the total spectral counts in the sample. It reduces bias toward larger proteins and allows more accurate comparison of relative protein abundance across samples.[50] emPAI, on the other hand, estimates protein abundance based on the ratio of the observed peptides to the theoretically observable peptides for a given protein. This approach accounts for protein size and the complexity of digestion, also providing a semiquantitative estimate of protein concentration.[51] AQUA method relies on synthetic stable isotope-labeled peptides spiked into the sample as internal standards. The MS signal of the endogenous peptide is compared with that of the labeled AQUA peptide, providing an absolute quantification of protein abundance. While not strictly "label-free" when isotopic standards are used, it is often grouped with LFQ approaches due to its relative quantification principle before normalization.[52] The absence of expensive reagents such as isotopic labels or specialized equipment makes LFQ a more affordable option for large-scale proteomics experiments, like on large cohorts of samples from patients (>100 patients).[53] However, the LFQ is more susceptible to experimental variations between runs, such as differences in sample preparation, chromatographic retention times, and ionization efficiency. Such variability can introduce noise into the quantification process.[49]

Labeled proteomics pipeline involves the introduction of isotopic or chemical labels into the sample, which enables the relative or absolute quantification of proteins across different experimental conditions. Labeling can be done either at the protein level or at the peptide level, depending on the labeling strategy.[54] The three most common labeling strategies are isobaric labeling (e.g., TMT, i.e., Tandem Mass Tag (ThermoFisher) or iTRAQ, i.e., isobaric tags for relative and absolute quantitation (Waters)), stable isotope labeling by amino acids in cell culture (SILAC), and labeling with chemical isotopes.[55,56] TMT (very similar to the iTRAQ kit) can label up to 32 different samples (32-plex) in a single experiment, allowing for multiplexed quantification. Each tag consists of an MS/MS reporter ion, a mass normalizer, and a reactive group that binds to primary amines on peptides. During MS/MS, the tags release reporter ions of different m/z values, which are used for relative quantification of the peptides.[57] The ability to multiplex samples increases throughput and allows for the simultaneous quantification of multiple conditions in a single run. This minimizes run-to-run variability and batch effects.[57]

SILAC, which stands for "stable isotope-labeled amino acids," is a metabolic labeling technique that introduces stable isotope-labeled amino acids (e.g., ^{13}C or ^{15}N) directly into the cell culture medium.[56] Cells incorporate the labeled amino acids into their proteins during growth, creating a direct comparison between "light" (unlabeled) and "heavy" (labeled) peptides in the MS spectra.[58] It is suitable to study dynamic processes, such as signaling pathways, protein turnover, and interactions in cellular studies, because it allows the incorporation of time points into the experimental design.[56] On the other hand, SILAC is only applicable to cells that can be cultured in vitro, limiting its use in clinical or environmental samples where metabolic labeling is not feasible.[56] Another labeling option is the chemical isotope labeling, which involves the use of stable isotope labels that are chemically introduced to peptides after protein digestion.[59] Common strategies include dimethyl labeling, that introduces dimethylation at the peptide N-termini, lysine residues and generates a mass shift between labeled and unlabeled peptides, or the ^{18}O Labeling, which works the same

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

Table 1. Pros and Cons of Label-Free and Labeled MS Proteomics

feature	label-free proteomics	labeled proteomics
sample preparation	simple, i.e., no labeling required	complex as it involves labeling steps
cost estimate	lower for sample preparation (no labeling needed); MS experimental might impact more due to not multiplexing and higher replicate required	higher due to labeling reagents for sample preparation. Multiplexing helps contain costs associated with MS runs
throughput	high, simultaneous analysis possible	moderate (limited to ~20 samples per experiment)
grouping and replicates	minimum three replicates per group are suggested due to the intrinsic variability associated with label-free workflows	single replicates (as low as 2 samples/group) are enough thanks to high reproducibility and variability
reproducibility	lower, more variability (peptide abundance is quantified on precursor intensity/area)	higher, reduced experimental variability (peptide abundance is quantified on reported ion intensity)
applicability	broad, all biological samples	limited, some techniques restricted to sample type, especially for the minimum amount of material required
flexibility in study design	high, new samples can be added any time before LC-MS/MS analysis	low, fixed once the experimental design has started

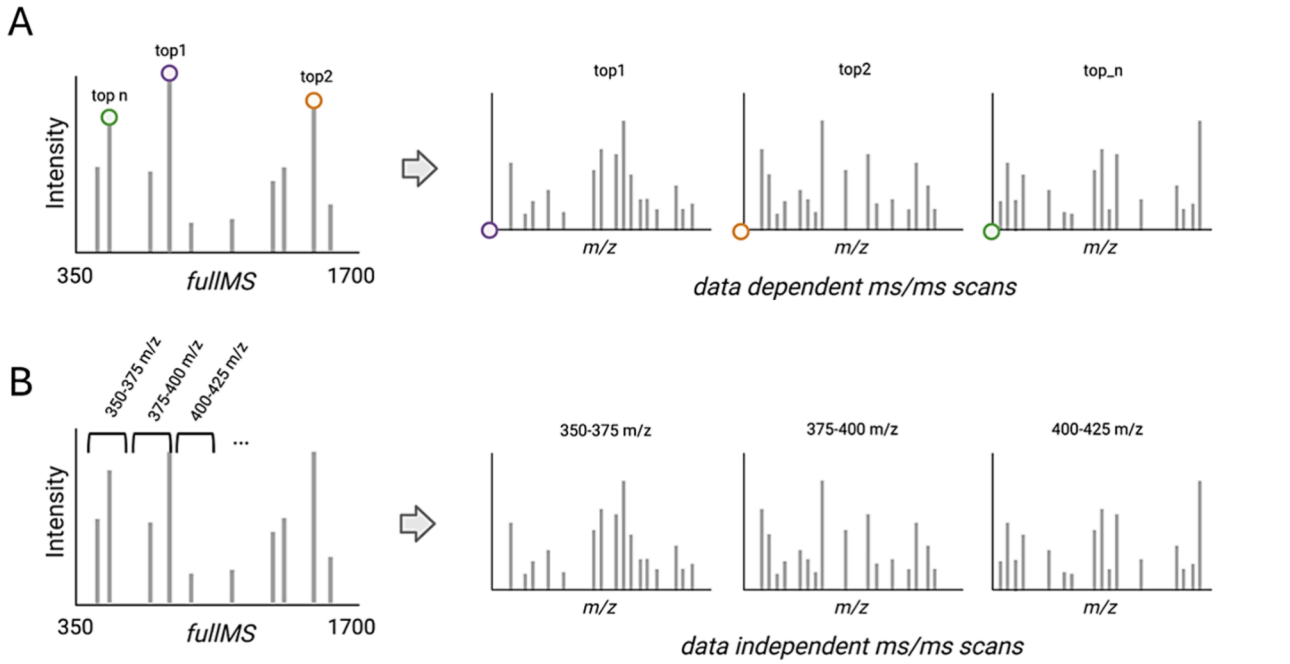


Figure 3. Data-dependent acquisition (DDA) (A) and data-independent acquisition (DIA) (B) schematic workflows. In DDA mode, the top n precursor ions are selected according to their intensities in MS1 scans and fragmented in MS/MS experiments every scan cycle. In DIA mode, the full-MS window is divided into smaller continuous m/z windows (usually 15−40 m/z), and all their precursors are fragmented in MS/MS experiments. The overlapping fragment ions for each small window are then deconvoluted by appropriate software.

but at the C-terminal carboxyl group.[60,61] Overall, label-free experiments are preferred in explorative experiments when large groups of samples or large cohorts of patients are being analyzed. This is due to the lower costs associated with sample processing and to the fact that labeled proteomics limits sample size to the one required in vendors' kit requirements, which is generally a few microliters. On the other hand, labeled techniques in a standard experiment result in higher accuracy and reproducibility, with as low as two replicates per group.[54] Also, labeling allows multiplexing, resulting in lower costs associated with fewer MS runs needed. Pros and cons of label-free and labeled MS proteomics experiments are reported in Table 1, while a schematic representation of the possible quantification experiments in bottom-up proteomics is reported in Figure 2B.

Additionally, data-dependent acquisition (DDA) and DIA strategies influence how mass spectra are recorded, balancing the depth of coverage with reproducibility and quantification accuracy. DDA begins with a full-MS scan to detect all ionized peptides in a sample.[62] From this scan, a limited number of the most intense precursor ions (typically 8−25) are selected in real time for fragmentation, balancing acquisition speed and resolution.[63] The resulting MS/MS spectra are then used for peptide identification.[62] While DDA provides high-quality spectra for abundant peptides, it can miss lower-abundance proteins in complex mixtures due to its selective nature. DIA, on the other hand, fragments all precursor ions within predefined m/z windows, allowing for broader and more consistent proteome coverage.[64] Instead of isolating individual precursors, DIA divides the full m/z range (e.g., 400−1200 m/z) into overlapping windows (e.g., 25 Da each), fragmenting all ions within each window simultaneously.[65] This results in complex, partially overlapping spectra that require sophisticated computational tools and complete spectral libraries to deconvolute and identify peptides accurately.[66] A schematic explanation of how DDA and DIA work is illustrated in Figure 3.

2.2. Current Instrumentations for Bottom-Up MS Proteomics and Operational Costs. All the most modern and cutting-edge models of MS instruments provide high sensibility and, when coupled to nano-LC systems, allow the investigators to perform medium- and high-throughput

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

Table 2. Comparison between Low-, Middle-, and High-Throughput MS Proteomics

feature	low throughput	middle throughput	high throughput
primary objective	deep proteome coverage and discovery of novel proteins/protein isoforms/PTMs	balanced coverage and scalability for comparative studies	rapid and large-scale profiling with high reproducibility
typical use cases	target discovery, pathway mapping, PTM analysis	drug response profiling, target validation, and mechanistic studies	biomarker discovery, patient stratification, clinical screening
sample capacity	<50 samples per batch	50−500 samples per batch	thousands of samples per batch
acquisition mode	usually DDA, sometimes DIA	DIA or multiplexed DDA (TMT/iTRAQ)	DIA or scanning SWATH with shorter gradients
LC setup	nano-LC, long gradients (90 - 240 min)	micro/nano-LC, moderate gradients (30 - 60 min)	high-flow or Evosep LC, short gradients (5 - 15 min)
flow rate	200−300 nL/min (nanoflow)	1−10 μL/min (microflow)	high-flow or Evosep
sample preparation	extensive manual processing; multiple fractionation steps (e.g., SCX, HILIC, high-pH RP)	streamlined digestion and cleanup; optional fractionation	automated digestion and cleanup using liquid-handling robots
throughput (samples/day)	<10	10−50	>100
reproducibility	high within-run precision, lower between runs	moderate to high reproducibility	very high reproducibility (DIA advantage)
limitations	time-consuming, low sample capacity, expensive	intermediate depth, requires optimized LC-MS setup	lower depth, risk of missing low-abundance proteins

experimental sessions with minimal operator occupancy and manual intervention required (a comparison between low-, middle-, and high-throughput MS proteomics is provided in Table 2).

High-resolution mass spectrometry (HRMS) analyzers play a key role in MS proteomics, with recent Orbitrap and quadrupole time-of-flight (qTOF) systems being widely used for all the omics disciplines, and recent HRMS dedicated to this purpose.[67] Orbitrap mass analyzers, from Exploris 480 to the newest Orbitral Astral, which was designed for deep-coverage proteomics workflows, provide increased mass accuracy and resolving power, making them ideal for deep proteome profiling and precise identification of posttranslational modifications.[68] This is also possible thanks to their ability to distinguish precise isotopic patterns and isotopic ratios, which can be elaborated by proprietary software programs to achieve high selectivity in peptide identification. Also, Orbitrap Astral, introduced in 2023, combines a new electrostatic mass analyzer designed for ultrafast acquisition to ion routing multipole (IRM), to control ion traffic between analyzers. This ensures higher acquisition rates with retained resolution, which results in shorter chromatographic gradients needed compared to "classical" instruments.[68] qTOF instruments, including the well-known Impact II UHR-QqTOF from Bruker, on the other hand, offer fast acquisition speeds and high sensitivity, especially when combined with ion mobility systems, which are particularly advantageous for discovery proteomics and DIA-based workflows.[69] The instruments provided with ion mobility (IM) are able to introduce a fourth separation dimension beyond LC, m/z, and fragmentation, resulting in improved identification of isobaric/isomeric peptides and lower spectral congestion, which is optimal for single-cell proteomics or highly complex samples, for instance. High-resolution MS represents one of the major financial barriers to implementing bottom-up proteomics in drug discovery. The Orbitrap Astral (Thermo Fischer Scientific, Waltham, Massachusetts, USA) costs are typically 1.0−1.2 M$ depending on the regional area, reflecting its cutting-edge speed and proteome coverage. The Orbitrap Exploris 480 (Thermo Fischer Scientific, Waltham, Massachusetts, U.S.), launched earlier, offers a more accessible yet robust solution, with prices generally in the range of $600,000−$800,000 depending on configuration, software packages, and geographical areas (prices are referred to as average for North American and European sellers). Same prices are required to purchase Impact II QTOF (Bruker Corporation, Billerica, Massachusetts, U.S.), Synapt-XS Q-IMS-TOF (Waters Corporation, Milford, Massachusetts, U.S.), and similar instrumentations. Beyond the initial investment, annual operational costs are significant. Indeed, comprehensive service contracts alone often account for 10−15% of the purchase price, while consumables such as LC columns, solvents, labeling reagents, and sample preparation reagents can add thousands of dollars per year in active laboratories.[70] Staff training and bioinformatics support for large MS data management and storage further contribute to recurrent expenses.[71] These costs underscore the importance of shared core facilities, consortia, and industry-academia partnerships, which can distribute the financial burden while ensuring access to advanced instrumentation.

3. MS PROTEOMICS FOR TARGET IDENTIFICATION

The first stage of drug discovery is the target identification and validation stage, which involves biochemical and histopathologic studies on the single proteins or associated pathways involved in the disease onset and progression.[72] The two main methods to address this stage (i.e., knockout studies or Western blotting) are often slow and limited by the need to focus on a small subset of proteins.[73] For simplification, these approaches can be divided into genetic-based and biochemical-based. One of the most applied genetics methods in target identification is the identification of a gene causing a resistance phenotype through a phenotypic screening.[74] Genetic screening can be either direct, aiming to identify the gene underlying phenotype, or reverse, but in both cases.[75] The biochemical approach to target identification focuses on protein expression and can be addressed through different tools.[76] The most traditional ones include Western blot, ELISA analysis, and other instrumental techniques, but the emerging role of MS proteomics has started to become central for this stage.[10] Indeed, MS proteomics allows for comprehensive proteomic profiling, identifying networks that are dysregulated in disease conditions and offering new therapeutic targets.[77] MS proteomics and chemoproteomics enable a middle- and high-throughput identification of proteins that are differentially expressed or modified in diseased versus healthy states (i.e., prioritize speed, reproducibility, and scalability rather than maximal depth of proteome coverage rather low-throughput techniques).[82] By analyzing protein

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

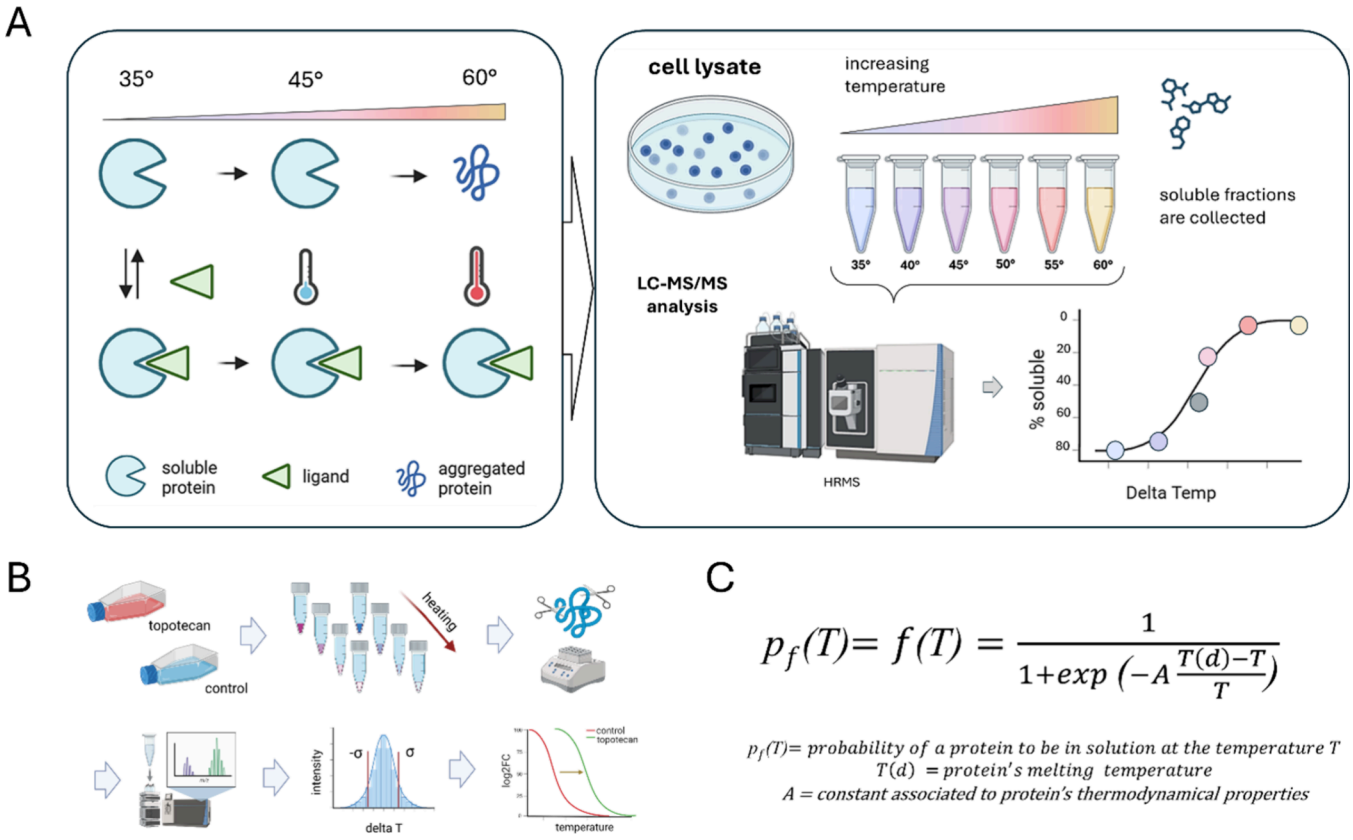


Figure 4. (A) TPP-MS experiment scheme. Adding more compound (ligand) to the protein lysate raises the stability of the target proteins and prevents its aggregation, thus precipitation, even at higher temperatures. The residual concentration of protein−ligand complex in the soluble phase is directly correlated to ligand affinity; thus, K_d can be extrapolated by a titration curve. The image was generated with BioRender. (B) Example of the application of TTP-MS by Fedorov et al., who exploited TTP-MS to characterize the binding and mechanism of action of topotecan in ovarian cancer cells. (C) Fedorov et al. proposed the equation to allow determining protein solubility curves, derived from the first thermodynamic principles. The target proteins were established to distribute outside 3σ (6.7 °C) of the obtained Gaussian distribution[82] of the melting temperature reported in the B panel of this picture.

abundances and their interaction networks, it is possible to determine potential drug targets that are key to the pathological processes of various diseases.[72] The two primary proteomics approaches, i.e., shotgun (discovery-based) and targeted proteomics, complement each other in facilitating target identification. Discovery proteomics, also known as untargeted or shotgun proteomics, is especially useful for identifying novel targets in disease contexts where the underlying mechanisms are poorly understood.[78] This approach allows the identification of thousands of proteins at once, which can then be prioritized for further investigation based on their involvement in disease pathways.[79]

3.1. Thermal Proteome Profiling MS. TPP-MS methods combine thermal denaturation with quantitative MS to identify changes in protein stability upon ligand binding or other environmental changes.[80] TPP-MS relies on the concept that proteins unfold and denature at specific temperatures and precipitate from the sample solution. The processes that can be altered when proteins are bound to small molecules or interact with other proteins or ligands.[80] By simplifying its description, TPP-MS is a translation of the targeted approach of cellular thermal shift assay (CETSA) onto the whole proteome.[81] In TPP-MS, a sample, often cell lysate or tissue lysate, is subjected to increasing temperatures to gradually denature proteins.[82] After heat treatment, the supernatant of the heat-exposed fractions is processed with a standard bottom-up proteomics workflow, and peptide mixtures are analyzed using MS to identify which proteins remain soluble at each temperature, thus being differentially expressed vs the samples not administered with putative ligands.[83] This is valid under the assumption that the ligand thermodynamically stabilizes protein folding, thus preventing heat-induced unfolding and precipitation.[84] Besides a specific drug target, TPP-MS can be coupled to bioinformatics tools to identify the biochemical pathways that are involved in the ligand's MoA through GO's functional enrichment.[85] TPP has been employed to identify drug targets in various studies in the last 15 years, thanks to the evolving MS technologies and the emergence of new platforms, either open access or licensed, for MS data organization and analysis.[86,87] The workflow and principles are illustrated in Figure 4A.

For instance, a recent study from Saei et al. demonstrated the use of TPP to identify thioredoxin reductase 1 (TXNRD1) as a target for auranofin, a gold compound used in treating rheumatoid arthritis.[88] The study revealed that auranofin binds to TXNRD1, stabilizing it against thermal denaturation, which facilitated the identification of this target through MS analysis and confirmed its on target activity.[88] By using TPP, Franken et al. demonstrated its utility in identifying cancer drug targets. They used multiplexed quantitative MS to analyze the thermal stability of proteins in K562 cell extracts in response to various kinase inhibitors, revealing direct interactions with specific proteins, including those involved in cancer signaling

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

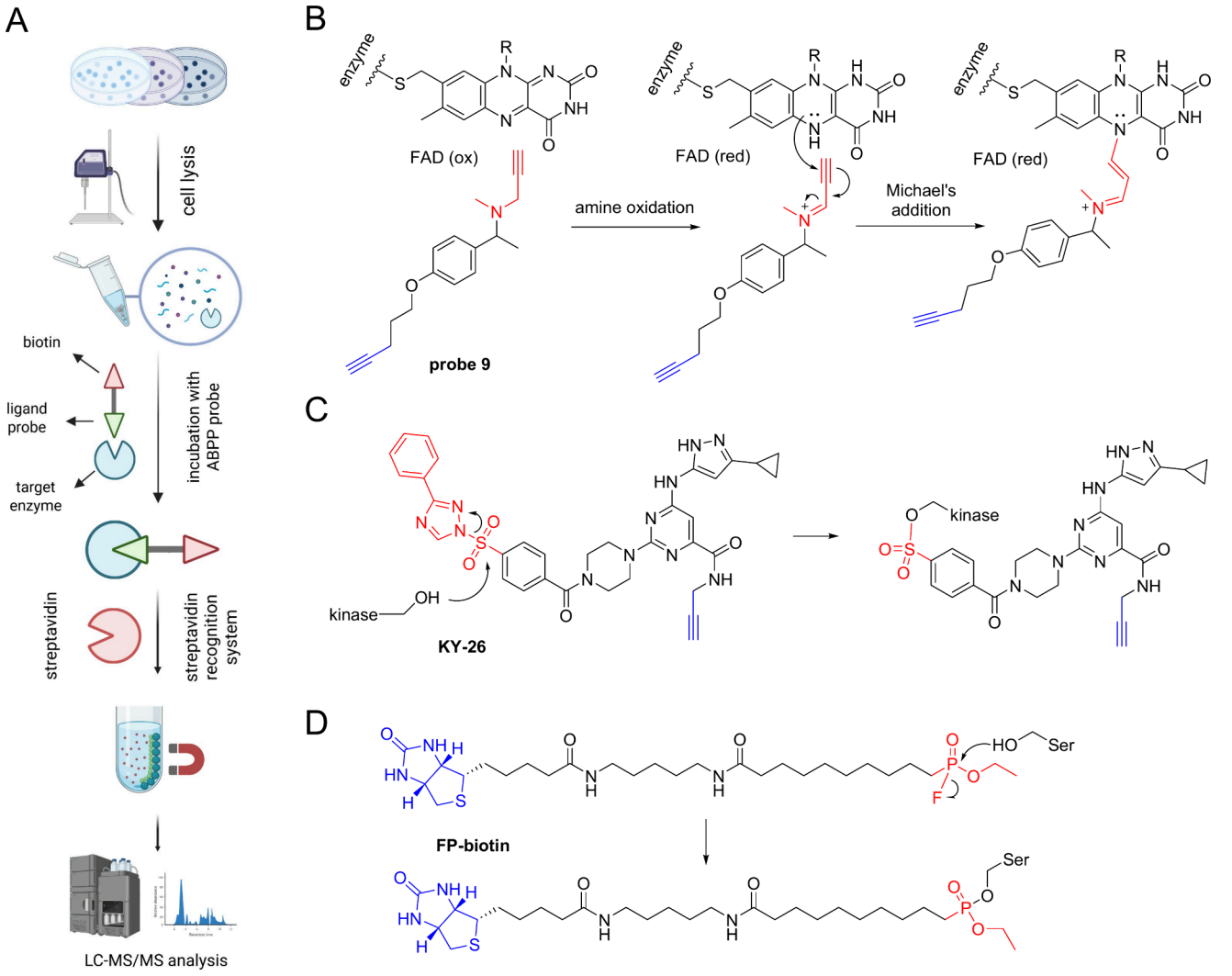


Figure 5. (A) APBB workflow from the cell lysate to MS analysis. The probe is designed so that it targets the catalytic (active) site of the enzyme but cannot be cleaved like the natural endogenous substrate. Thus, the enzyme probe can be fished through a recognizing site, e.g., a biotin moiety, through the biotin−streptavidin system. The image was generated with BioRender. (B−D) Some examples of ABPPs. (A) Probe 9 was synthesized by Krysiak et al. exploited the ABPP strategy to elucidate the enzymatic activity of flavin-dependent oxidases. (B) Probe **KY-26** exploits the electrophilicity of the sulfonamide group to react with tyrosine and lysine residues of several kinases. (C) FB-biotin probe was designed by Tan and colleagues to target clan CA proteases in *P. falciparum*.

pathways like the BCR-ABL path, and the AKT protein.[89] This work highlighted how TPP can elucidate the mechanisms of action for therapeutic agents and identify potential off-target effects.[89] In another comparative study by George et al., TPP was employed to analyze the stability of proteins in acute myeloid leukemia cells treated with losmapimod, a known inhibitor of the mitogen-activated protein kinase MAPK14 (p38α). The results demonstrated that TPP could effectively detect target engagement of losmapimod with MAPK14 and its downstream target MAPKAPK3.[90] Similarly, TPP was employed in the Drug Discovery pipeline of nirlotinib for the identification of target/off-target activity by Marín-Rubio et al, that proved that unlike imatinib, it binds directly also to p38α and inhibit the p38α-MK2/3 signaling axis, which suppressed pro-inflammatory cytokine expression and innate immunity markers in activated monocytes derived from acute myeloid leukemia (AML).[91] In that work, the authors indirectly propose p38α, an off target of the second generation of the tyrosine kinase inhibitors, as new drug target for AML.[91] TPP was also used for natural products target screening and identification.[82] This is the example of topotecan, a well-known and characterized TOP1 inhibitor with antiproliferative activity, that was demonstrated to affect the thermal stability of 14 proteins over the established threshold, all of them belonging to DNA replication and cellular metabolism processes.[82] Workflow performed by Fedorov et al. is illustrated in Figure 4B,C. Overall, TPP-MS offers a more comprehensive alternative to traditional methods, such as differential scanning fluorimetry (DSF) and genetic target validation. Unlike DSF, which requires purified proteins and tests one target at a time, TPP-MS operates directly in complex cell lysates or intact cells, enabling a proteome-wide assessment of ligand-induced stability shifts in a single experiment. This accelerates target deconvolution while preserving the physiological context. Moreover, while genetic knockouts or overexpression systems provide indirect evidence of target involvement, TPP-MS delivers direct, label-free confirmation of compound-protein interactions and selectivity.

3.2. ABPP MS Proteomics. ABPP is an MS chemical proteomics technique that has gained significant traction in drug discovery for target validation and MoA elucidation.[92] ABPP is

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

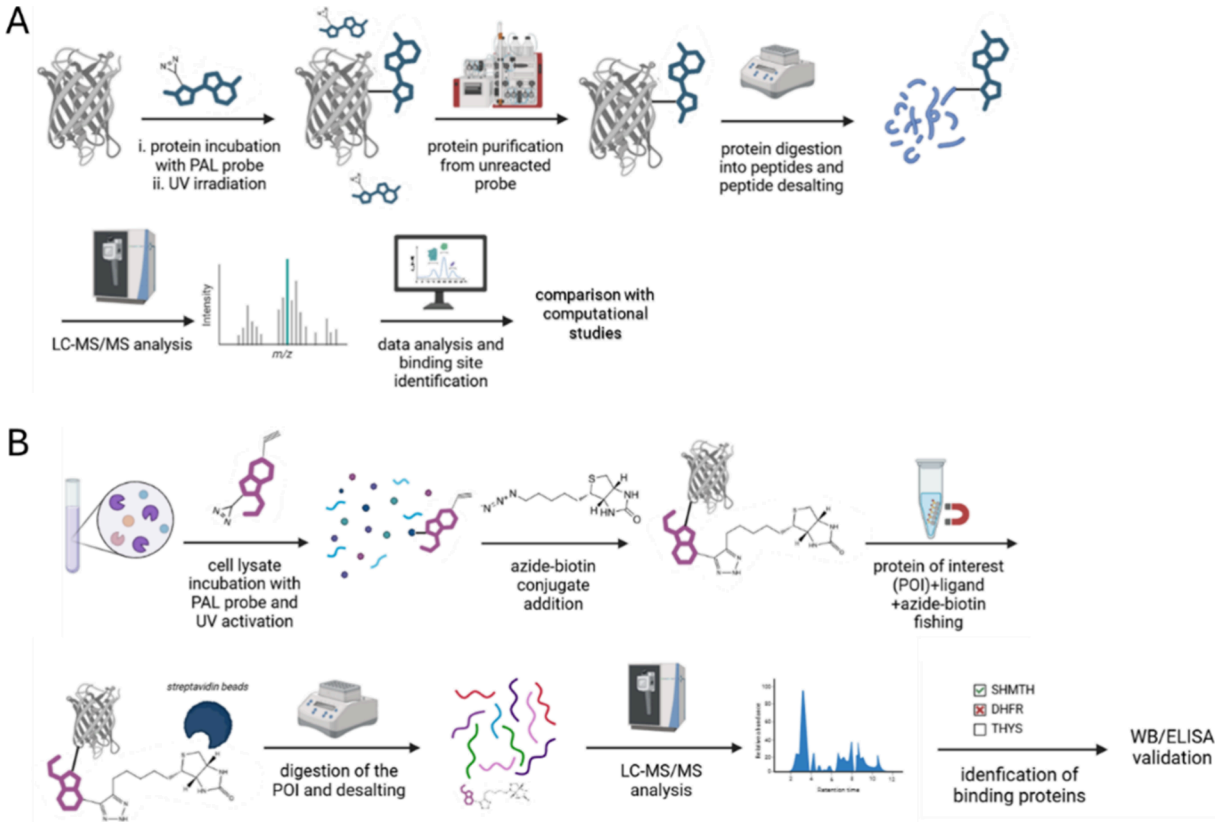


Figure 6. Photoaffinity labeling (PAL) mass spectrometry workflows. A single protein of protein lysate from a cellular experiment (A) is incubated with the PAL probe and photoactivated under a specific UV wavelength and timing. In the case of cell lysate, an azide-biotin bifunctional molecule is added to "fish" the POI and react with the alkyne moiety of the PAL probe. A triazole is formed through click chemistry, and the construct can be recognized with streptavidin beads of affinity columns. (B) Target protein identification in cell lysate. In both pipelines, the experiment includes protein denaturation, cysteine alkylation, hydrolysis into smaller peptides, and MS analysis. The image was created with BioRender.

especially useful for studying enzymes, as it allows profiling their activity in complex biological systems, such as cell lysates or tissue lysates.[93,94] ABPP involves the use of activity-based probes (ABPs), chemical probes designed to interact with active forms of enzymes, often targeting catalytic or reactive residues with a covalent moiety and a recognition system (i.e., biotin−streptavidin).[92] ABPPs are highly selective for enzyme families, depending on the reactive group used (e.g., targeting serine hydrolases, cysteine proteases),[95] and are designed to covalently modify active enzymes, allowing their capture with different recognition systems like biotin−streptavidin and MS analysis.[95] A general workflow for an ABPP experiment is represented in Figure 5A. ABPP is highly specific for active enzymes (as opposed to merely present proteins), distinguishing between active and inactive or nonfunctional enzyme populations.[96] By comparing samples treated with a small-molecule inhibitor (or drug candidate) against untreated controls, it can assess how the inhibitor or drug candidate modulates the activity of specific enzymes, enabling target identification and off-target analysis.[96]

Krysiak et al. exploited the ABPP strategy to elucidate the enzymatic activity of flavin-dependent oxidases, in particular monoaminoxidases (MAO A/B).[97] The probes that the group developed, based on existing MAO inhibitors like selegiline, allowed for selective labeling of enzymes within living cells and offers an effective tool for the drug discovery pipeline of the MAOs both in vitro assays using recombinant proteins and in more complex biological samples such as mouse tissues and human brain cancer cells.[97] The reaction mechanism of this probe is represented in Figure 5B. In 2017, McCloud et al. successfully exploited the ABPP pipeline to design a pan-kinase probe (**KY-26**, Figure 5C) able to target 133 endogenous kinases even in the presence of high ATP concentrations by targeting a conserved lysine in the ATP-binding site.[98] Starting from a sulfonyl fluoride-containing probe (XO44[99]), the authors synthesized a sulfonyl-triazole analogue that contains a triazole with enhanced leaving group ability in order to modify tyrosine and lysine residues, providing an advanced chemical proteomics method for kinase profiling in live cells, expanding the coverage of kinases beyond previous experiments with noncell-permeable probes like kinobeads. Recently, this target identification approach was employed also to accelerate the drug discovery process of parasitic diseases like malaria and leishmaniases by focusing on cysteine and serine peptidases.[100,101]

Focusing on malaria treatments, Tan et al. introduces new broad-spectrum ABPs designed to target clan CA proteases in *Plasmodium falciparum*, which play a critical role in the malaria parasite's life cycle and represent an emerging potential drug target.[100] The authors present new dipeptidic vinyl sulfone probes that can efficiently target both endopeptidases and dipeptidyl aminopeptidases in *P. falciparum* and mammalian cells, with a free N-terminal tryptophan and a fluorophore, which are partially or totally cell permeable[100] With the aim to explore serine proteases (SPs) in *Leishmania* spp as potential drug targets, given their critical roles in parasite survival and infection, two years later, Porta et al. combined ABPP and (iTRAQ)-based quantitative MS proteomics, as reported in Figure 5D.[101] A library of fluorophosphonate probes was used to detect and identify active serine hydrolases in different life stages

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

Table 3. Chemical Requirements of a PAL Probe and the Corresponding Design Considerations to Be Considered during Experimental Planning

component	function	chemical features	design considerations
recognition element	binds selectively to the target protein	derived from the bioactive compound (parent molecule), thus must retain its activity	structure−activity relationship (SAR) must be preserved, i.e. probe should compete with native ligand
photoreactive group	forms covalent bond upon UV activation	small, stable in the dark. Activated by UV (typically 254−365 nm). Generates reactive species (carbene, nitrene, radical, etc)	diazirines and aryl azides are less bulky and more selective. Benzophenone is more stable but bulkier
linker/spacer	minimizes steric hindrance and provides flexibility	chemically inert; variable length; often flexible and hydrophilic	length must allow photoreactive group to reach protein surface without interfering with binding
reporter tag/ handle	enables enrichment, visualization and/or detection	must be bio-orthogonal, small enough not to disrupt binding, reactive for click chemistry	tag is often introduced via click chemistry post-labeling to reduce probe size in cell
UV activation wavelength	needed to activate photoreactive group	match the wavelength with the photophore used	choose based on cell/tissue compatibility and minimal biological damage
solubility/ permeability	ensures bioavailability and cell penetration	balanced logP. Excessive hydrophobicity or polarity should be avoided	compound should retain the activity of the parent molecule
synthetic accessibility	practical synthesis and modular design	compatible with multistep organic synthesis	avoid labile intermediates or protecting group conflicts
photostability	prevent degradation prior to UV exposure	must remain stable under storage and biological assay conditions	store in dark. Use new stock solutions and photoactivation can occur unintentionally.

of Leishmania, and two proteins of *L. Mexicana* (carboxypeptidase (LmxM.18.0450) and prolyl oligopeptidase (LmxM.36.6750) were identified as possible druggable targets.[101] The success and versatility of the ABPP approach to characterize cell proteome, e.g., after the administration of investigational compounds, pushed the bigger chemical suppliers to commercialize standard ABPP kits by encompassing the chemical moiety of the most common protein motifs, including the ATP-binding site of human kinases. For instance, a kit simplification of a kinase-ABPP was employed in the medicinal chemistry studies on gastric cancer (GC) by Choi et al.[102] In fact, the authors exploited desthiobiotin-ATP, which is able to specifically and covalently target the catalytic lysines of kinase to test the activity of the heat-shock protein 90 (HSP90) inhibitor **AUY922**. Through the whole cell ABPP assay, the microtubule-associated serine/threonine kinase-like (MASTL), known for its involvement in mitotic control and cancer progression, was identified as specifically upregulated by HSP90. Further MS proteomics and bioinformatic studies were employed by the authors to validate this protein as a GC target.[102] In the context of drug discovery, ABPP-MS is more informative than traditional enzyme assays (e.g., colorimetric or fluorometric substrate-based assays) that measure activity toward one purified enzyme at a time. ABPP-MS can monitor hundreds of enzymes directly in complex proteomes or cells, capturing native context and competition with inhibitors. Also, ABPP-MS is superior and safer to using radiolabeled compounds (e.g., [^{3}H]- or [^{14}C]-labeled inhibitors) and to "traditional" affinity chromatography/pull-down assays, which rely on immobilized compounds, often yielding nonspecific binders or losing weak/transient interactions, in contrast to ABPP probes that are selective only toward active enzymes.

3.3. PAL MS Proteomics (PAL-MS). Unlike ABPP, which is activity-specific, as it labels only the active forms of enzymes through a covalently stable bond, PAL uses photoactivatable probes, which typically contain a photoreactive group that becomes activated upon exposure to light (usually UV light).[103] Once activated, this group forms a covalent bond with any protein in proximity.[104] Unlike ABPP, PAL is not activity-specific. Indeed, it labels proteins based on spatial proximity to the probe rather than their catalytic activity.[11] PAL probes are generally used to study noncovalent interactions between small molecules (such as ligands, inhibitors, or drugs) and proteins.[105] PAL probes consist of a photoreactive group (e.g., diazirines, aryl azides, or benzophenones) that is inert until exposed to UV light, a recognition element (such as a drug, ligand, or small molecule) that binds noncovalently to the target protein before photoactivation, and, often, a reporter tag (e.g., biotin or fluorescent group) is also included for detection.[106] The photoreactive group must be stable under experimental conditions but reactive enough to form a covalent bond with the target upon light activation. Proper placement of the photoreactive group on the probe is crucial.[9] If the photoreactive group is too far from the interaction site, it may not form a covalent bond with the target, even after light activation. Probes must have balanced solubility to ensure they reach their target, especially if the target is in a hydrophobic environment (e.g., within cell membranes) or a hydrophilic one (e.g., cytosol).[9] Experimental PAL workflow is represented in Figure 6. PAL probes can capture transient interactions or noncovalent binding events that are often difficult to observe by other techniques.[107] The specificity in PAL typically comes from the binding affinity of the recognition element (e.g., a drug or ligand) to its target before photoactivation, but once UV light is applied, nearby proteins can also be labeled, resulting in potential off-target labeling.[108] On the other hand, it requires ad hoc synthetic pathway to obtain a suitable chemical probe encompassing both the UV-activable moiety and the recognition portion (usually, an alkyne able to react with an azide-biotin system through copper-mediated click chemistry).[109]

The chemical design of probes is a critical aspect of PAL, requiring a delicate balance between functional necessity and minimal perturbation of the original drug's properties. PAL probes must meet specific criteria (as outlined in Table 3) while closely preserving the parent molecule's chemical structure, affinity, and potency for its target(s) within cellular environments.[110] This dual requirement applies to both the photoreactive moiety and the pull-down moiety, commonly an alkyne system.[111] The photoreactive group, whose steric hindrance significantly reduces upon photoactivation, must be carefully selected to avoid altering the drug's core pharmacophore. Preliminary assays are essential to confirm that the modified probe retains its intended target affinity.[104] Furthermore, to prevent introducing bias in binding site proximity estimations,

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

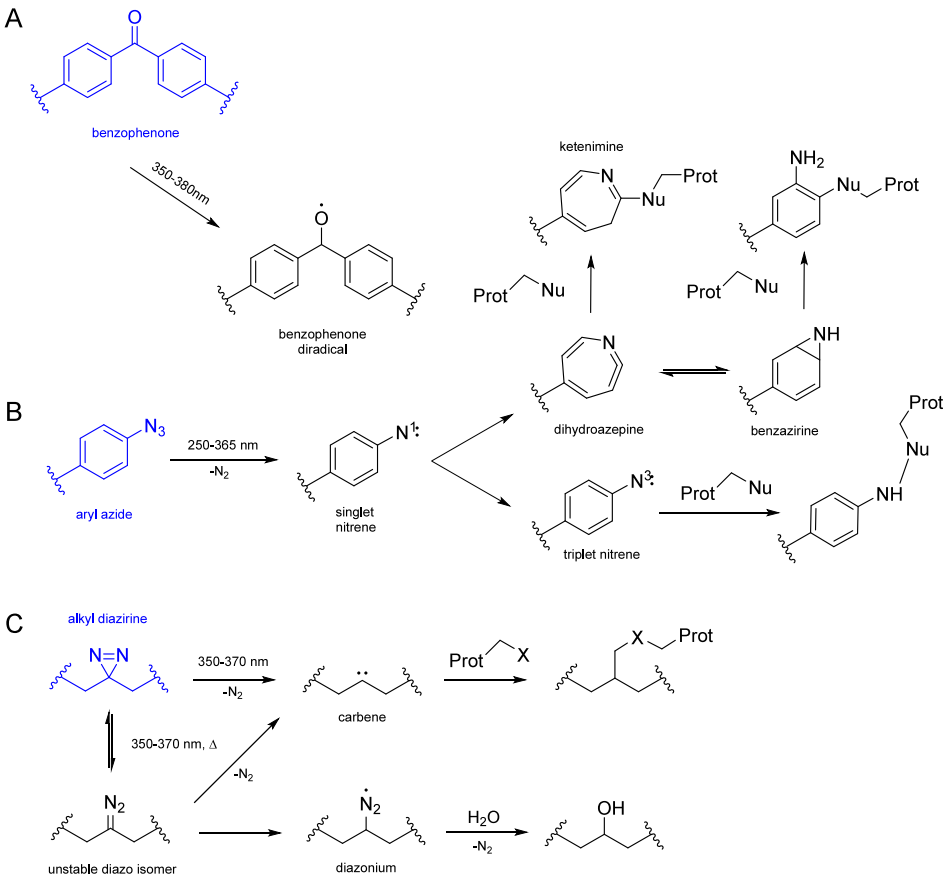


Figure 7. Reaction scheme of the three main classes of PAL photoreactive moieties: (A) benzophenones, (B) aryl azides, and (C) diazirines with their UV activation pathways. On the left, in blue, the initial species is reported.

the reactive unit should be directly attached to the original drug scaffold, ideally without a linker.[106] In contrast, the design of the recognition (pull-down) system is more straightforward. Its primary role is to be readily accessible for interaction with the "fishing" system within the cell lysate. Therefore, spacers, such as 2−3 units of polyethylene glycol (PEG) or poly glycines, are often incorporated to enhance its exposure and optimize retrieval efficiency.[104]

One commonly used photoreactive group in PAL is **aryl azide**. Aryl azides are activated by UV light at wavelengths around 250−365 nm, where they undergo a transformation that releases nitrogen gas and produces a highly reactive nitrene intermediate.[112] Nitrenes can insert into a variety of bonds, such as C−H, N−H, and O−H, allowing the probe to covalently attach to its target. A key advantage of aryl azides is their high reactivity, which makes them suitable for labeling a wide range of biomolecules. However, nitrenes are also short-lived and unstable, which can result in off-target interactions and nonspecific labeling.[113] Additionally, the need for UV light activation at low wavelengths can damage biological samples, including proteins and nucleic acids, limiting the applicability of aryl azides in certain systems, particularly in live-cell studies, where preserving biological function is crucial.[109]

Another important class of photoreactive moieties used in PAL is **diazirines**. Upon UV light exposure at longer wavelengths (around 350−370 nm) than arylazides, diazirines decompose to produce a reactive carbene species. Carbenes are highly reactive and can insert into C−H and N−H bonds, making diazirines versatile for covalently attaching to a wide variety of targets.[114] One of the key advantages of diazirines over aryl azides is their longer activation wavelength, which reduces the likelihood of photodamage to biological systems.[111] Moreover, diazirines are small and compact, meaning they can be incorporated into probes without significantly disrupting the structure and function of the molecules being studied. This makes them particularly useful in labeling complex environments like proteins, lipids, or membranes.[115] However, alkyl diazirines tend to be less reactive than aryl azides, which means that careful control of light exposure and experimental conditions is necessary to achieve efficient labeling.[116]

A third photoreactive group that is occasionally used in PAL is benzophenone. Unlike aryl azides and diazirines, benzophenones are activated by light at longer wavelengths, typically in the near-UV or visible range (350−380 nm). Upon light activation, benzophenones form a triplet-state excited intermediate that can abstract hydrogen atoms from nearby C−H bonds, generating a covalent bond between the probe and its target. Benzophenones are more selective than both aryl azides and diazirines because they primarily react with C−H bonds, reducing the chances of nonspecific labeling.[117] Furthermore, their ability to be activated at longer wavelengths makes them ideal for studies in living cells or tissues where minimal photodamage is essential. However, benzophenones require relatively long exposure times to achieve effective labeling, which can limit their utility in fast-acting or dynamic biological processes.[118] Thus, the choice of photoreactive group must be carefully matched to the specific requirements of the experiment, balancing factors such as reactivity, photowavelength,

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

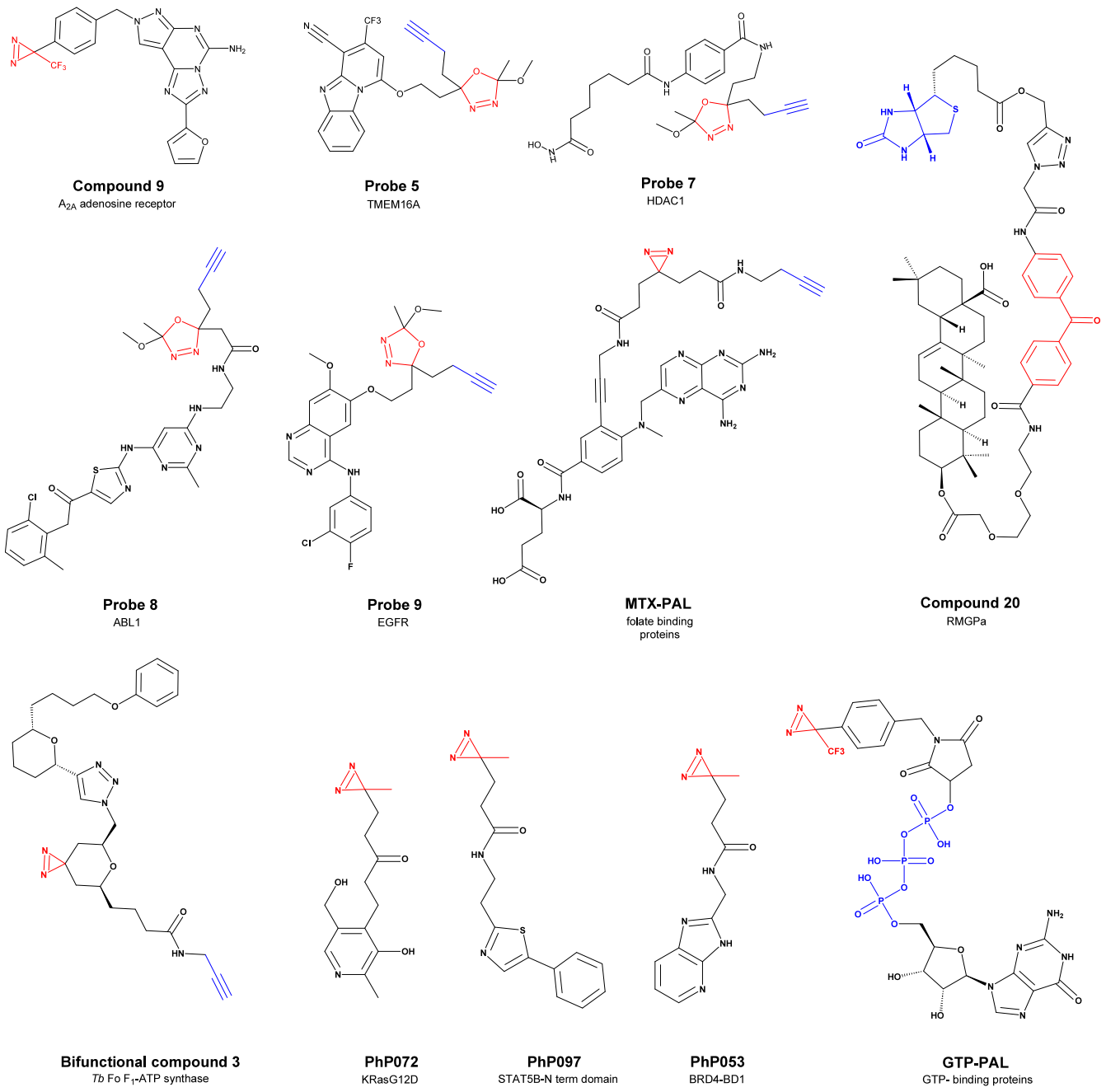


Figure 8. Some examples of PAL probes from the literature with different photoactivable moieties. Blue moieties represent photoactivable fragments, while red moieties are part of the cellular recognition system with the fishing probe, either involved in a covalent mechanism (e.g., terminal alkyne for copper-catalyzed azide−alkyne cycloaddition) or triphosphate for trivalent metals complexation.

specificity, and compatibility with the biological system under study. A scheme reporting the three main families of PAL probes and their UV reactivity is reported in Figure 7.

Since aryl and alkyl diazines are the most used class of photoreactive probes in chemical biology and biochemistry, also outside PAL purposes, their reactivity toward amino acidic residues was recently investigated by West et al., which tested the reactivity of FMOC-protected small aryl and alkyl azides toward equimolar mixtures of N-acetyl, O-methylated natural amino acids.[119] The authors demonstrate that alkyl and aryl diazirines follow distinct labeling mechanisms. Alkyl diazirines preferentially react with acidic amino acids, such as glutamic acid and aspartic acid, through a reactive diazo intermediate, which is generated prior to carbene formation. This reactivity is strongly influenced by pH, with increased labeling occurring under more acidic conditions, where these residues are protonated and more susceptible to nucleophilic attack. In contrast, aryltrifluorodiazirines primarily label proteins through a short-lived carbene intermediate, which inserts indiscriminately into multiple amino acids but shows a particular preference for cysteine. Unlike alkyl diazirines, the labeling efficiency of aryl diazirines is largely independent of pH, indicating that the carbene pathway is dominant and relatively unaffected by local protonation states.[119] Proteomic analysis further confirms that alkyl diazirines show a strong preference for labeling membrane proteins, which often possess acidic regions with elevated pK_a

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

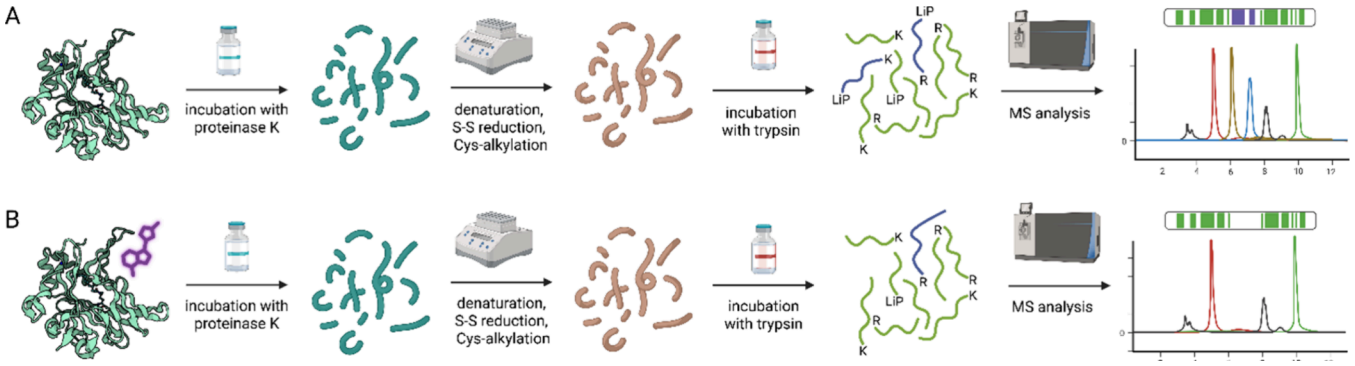


Figure 9. LiP-MS experimental schemes without (A) and with (B) tested inhibitor(s). Cell (or tissue) lysate is then incubated with low concentrations of Proteinase K, thermally denatured, and thiols are reduced with DTT, treated with iodoacetamide, and finally redigested with trypsin to obtain smaller peptides. Resulting peptides are analyzed by LC−MS. Experiments with ligand incubation are compared to control samples, and the least covered sequence regions are identified from a differential analysis.

values due to their lipid environment. These findings suggest that the local charge and electrostatic environment of proteins play a crucial role in determining the outcome of PAL experiments and that the choice of the most suitable UV photoactivable moiety determines the good outcome of the experiment.[119]

The PAL technique is more recent than ABPP; thus, its applications are still confined to a limited area of drug discovery research. It has already been employed with success in different fields. For instance, in 2017 Muranaka et al. successfully characterized the exploitable binding sites of the A_{2A} adenosine receptor with PAL, **compound 9**, by exploiting a diazirine-containing chemical probe.[120] This year, Bon et al. have synthesized a series of oxadiazoline probes that, upon UV activation, generate irreversible diazoketones, which were applied to modify the HDAC1 inhibitor vorinostat (**Probe 7**), the ABL1 inhibitor **Probe 8**, the TMEM16A inhibitor **Probe 5**, and EGFR inhibitor gefitinib (**Probe 9**)[121] (Figure 10). Their medicinal chemistry was employed to develop **OTX-015** (**MK-8628**) and validate its binding to BRD4-BD1 in a dose-dependent manner without altering binding affinity, resulting in efficient treatment for patients with acute leukemia.[121] With the purpose of generating a tool to screen the medicinal chemistry of the folate cycle, Takamura et al. successfully designed and tested on *E. coli* lysates a methotrexate (**MTX**) analogue carrying a side spacer linked to the p-aminobenzoic acid of **MTX**, which carry both a UV photoreactive diazirine moiety, and an alkyne for probe fishing (Figure 8).[122]

The team demonstrated the specificity of the probe versus the folate moiety containing protein, with no binding, even at high concentrations, to albumin and other serum abundant proteins providing a useful chemical biological tool to investigate microbial folate-binding proteins in health and microbiome-related diseases.[122] Similarly, with a remarkably innovative approach for that time, Kaneda et al. successfully exploited previous ABPP experiments to design of GTP-binding protein (diazirine-based **GTP-PAL**, Figure 9) probes to investigate the selectivity of human GTPases.[123] In this case, the authors have taken advantage of the triphosphate moiety as a recognition motif for the Fe^{III} IMAC stationary phase. Starting from the same ligand structure, George Cisar et al. exploited the benzophenone derivative with a more common alkyne residue for the same purpose.[124] Despite the fact that the chemical modification of natural products is more challenging than synthetic-derived molecules, PAL-MS was anyway employed also for the characterization of the pharmaceutical activity of natural products isolated from extracts. This was the case of Zhang et al., who exploited the benzophenone photoactivable moiety to synthesize a derivative of oleanoic acid.[125] Through this chemical construct, his team investigated all the potential ligands of oleanoic acid as therapeutical targets.[125] Parallelly, Tulloch et al. developed a natural product library of simplified analogues of acetogenin-type ether and used PAL to investigate the target. Their lead compound, a bis-tetrahydropyran 1,4-triazole inhibitor, was modified into a bifunctional photoaffinity probe.[126] All of the chemical structures of the examples of PAL here discussed are reported in Figure 8.

Through this, the team identified the FoF1-ATP synthase, specifically targeting its α- and β-subunits, as the major target, disrupting the parasite's ATP production via oxidative phosphorylation.[126] PAL was also used for FBDD pipelines. Indeed, Ábrányi-Balogh et al. developed a screening approach combining evolutionary-optimized fragment pharmacophores with a diazirine photoaffinity handle, activated by a photocatalyst.[127] Thanks to this, the authors identified binding sites on BRD4 using a fragment that binds in both the primary acetyllysine site and additional sites, discovered new allosteric sites on the oncogenic KRasG12D protein, previously considered challenging to target, and designed a library of fragment hits against the N-terminal domain of STAT5B, which has no known ligands, with evidence of inhibiting cancer cell viability.[127] PAL and ABPP are central tools in chemical biology. While PAL is suitable for mapping noncovalent binding interactions between small molecules and proteins through proximity-based photolabeling, ABPP is uniquely suited to probe functional enzymatic activity by covalently modifying active-site residues of enzymes in an activity-dependent manner. Their distinct but complementary mechanisms make them powerful strategies for elucidating target engagement, mechanisms of action, and off-target profiles in complex biological systems. Their comparison is reported in Table 4. Overall, PAL-MS represent a solid and more valuable alternative to single amino acid mutation (like alanine scanning) and mutant mapping for target ID and binding site description. Compared to these older techniques, PAL-MS results are more robust, i.e., also allowing to work with the intact protein, and therefore the system is more adherent to the biological system represented. Also, considering time needed to purify and test a library of a protein mutant, despite not being a short-lasting experiment, PAL-MS may faster these preliminary processes. Furthermore, when combined to MD and computa-

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

Table 4. PAL's Strength and Limitations in Comparison with the ABPP Technique

feature	ABPP	photoaffinity labeling (PAL)
mechanism	activity-based covalent labeling of enzymes	proximity-based covalent labeling via UV activation
specificity	highly specific for active enzymes	depends on the binding affinity of the recognition element, i.e. less specific after UV activation
aims	enzyme profiling, target identification, mechanism of action studies	mapping noncovalent drug−protein interactions, studying protein complexes
proteins studied	enzymes	no limitations
labeling type	covalent modification of active sites	proximity-based cross-linking after UV activation
advantages	functional insight into enzyme activity	captures transient interactions and studies nonenzymatic proteins
limitations	limited to enzyme profiling	potential nonspecific labeling at high μM K_d's. Chemical probe might be nonsynthetically accessible

tional models, as in the previously mentioned works, it represents a valuable tool for the early phases of drug discovery.

3.4. Limited Proteolysis MS. LiP coupled to MS (LiP-MS) has recently emerged as a highly informative chemoproteomics strategy to study target identification and binding site characterization.[128] Its principle lies in the fact that ligand binding alters the structural dynamics and accessibility of proteins, as in their folded state they are only partially accessible to proteases, and exposure of flexible regions or pockets determines cleavage susceptibility.[129] When a small molecule interacts with a protein, the induced conformational stabilization or rearrangement changes the landscape of protease accessibility, thus modifying the proteolytic cleavage pattern. LiP-MS leverages this principle by applying a controlled, partial digestion of the proteome, followed by global peptide analysis with MS, to detect changes in proteolytic fingerprints between ligand-treated and untreated samples.[130] These differential patterns are then interpreted as signatures of direct or indirect molecular engagement. An LiP-MS experiment begins with the preparation of a biological system, very often cell culture lysate or an organelle, in which protein conformation is maintained in a near-native state.[131] The proteome is divided into two or more parallel samples, one exposed to the compound of interest (often at increasing concentrations) and the other serving as a reference control.[132] The key step is the application of a nonspecific protease, most commonly proteinase K, under mild/diluted conditions and controlled temperature, which ensures LiP rather than complete digestion.[133] The enzyme is allowed to act briefly, long enough to generate cleavage at solvent-exposed and flexible regions but not to degrade proteins entirely. This step generates structurally informative cleavage products, the distribution of which depends sensitively on the conformational state of the proteins.[134] The reaction is then quenched with a reducing agent, remaining proteins/peptides are denatured, and the resulting protein fragments undergo a second digestion with a site-specific protease such as trypsin, which standardizes peptides for MS detection and quantification. Quantitative analysis then identifies peptides whose abundance significantly differs between the two conditions, namely, conformotypic peptides.[135] Unlike affinity-based methods, LiP-MS does not require derivatization or immobilization of the ligand, which can alter its binding properties or introduce artifacts. It also circumvents the need for engineered protein tags, which are not always feasible in complex systems or clinical samples. All these premises make LiP-MS cheaper in terms of economic resources and time spent on sample processing. Because LiP-MS reports on conformational changes, it goes beyond mere detection of binding and can capture allosteric effects or secondary structural rearrangements that are invisible to classical pull-down approaches.[136] Nevertheless, several limitations must be recognized. Proteins expressed at low abundance or proteins with few solvent-exposed protease cleavage sites may not yield a sufficient signal to detect changes. The method also requires careful optimization of protease concentration and digestion time.[133] Overdigestion can obscure site-specific effects by generating excessive fragmentation, whereas underdigestion reduces the discriminatory power of the assay. Moreover, not all observed proteolytic differences directly correspond to ligand binding. Some may reflect indirect consequences such as protein stabilization within a complex, conformational shifts caused by downstream signaling, or even protease preference biases.[133] A schematic representation of an LiP-MS general experiment is depicted in Figure 9.

LiP-MS was successfully used in numerous medicinal chemistry programs aimed at the identification of drug targets. In particular, but not limited to, after phenotypic screening and with vegetal extracts with pharmaceutical properties. This is the case of Zhao et al., who found that cinobufagin can revert bortezomib resistance in myeloma cells by promoting SEC62 expression, thus its association with the partner TRPM4, that resulted in the activation of necrosis by sodium overload (NECSO) pathway. In particular, the authors demonstrated Cinobufagin dose-dependent binding to TRPM4 by combining MD with LiP-MS performed on 8226-BTZR cells.[137] Similarly, Yong-Chun et al. demonstrated through LiP-MS that Sanguinarine halts oral squamous cell carcinoma by binding to PKM2/TFEB. The researchers treated CAL27 cells with three different concentrations of sanguinarine, pretreated with Proteinase K, and with pancreatic lysate enzymes after denaturation. The only six proteins that resulted as differentially expressed between treated cells and control cells were considered putative targets. Among them, PKM2 had the lower p-value, indicating that it is the preferred target.[138] In the field of marine drugs, Gracilioether A was found to be a micromolar inhibitor of USP5 (ubiquitin carboxyl-terminal hydrolase 5) by DART (drug affinity-responsive target stability) combined with LiP-MS. The authors exploited the former technique to perform target identification studies, while LiP-MS was used here to decipher the Gracilioether A binding site to USP5.[139] With the same workflow, the same group also identified the polybrominated Mycalin A as a GRP75 inhibitor, with consequences on p53 cytosolic retention and cell survival.[140]

Overall, TPP-MS, PAL-MS, ABPP-MS, and LiP-MS represent four valid chemical proteomics strategies for target identification and other steps of the drug discovery pipeline. They all possess pros and cons according to the aim of the research context, and their use must be carefully evaluated also according to the financial resources dedicated to the considered project, the available scientific expertise (e.g., despite its precious outcome, PAL-MS strictly necessitates a solid organic chemistry expertise and equipment, and TPP-MS requires physical chemistry expertise to obtain insight into thermodynamical data), and the available time. Furthermore, all these MS chemoproteomics techniques require a preliminary phase to optimize the working conditions. A brief comparison of the prerequisites, time, and costs associated with each technique is illustrated in Table 5,

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

intended as a simplified preliminary guide to choose the best experiment.

3.5. Biofluid-Based MS Proteomics. Biofluids, such as plasma, serum, cerebrospinal fluid (CSF), urine, and saliva, provide a preclinically accessible and usually little invasive source of information for target identification in Drug Discovery.[141] In contrast to cell or tissue proteomics, which often requires invasive procedures and is limited to a single time point, biofluid MS proteomics allows longitudinal sampling and can represent a systemic view of disease-associated protein dysregulation that can be used for novel drug target investigations.[142] These features make biofluid proteomics particularly suited for translational applications, where the discovery of novel therapeutic targets must be aligned with clinical feasibility.[143] The principal challenge of biofluid-based MS proteomics is related to its intrinsic dynamic range of protein concentrations.[143] Plasma, for instance, contains proteins spanning over 10 orders of magnitude in abundance, with albumin and immunoglobulins accounting up to 80% of total protein content.[144] Such highly abundant proteins (namely, HAPs) frequently mask low-abundance proteins (LAPs) that may carry greater biological relevance as disease-associated targets.[144] To overcome this limitation, workflows typically include immunodepletion of the most abundant proteins, often through affinity chromatography columns, followed by fractionation strategies to reduce sample complexity.[145] These solutions are usually built to remove either only albumin and immunoglobulins or to include antibodies also toward the complement and coagulation system.[146] Applications of fluid-based MS proteomics for target identification are broad. Under oncology and chronic inflammatory conditions, plasma proteome profiling has uncovered dysregulated proteins of complement and coagulation cascades, which serve as potential therapeutic targets as well as biomarkers for disease progression. This is the example of Mazidi et al., who used plasma-based proteomics to identify drug targets for ischemic heart disease (IHD).[147] The authors initially identified 13 proteins that showed potentially causal associations with IHD, including N-terminal prohormone of brain natriuretic peptide and proprotein convertase subtilisin/kexin type 9. The combination of MS proteomics results with transcriptomics showed that FURIN is a potential novel target and matrix metalloproteinase-3 a potential repurposing target for IHD.[147] CSF MS proteomics also has become an indispensable tool in neurodegenerative research, where altered levels of synaptic, axonal, and myelin-associated proteins provide both mechanistic insight and candidate drug targets for diseases such as Alzheimer's disease, multiple sclerosis, and other neurodegenerative conditions.[148,149] This is the main purpose, for example, of the Global Neurodegeneration Proteomics Consortium (GNPC, https://www.neuroproteome.org/), a public−private partnership to build one of the world's largest harmonized proteomic data sets, which now contains protein measurements from multiple platforms from more than 35,000 biofluid samples from CSF and plasma.[150] Urine proteomics has been particularly valuable in nephrology drug discovery, where proteins associated with tubular dysfunction or glomerular injury not only serve as indicators of drug-induced nephrotoxicity but also reveal novel pathways involved in disease progression, as in the case of chronic kidney disease (CKD).[151] More recently, salivary proteomics, either whole cell or cell-free, has gained traction as a noninvasive method to identify systemic immune alterations,

Table 5. Comparison of Prerequisites, Time, and Costs Associated with TPP-MS, PAL-MS, ABPP-MS, and LiP-MS

technique	economic cost	scientific expertise needed	time needed	software commonly used
TPP-MS	medium-high (many MS experiments, reagents, cell cultures, etc.)	medium-high (thermal shift expertise required and physical-chemical interpretation of results)	medium-long (several days experiment + data analysis), ∼1−3 weeks	TPP-Analyzer (R), MaxQuant, Perseus, FragPipe
PAL-MS	high (probe synthesis + probe tests, MS experiments)	high (chemical probe design + proteomics)	long (weeks for probe design, test, and synthesis, days for experiment), >1 month	MaxQuant, Proteome Discoverer, Skyline, in-house scripts
ABPP-MS	high (activity probes + enrichment + MS experiments)	high (chemistry for probe design + MS proteomics)	medium-long (probe synthesis + days for experiment), ∼1 month	MaxQuant, Proteome Discoverer, custom R/Python scripts
LiP-MS	medium-low (protease + MS experiments, no probe synthesis needed)	medium (proteomics expertise, protease optimization)	medium-short (few days), ∼1−2 weeks	MSFragger, FragPipe, MaxQuant, Skyline

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

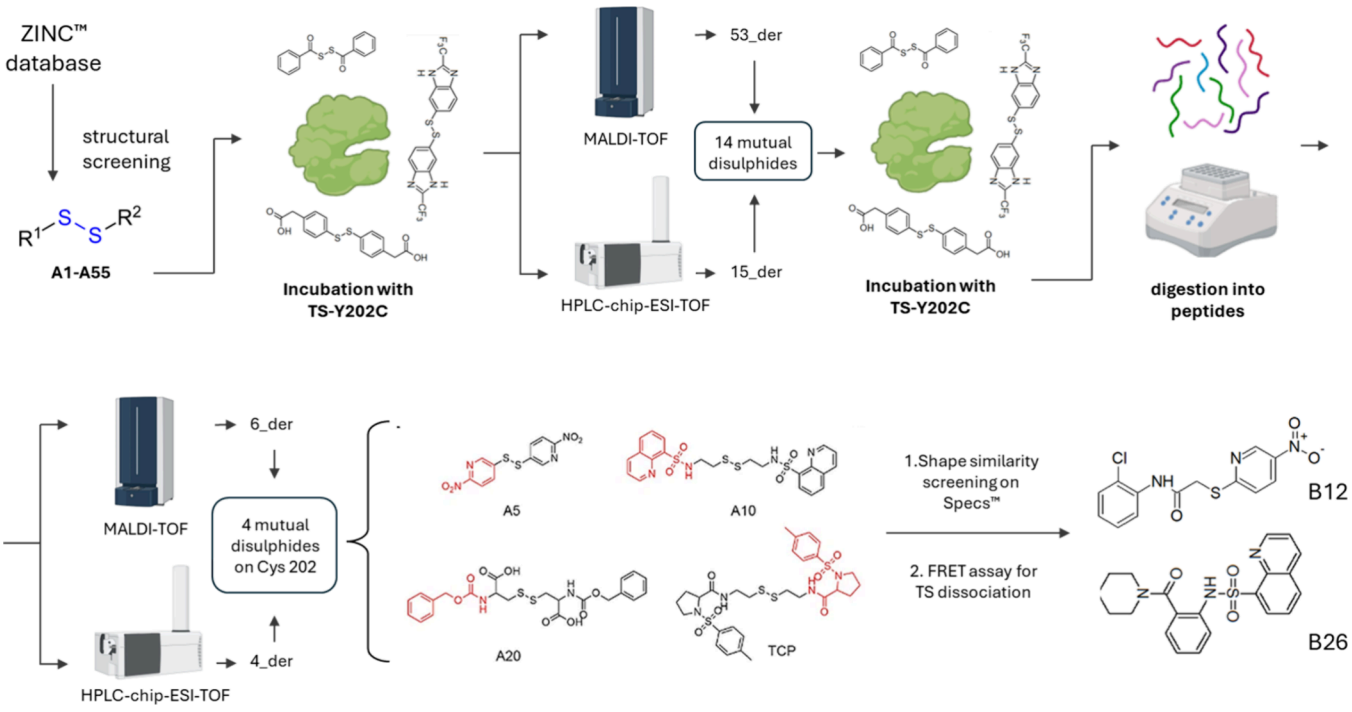


Figure 10. Experimental pipeline used by Costantino and colleagues[173] to map reactive cysteines on mutated TS surface. The authors exposed the TS C195S-Y202C variant to a library of 55 disulfides from a larger database in ZINC and analyzed both derivatized intact protein with MALDI-TOF in a top-down experiment and in HPLC-chip-ESI-TOF. The candidates that were found bound in both MALDI-TOF and ESI-TOF experiments ($n = 14$) were digested into peptides with ESI-MS and submitted to sequencing analysis to evaluate Cys180, Cys199, Cys202, and Cys210 binding. Only disulfides that bound selectively to Cys202, not involved in catalytic pocket reaction nor in mRNA interaction, were selected for optimization ($n = 4$). Those fragments were submitted to shape similarity on the Specs data set. The best 25 candidates (**B1**−**B26**) were submitted to the FRET experiment for their ability to dissociate TS, and **B12** and **B25** were selected as the best hit compounds. This allowed to develop compound **E7**, a TS destabilizer lead compound.[173] The image was created with BioRender.

demonstrating that even peripheral fluids can be informative for target discovery.[152]

4. MS PROTEOMICS FOR HIT IDENTIFICATION AND VALIDATION

4.1. MS-Guided FBDD. Hit identification is a critical step in the early stages of drug discovery, where molecules with biological activity against a target are identified from large chemical libraries.[153] MS-based proteomics facilitates hit identification by providing insights into the molecular mechanisms of potential drug candidates.[153] For instance, MS Proteomics has become a key tool in the FBDD process, through high specificity in detecting fragment-target interactions, elucidating the binding mechanisms, and validating hits at an atomic level besides X-ray crystallography.[154] FBDD is a powerful approach used to identify small chemical fragments that bind to a biological target.[155] These fragments are typically much smaller than conventional drug-like molecules, containing fewer than 20−30 atoms.[156] Once these fragments are identified, they are optimized into more potent compounds through various strategies.[157] FBDD offers several advantages over X-ray crystallography or other techniques, including higher hit rates (smaller fragments are more likely to bind due to their reduced complexity), better chemical space exploration, as fragments represent a broader range of chemical diversity, and efficient optimization, as MS identified fragments serve as starting points to grow into potent, drug-like molecules.[158] FBDD has successfully been used in the discovery of several approved drugs (e.g., vemurafenib for cancer and venetoclax for chronic lymphocytic leukemia).[159,160] Native MS is a powerful technique for this purpose because it allows for the direct detection of intact protein−ligand complexes without requiring labels or complex secondary reagents.[161] In native MS, the protein and fragment are ionized under conditions that preserve the noncovalent interactions.[79] When a fragment binds to a protein, the mass of the resulting protein-fragment complex increases, which can be directly measured by a mass spectrometer. The comparison between the mass of the unbound (apo) protein and the bound (holo) protein-fragment complex provides evidence of binding.[162] Additionally, native MS can detect multiple fragment-binding events if the protein has more than one binding site, offering insights into the stoichiometry of fragment binding.[163,164] For example, if a protein has two identical binding sites, native MS can show whether one or both sites are occupied by fragments, which is crucial information for subsequent fragment optimization and drug design.[12] To facilitate high-throughput fragment screening, pooling strategies are sometimes used in MS-based FBDD. Small libraries of fragments are pooled together and incubated with the target protein, and MS is used to detect which fragments bind.[79] In this strategy, MS can deconvolute the pooled fragments to determine which specific molecules are responsible for binding.[165] An example of FBDD through native MS is the exploration of CYP121 as a target for drug development against antibiotic-resistant strains of *Mycobacterium tuberculosis* by Kavanagh et al., who combined X-ray crystallography, native MS, and ITC to discover a new hit compound, **25a**, with an affinity of 15 nM for CYP121.[166] The authors began their research using X-ray screening of low molecular weight fragments. These fragments are small

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

molecules, and their initial binding to the recombinant protein was relatively weak in the mM affinity range. Through a process of merging (linking) these fragments, they synthesized a new series of 5-amidopyrazole derivatives. To assess the improved binding affinity of these new compounds, they tested them in a dose-dependent manner using native MS. This method allowed them to precisely determine the 5-amidopyrazoles binding affinity by analyzing the dose dependency of the concentration of MS complex generated ultimately helping them identify the most effective antibacterial agent.[166] With a similar approach, Vaaltyn et al. identified a small cohort of molecular fragments whose binding to the TPR2AB region of the target, HOP inhibits the complex formation between HOP and the HSP90 C-terminus.[162] Their experimental observations lead to the discovery of losartan as a weak but consistent HSP90 inhibitor.[162]

4.2. Cross-Linking MS Proteomics. Cross-linking (XL)-MS is an MS technique used to investigate protein−ligand interactions, including fragment binding in FBDD. XL-MS involves the use of chemical cross-linkers that covalently bind specific amino acid residues that are spatially close in a protein or between proteins.[167] When applied to FBDD, XL-MS can provide valuable structural information about how a fragment interacts with its protein target, including identification of the binding site, conformational changes upon fragment binding, and insights into the overall architecture of the protein or protein complex.[168] XL-MS involves the use of chemical cross-linkers that react with specific amino acid side chains, such as lysine, cysteine, or serine, to covalently link them if they are within a certain distance from one another (typically 10−30 Å).[169] When a fragment binds to a protein, it can induce conformational changes, stabilize certain regions of the protein, or bring specific amino acid residues into proximity. The resulting cross-linked products are digested into peptides, analyzed by MS proteomics, and mapped onto the known or predicted protein structure.[170] XL-MS works by utilizing cross-linkers that are bifunctional chemical reagents capable of covalently binding two amino acid residues within a defined distance range. The cross-linking process is driven by the proximity of the reactive residues when a protein assumes a particular conformation, making it an effective tool for capturing dynamic interactions.[171] Cross-linkers come in various forms, including the homobifunctional cross-linkers, i.e., reagents that contain two reactive groups that target the same type of residue (e.g., two lysines), heterobifunctional cross-linkers, which have two different reactive groups that can target different types of residues (e.g., lysine and cysteine), and cleavable cross-linkers, which are designed to be cleaved under specific conditions (e.g., reduction, UV light), making it easier to identify the cross-linked peptides in the MS.[172] An example of XL-MS in FBDD is represented by the hit identification process of innovative thymidylate synthase (TS) destabilizers with anticancer activity described by Costantino et al.[173] In their drug discovery pipeline aimed at finding innovative allosteric thymidylate synthase (TS) inhibitors, the authors used MALDI-TOF and LC-ESI-MS to investigate the affinity of a TS mutant enriched with cysteines (TS C195S−Y202C variant) vs a library of thiols and disulfides. More than 80 reactive compounds were tested, and this strategy allowed to identify two bioisosterically analogues of the bound fragments, as hit compounds, namely **B12** and **B26**, were optimized and used to develop the lead compound **E7**.[173] The workflow is illustrated in Figure 10.

With a similar approach, Fan et al. investigated the acylation pocket of the transcriptional enhancer-associated domain 2 (TEAD2) with a library of small ligands similar to the previously identified flufenamic acid, containing an acrylamide warhead able to be attached by Cys380.[174] MS proteomics displacement assays lead to the identification of the hit compound **MYF-01−037** as highly affine to the acylation pocket, which led to the identification of the lead compound **MYF-03−069**, active in cells in the nanomolar range with an optimal PK for oral administration.[174] Despite still being a bit underused compared to other techniques, the integration of FBDD and cross-linking proteomics with MS provides a technically advanced platform for mapping ligand−protein interactions with high sensitivity and resolution. In contrast to traditional FBDD methods, such as NMR, X-ray crystallography, or SPR/GCI, which require purified proteins and large sample quantities and often overlook low-affinity or transient binders, MS-based proteomics enables fragment screening directly in complex lysates or intact cells. This allows the identification of interactions in the native proteomic context. XL-MS enhances this further by covalently stabilizing weak fragment-protein complexes through reactive linkers, facilitating subsequent MS-based localization of binding sites at the peptide or residue level. These combined strategies provide a quantitative, proteome-wide view of ligand engagement, extend chemical coverage to traditionally "undruggable" targets, and accelerate hit validation and MoA studies compared with conventional structure-based or biophysical screening techniques.

5. MS PROTEOMICS FOR LEAD OPTIMIZATION: MOA AND OFF-TARGET ACTIVITY

MS proteomics contributes to the refinement of drug candidates by helping define the SAR, which correlates changes in the chemical structure of compounds with their biological activity.[21] MS proteomics allows researchers to rapidly assess how modifications to a compound affect its ability to bind to the target protein, its efficacy, and its specificity.[21] This accelerates the identification of lead compounds with the best activity and lowest toxicity.[21] Once validated, hits undergo lead optimization, where the chemical properties of these compounds are refined to enhance their PK and PD characteristics.[175] MS proteomics continues to play a vital role in this process and is often exploited during these phases as label-free MS proteomics or TMT/iTRAQ labeled to gain insight into the general mechanisms of action identification and as hydrogen−deuterium exchange MS (HDX-MS) to characterize the conformation of a target protein that changes upon binding a ligand.[176] Also, coimmunoprecipitation followed by MS (Co-IP-MS) is another technique that plays a central role in hit and lead development up to the preclinical stages. Indeed, it allows us to quantify the impact of a drug candidate on protein complex formation; thus, it represents a gold standard in the development of PPI inhibitors.[177] During the stages of drug discovery, time-resolved experiments (i.e., comparing different time points after compound administration) or dose-resolved experiments (i.e., comparing different compound doses) are often used to elucidate different biochemical pathways involved in their mechanisms of action.[178]

Recently, Lee et al. published an article on time-dependent changes in the proteomes of cancer cells as they expand. These changes, particularly in cell division, adhesion, and metabolism, begin as early as the second day of culture.[178] The study uses label-free MS proteomics to screen a library of natural products

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

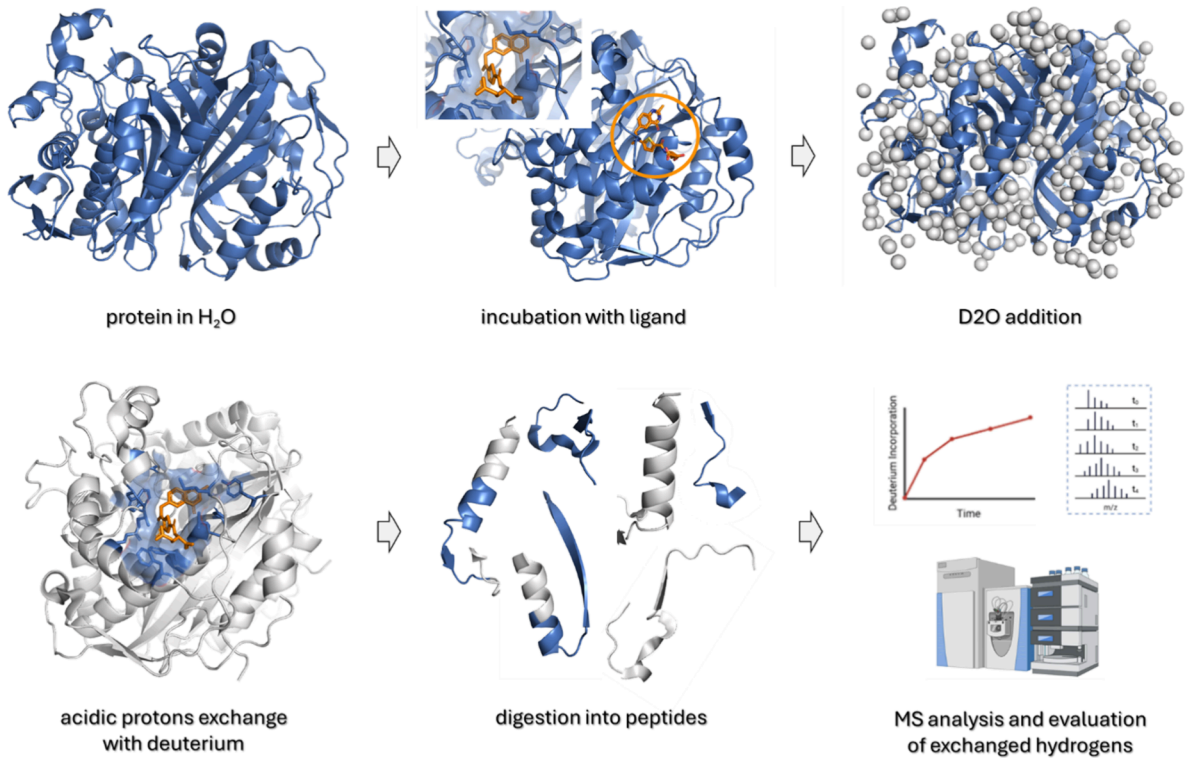


Figure 11. HDX experiment scheme. After ligand incubation, D2O is added for an established timelapse before protein digestion. And acidic pH quenches the reaction and unfolds the protein and analyzed with HRMS. HR instruments like TOF or Orbitrap can discriminate small hydrogen mass shifts, which makes it easy to identify which peptides and single amino acid underwent major deuterium incorporation. In the present figure, the protein with its hydrogens is represented in blue cartoon, D_2O is represented as gray spheres and protein residues that have incorporated deuterium as gray cartoons. Thymidylate synthase in complex with methotrexate (PDB:5X66) was used to generate this image on BioRender.

and identifies four pentacyclic triterpenes that are effective against cancer cells under prolonged growth, and it represents a good example of time-resolved exploitation of MS proteomics for broad-range MoA investigation.[179] The authors discovered that some tested compounds selectively target highly overgrown cells, offering potential avenues for treating drug-resistant cancers, and they demonstrated that caution should be taken in the evaluation of anticancer drugs using differently old tumor cells, introducing the concept that some compounds are stage-specific.[179] Last year, Chang et al. explored how lysine deacetylase inhibitors (KDAC inhibitors or KDACis) affect cellular processes at the proteome, acetylome, and phospho-proteome levels through dose-dependent proteomics.[180] The research systematically profiles the impact of 21 KDAC inhibitors (KDACis), revealing that many acetylation sites respond to these drugs in a dose-dependent manner, and concluded that KDACis, especially HDAC1/2/3 and HDAC6 inhibitors, share many common acetylation targets. Through this study, the authors offer a valuable resource for investigating drug mechanisms in different subtypes of cancers and propose a new MoA for panobinostat, which also induces cytosolic accumulation of the paralogous lysine acetyltransferase KATs p300 and of the CREB-binding protein CBP.[180] With a similar purpose, through the use of Functional Identification of Target by Expression Proteomics (FITExP), which consists of a differential MS-based proteomics analysis of cells treated with inhibitor(s) vs control cells, Saei and colleagues deciphered the specific target(s) and MoA or a library of 56 compounds belonging to 19 different classes of chemical compounds on A549 human lung adenocarcinoma cells. This experiment resulted in the publication of ProTargetMiner (http://protargetminer.genexplain.com), a curated, publicly available, expandable proteome signature library that represents a valuable tool in drug discovery.[181] Ruprecht et al., by combining middle throughput label-free MS proteomics and biostatistics, characterized the molecular mechanisms of action of 50 drugs employed to treat lung cancer, including HDAC inhibitors, TK inhibitors, CDK inhibitors, and MAP2K1/2-MAPK1/3 inhibitors.[182] Through PCA analysis and aggregation analysis of the differentially expressed proteins for each cell treatment, the authors mapped the biochemical networks modified by each drug class and defined their biochemical fingerprint around the specific target.[182] MS proteomic data also led the team to discover that inhibition of mitochondrial function is an off-target mechanism of the MAP2K1/2 inhibitor **PD184352**, and that the ALK inhibitor ceritinib modulates autophagy.[182] Still using MS proteomics, El-Baba et al. identified 197 proteins affected by **ST1926** administration, a synthetic retinoid used to treat glioblastoma multiforme (GBM), which targets POLA1, revealing its broad impact on key cancer pathways.[183] The investigators discovered that **ST1926** treatment significantly reduces POLA1 levels, leading to cell cycle arrest, DNA damage, and apoptosis in GBM cells.[183] An innovative, high-throughput platform for the cellular screening of drug phenotypes by dose-resolved expression proteomics has lately been introduced by Eckert et al.[184] The DecryptE platform enhanced the understanding of the MoA of drugs, focusing on the effects of 144 drugs and research compounds across 8000 proteins, by generating over 1 million dose−response curves of protein expression.[184] The authors discovered that the proteome modulation is strictly dependent not only on their MoA but also on the concentration used, i.e., some drugs, such as HDAC

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

inhibitors, affect numerous proteins, while others, like certain kinase inhibitors, cause more specific changes.[184] All of these case studies exemplify how **MS proteomics**, typically a label-free pipeline to accommodate more conditions, has been utilized to investigate the **MoA** of one or more compounds.

5.1. Hydrogen−Deuterium Exchange MS Proteomics. Hydrogen−deuterium exchange MS (HDX-MS) is a chemoproteomics technique that plays a central role in drug discovery by providing insights into protein structure, dynamics, and interactions with potential therapeutic agents, either small molecules or biopharmaceuticals.[185] This technique allows investigators to characterize how proteins change conformation upon binding with ligands, revealing the main binding sites, strength, and the effects of various molecules on protein stability.[186] By measuring the rate of deuterium incorporation into protein structures, HDX-MS can reveal the dynamics of protein−drug interactions, enabling the identification of allosteric or secondary binding sites, and help in the phase of drug optimization.[187] Briefly, the protein or protein mixture previously purified is incubated with the investigated ligand(s) at different concentrations in the presence of D_2O, to allow the protein's mobile protons to exchange with deuterium.[187] It is assumed that a good ligand can screen the protein's protons close to its binding site and engage in protein−ligand interaction.[187] The protein is then quickly digested into peptides, and the rate of deuterium incorporation is then measured by high-resolution MS, as reported in Figure 11.[188]

The ability to analyze also large complexes and challenging targets, such as membrane proteins and disordered regions, makes HDX-MS, particularly valuable in the development of antibodies.[189] Furthermore, advancements in tools and data processing for HDX-MS experimental interpretation have enhanced the efficiency and accuracy of HDX-MS.[190] Despite being like the technology of the previously described PAL, HDX-MS has the advantage of needing no chemistry or probe design, making it ideal to investigate lead compounds with their original structure and no modifications.[191] Also, correlating the HDX-MS profile of large numbers of ligands during the hit optimization phase with their functional outputs enables the development of SARs and the delineation of hit scaffold classes based on functional selectivity.[192] Ow et al. exploited this technique to explore the novel binding site of garadacimab (**CSL312**) on the Hageman factor (FXII), an essential component in the intrinsic coagulation cascade and a therapeutic target for the prophylactic treatment of hereditary angioedema.[193] The authors revealed an additional "side pocket," complementarity-determining regions (CDRs) close to the antibody main recognition site, which can be exploited as a potential paratope for the development of new, more selective FXII inhibitors.[193] Similarly, Woods et al. investigated the tyrosine phosphatase 1B (PTP1B) structure, a validated therapeutic target for obesity, diabetes, and certain types of solid cancer.[194] In particular, by HDX-MS, they mapped a particular rigid backbone sequence involved in the binding with the active-site inhibitor **TCS401**, which may serve for orthosteric compound design and optimization. Also, they mapped allosteric binding sites of PTP1B for small-molecule inhibitors.[194] Another example of HDX-MS-driven compound optimization is represented by the study of Song et al. on SARS-CoV-2 3-chymotrypsin-like proteases (3CLpro or M^{pro}), through which an allosterically regulated dimerization of this protein that affects its activity was characterized. Based on the HDX-MS data, the authors propose the scaffold of gastrodenol as a lead compound with dissociative activity on 3CPLpro for further investigation.[194] In a similar article about another viral target, the activity of lofanarib in inhibiting fusion protein was investigated to treat the early stages of respiratory syncytial virus infection. The time-dependent HDX-MS assay demonstrated that lofanarib interacts with 10 key residues of the fusion protein (EC_{50} 57.7 ± 15.4 nM in HEp-2 cells) to block its conformational exchange that triggers viral fusion and puts the basis to develop more specific inhibitors from this lead compound.[194] Overall, the literature demonstrates that HDX-MS provides several key advantages over a wide range of conventional techniques used to identify ligand-binding sites, including X-ray crystallography, cryoelectron microscopy (cryo-EM), nuclear magnetic resonance (NMR), surface plasmon resonance (SPR), and differential scanning fluorimetry (DSF). Unlike crystallography or cryo-EM, which require stable, purified, and often crystallizable proteins, HDX-MS operates in solution under near-physiological conditions and can probe dynamic, flexible, or disordered regions that are invisible in static structural models. Compared with NMR, HDX-MS is faster, is compatible with larger proteins and complexes, and requires much lower sample quantities. While SPR and DSF report on global binding events or thermal stability changes, HDX-MS uniquely provides site-resolved information by detecting changes in the backbone amide hydrogen exchange upon ligand binding. This enables the precise localization of binding regions, the identification of allosteric effects, and the characterization of conformational dynamics across the entire protein. Overall, HDX-MS delivers a powerful combination of structural sensitivity, throughput, and applicability, making it superior for mapping ligand-binding sites in complex protein systems.

5.2. Co-Immunoprecipitation Followed by MS (Co-IP-MS). Co-IP-MS is an MS proteomics technique used to study protein−protein interactions (PPIs) by combining Co-IP with MS proteomics.[195] The main purposes of Co-IP-MS are the mapping of protein interaction networks (PINs) within a cell or biological fluid in a quali-quantitative network, often before and after administration of a specific compound, or in the presence and absence of a metabolic disease, which makes it a gold standard technique to investigate the efficacy of protein:protein inhibitor candidates.[196] The considered cells or tissues, properly treated, are broken open under specific conditions that preserve protein interactions, e.g., mechanical lysis over chemical detergents.[197] A specific antibody targeting the protein of interest, known as the "bait" protein, is then introduced in the system, which is usually bound to a support, such as Protein A/G beads, through a Schiff's base.[198] This helps isolate the bait protein along with any associated binding partners ("prey" proteins).[199] After incubation, unbound proteins and other cellular components are washed away with a specific buffered solution, leaving only the bait protein and its prey attached to the beads.[200] Typically, low-salt or high-salt conditions are used to help disrupt specific interactions, along with detergents like Triton X-100 or NP-40 and reducing agents. The isolated protein complex is then subjected to a standard bottom-up MS proteomics workflow, and from the composition of the quantified peptides, it is possible to determine which proteins interact with the bait protein under determined conditions.[201] Sometimes, cell lysate PPIs can be stabilized through chemical cross-linking, e.g., DSP (dithiobis-succinimidyl propionate) to allow the identification of even short-lived complexes.[202] In this case, the covalent modification needs to be reverted with mild reducing agents like TCEP before MS analysis.[203] Figure 12

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

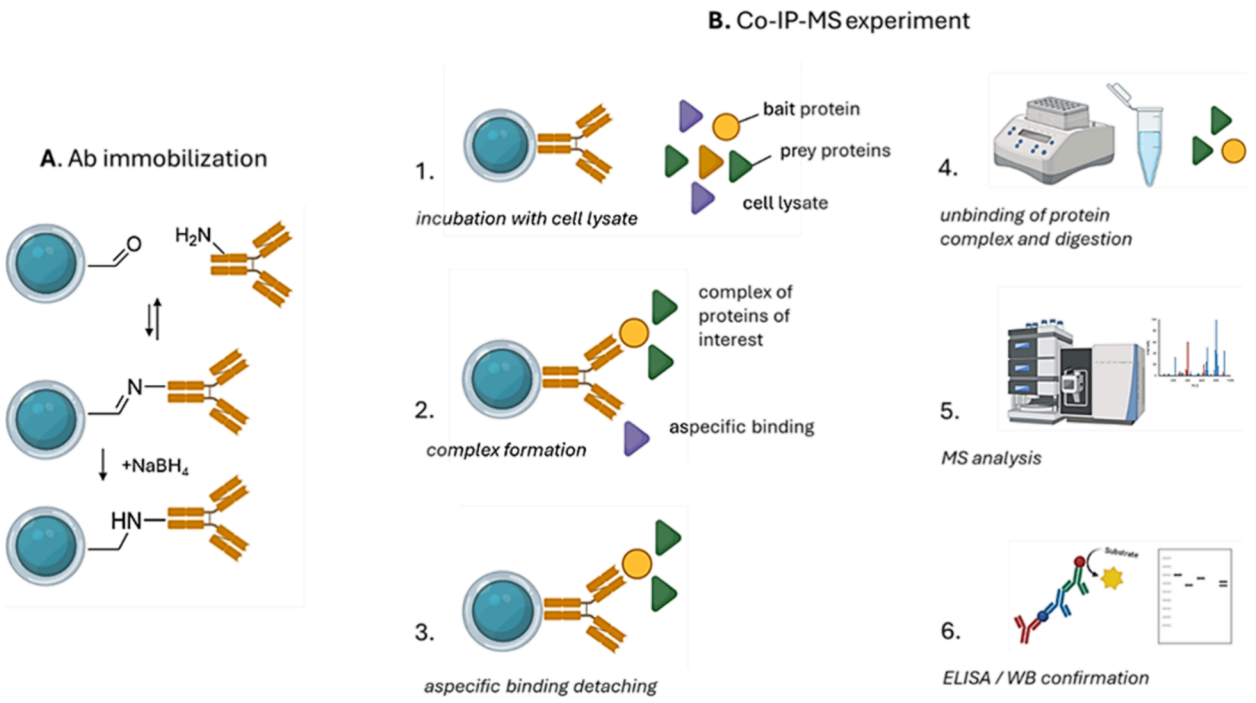


Figure 12. Schematization of a Co-immunoprecipitation-MS experiment. (A) Appropriate antibody toward bait protein is incubated with agarose beads functionalized with aldehyde-carrying ketones, which reversibly react with lysins of the antibodies by giving a Schiff's base. The formed imine is reduced by $NaBH_4$ to stabilize the bond. (B) Agarose beads properly functionalized with antibodies are incubated with cell lysate (1) to form the complex between protein of interest and prey proteins, including nonspecific binding (2). Nonspecific bound proteins are detached by varying salt or pH concentration (3), then the formed complex is detached and digested with a standard bottom-up MS procedure (4). Obtained peptides are analyzed with LC−MS/MS (5), and an orthogonal analysis with either ELISA, WB, or other immunohistochemical techniques is performed (6). The image was created with BiorRender.

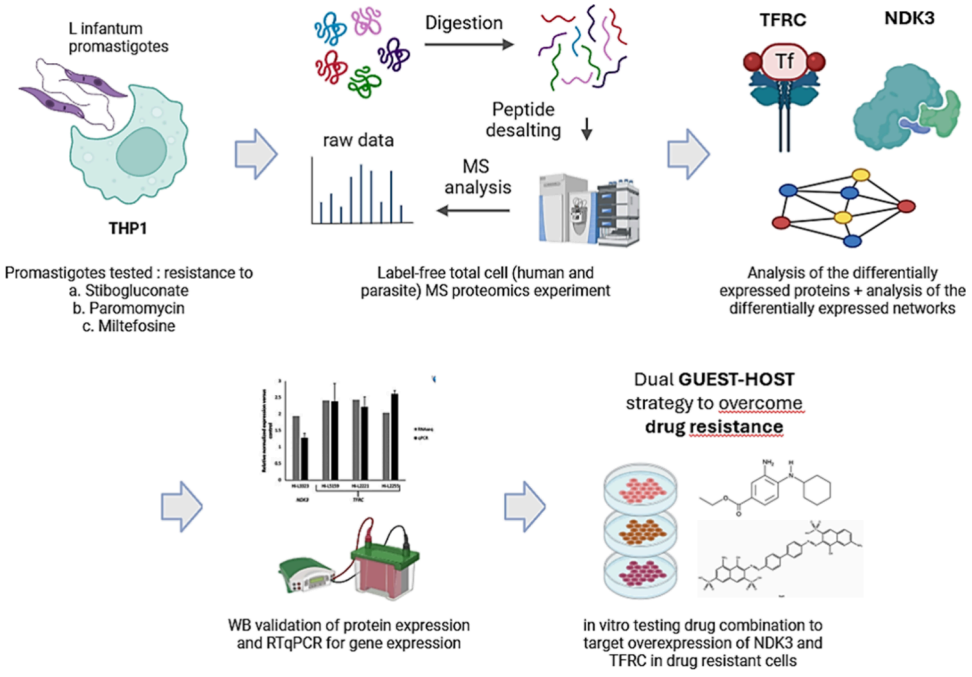


Figure 13. Examples of MS proteomics used for the characterization of drug resistance patterns in *L. infantum*. Tagliazucchi and colleagues exposed three drug-resistant *L. infantum* strains from clinical isolates to THP-1 macrophages (human). The system in acute infection was submitted to total MS proteomics. Human proteins were distinguished from the parasitic ones and quantified. NDK3 and TFRC were marked as overexpressed in miltefosine and paromomycin strains with respect to nonresistant *L. infantum* strains. The authors validate their findings through WB and quantitation of NDK3 and TFRC transcripts by RT-qPCR. As the authors' hypothesis includes NDK3 and TFRC being upregulated in host cells by the guest for iron and nucleoside metabolic supply, contributing to drug resistance, ongoing experiments include the administration of NDK3 and TFRC inhibitors in a perspective to also reduce the resistance index.[214] The figure was realized with BioRender and is based on the content of the referenced publication.

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

reports a schematic workflow for bead immobilization (Figure 13A) and affinity capture (Figure 13B). This was the case of Liu et al., who demonstrated the antinflammatory activity of isoliquiritigenin (ISL) by combining data from RT-qPCR and Co-IP-MS experiments and proving that ISL covalently binds to MS2 and disrupts the formation of the LPS/MD2/toll-like receptor 4 complex, thus resulting in significantly alleviated lung injury in LPS-induced mice.[204] Following a similar experimental pipeline, Zhihuai et al. investigated the pharmacological activity of liquidambaric acid (i.e., betulonic acid, or HY-N1451-31) in the disruption of STAMBPL1/NRF2 complex, which would promote DUB activity and eliminate ubiquitin molecules attached to NRF2, thus protecting it from proteasome-mediated degradation.[205] Several years before, Gorini et al. succeeded in Co-IP experiments with dynamin-1 coprecipitating BKCa channels from native mouse brain preparations.[206] This interaction was reproducibly detected by Western blotting, confirming the association between these two proteins and suggesting a specific and stable interaction under physiological conditions, pointing toward a functional relationship between dynamin-1 and BKCa channel complexes in neuronal tissue.[206] In their works, Hossain and colleagues also integrated Co-IP-MS in their innovative mathematical-based drug discovery decision-tree workflow, which combines a novel MS proteomics assay and a mathematical PK/PD model specifically designed for covalent drugs. Indeed, the model was validated for two approved drugs and proposed for a class of promising ALS drug candidates targeting SOD1.[206]

Overall, the preference of MS-guided Co-IP offers distinct advantages within the drug discovery pipeline, particularly when studying target engagement and MoA, compared with other common approaches such as surface plasmon resonance (SPR), cellular thermal shift assays (CETSA), or differential scanning fluorimetry (DSF). Unlike biophysical assays such as SPR or DSF, which rely on purified recombinant proteins and primarily measure direct ligand−target binding in vitro, Co-IP-MS provides information within a cellular or native physiological context. This allows the identification of drug−target interactions as they occur in living systems, preserving their natural conformation. As a result, Co-IP-MS offers a more realistic view of target engagement, especially for compounds acting within multiprotein assemblies or signaling complexes. Another major advantage of Co-IP-MS is its ability to capture both direct and indirect effects of compound binding. While thermal stability-based methods such as CETSA detect changes in protein folding upon ligand binding, they cannot directly reveal how those interactions affect the broader interactome. In contrast, Co-IP-MS can detect changes in the composition or abundance of protein complexes following drug treatment, highlighting secondary binding partners or interaction networks that are stabilized, disrupted, or newly formed. This is particularly valuable for drugs that act allosterically, modulate scaffolding proteins. Furthermore, Co-IP-MS offers unbiased identification of interacting partners without prior assumptions about the molecular mechanism, making it a powerful discovery tool for elucidating mechanisms of action of novel compounds or phenotypic hits. When compared with label-based affinity methods such as DARTS, Co-IP-MS has the advantage of relying on endogenous PPIs rather than on chemical derivatization of the ligand, which can sometimes alter binding properties.

6. MS PROTEOMICS FOR THE CHARACTERIZATION OF DRUG RESISTANCE MECHANISMS AND NEW DRUG COMBINATIONS

MS proteomics is increasingly recognized as a powerful and accurate tool for characterizing drug resistance in various diseases, particularly cancer, bacterial diseases, and parasitic infections.[207] By analyzing the proteomic profiles of resistant and susceptible cell lines, tissues, or clinical isolates, or by recreating in vitro certain conditions to induce a resistant phenotype, it is possible to identify specific proteins and map differentially regulated pathways that associated with or cause resistance mechanisms.[16,208] This allows to understand key proteins responsible for therapeutic failure, to improve compound optimization, and to study possible drug combination to escape or delay the onset of resistance.[209,210] For instance, MS Proteomics evidenced in HER2+ breast cancer 3D cell models (approximately 15−30% exhibit HER2 amplification, triggering enhanced signaling through various pathways, notably the PI3K/Akt and MAPK pathways, which promote cell proliferation and survival), the MS proteomic characterization revealed the involvement of mitochondrial complex I in acquired resistance to trastuzumab, highlighting how metabolic pathway modulation can influence treatment outcomes.[211] This could provide personalized therapeutic strategies, increasing the efficacy of drug treatment while reducing its side effects.[211] HDAC inhibitors represent important drugs for cancer therapy, but the unclear resistance mechanisms greatly limit their clinical applications.[212] The contribution of Hao et al., who exploits the (TMT)-based MS proteomics approach on HDACi-sensitive and -resistant cell lines, identifies a set of proteins responsible for HDAC-resistant phenotype (including MRTO4, PES1, WDR74, and NOP16), and proposes a combination strategy that targets both HDAC and one of the main causes of HDCA mutations to delay the onset of drug resistance during chemotherapy.[212]

Additionally, label-free MS proteomics methods have been employed to analyze the global proteome of drug-resistant bacteria, such as *Acinetobacter baumannii,* revealing changes in key virulence proteins that contribute to antibiotic resistance, such as the membrane efflux pumps AdeB and AdeDE, which actively expel a wide range of antibiotics, contributing to its multidrug-resistant (MDR) phenotype. This approach allows for the identification of proteins that may serve as biomarkers for resistance or potential therapeutic targets.[15] A very similar pipeline was used by Abril et al. to do a global shotgun proteomics characterization for pathogenic *Listeria* species, particularly *L. monocytogenes*, and describe their phenotypes of antibiotic resistance.[213] The study identified proteins associated with beta-lactamase activity, penicillin-binding proteins, and other resistance mechanisms to antibiotics such as gentamicin, kanamycin, fosfomycin, and tetracycline. Proteins like internalins (InlA, InlB), which are involved in bacterial entry into host cells, and peptidases that aid in host invasion were highlighted, were mapped as key player of the phenotype resistance, and the Lactococcin 972 family was proposed as new target for a drug combination approach.[213] MS Proteomics is also employed to gain insights into parasitic chemoresistance mechanisms. It is the case, for instance, of Saboia-Vahia et al., that compared the protein expression profiles of a miltefosine-resistant *L. infantum* strain with that of the wild-type (WT) strain.[214] They experimentally induced resistance in the parasites through six months of continuous exposure to increasing doses of

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

Table 6. Summary of MS Proteomics in Drug-Resistant Studies, with Their Related Outcomes and the Potential Outcomes within Medicinal Chemistry Programs

general concept	actual outcomes	potential achievements
use of resistant vs sensitive models (cell lines, tissues, bacterial/parasitic/viral isolates) to compare global proteomic profiles	identification of key differentially expressed proteins	discovery of novel therapeutic targets
	mapping of deregulated pathways	identification of existing inhibitors to block or modulate dysregulated pathways
	functional validation, *i.e.*, metadata confirmation, orthogonal validation (e.g., WB, ELISA)	experimental administration and testing molecules on resistant cells/isolates
induction of resistance in vitro through prolonged drug exposure and subsequent proteomic comparison	generation of resistant strains/models	prediction of clinical resistance evolution
	discovery of resistance mechanisms	development of early intervention therapies
		biomarker discovery for resistance monitoring
host−pathogen interaction proteomics to identify guest-induced changes in host cells	separation and quantification of host vs parasite proteome	targeting host cell pathways hijacked by parasites
	identification of guest-induced host reprogramming to sustain guest survival	dual targeting (host + pathogen) strategies
		potential reduction of resistance development

miltefosine and identified proteins like ABC transporters, sterol biosynthesis enzymes and phospholipid transporting ATP-ases as potentially involved in the resistant phenotype.[214] Similar results but on THP-1 infecting promastigotes were obtained by Tagliazucchi et al., who identified an overexpression of energetic (TCA) metabolism, fatty acid biosynthetic proteins and membrane lipid metabolism in miltefosine and antimonial resistant *L. infantum* clinical isolates (Figure 13).[215]

Only recent literature discloses the potentialities of MS proteomics and kinomics in identifying effective drug combinations compared with more conventional molecular and cellular techniques such as genomics, transcriptomics, WB, or immunoassays. Unlike genomic or transcriptomic profiling, which reveals only potential regulatory changes (i.e., on a genic base), MS proteomics reveals true functional effectors of drug response. This allows the overall characterization of adaptive signaling rewiring, compensatory pathway activation, and metabolic remodeling that underlie resistance phenotypes. In contrast to targeted techniques like WB, ELISA, or flow cytometry, which are limited to predefined proteins and require specific antibodies, MS proteomics provides a system-wide coverage without prior knowledge of targets (i.e., off-target information), and phosphoproteomics can uncover alterations in kinase signaling networks driving resistance. Moreover, by integrating whole system proteomic with phosphoproteomic data, MS enables rational design of synergistic drug combinations that cotarget reactivated or compensatory pathways, an insight often missed by single-*omic* or targeted assays. Overall, the purposes of MS proteomics in drug-resistant studies can be summarized in Table 6.

7. MS PROTEOMICS IN TOXICOLOGY AND ADVERSE OUTCOME PATHWAYS IDENTIFICATION

One of the modern approaches that has changed toxicology and safety assessment in drug discovery is the adverse outcome pathway (AOP) framework.[216] AOPs provide a structured way of connecting molecular-level interactions caused by chemicals (including drugs) to adverse health outcomes.[217] AOPs facilitate the prediction of potential hazards at the early stages of drug development, helping to reduce late-stage failures in clinical trials due to safety concerns.[218,219] An AOP is a conceptual framework that describes the progression from a molecular-initiating event (MIE), which is the first interaction between a drug or other chemical and a biological target, through a series of intermediate steps, ultimately leading to an adverse outcome at the individual or population level.[220,221] These steps form a cause-effect chain, linking biological mechanisms to phenotypical effects that can be measured and monitored.[222] A typical AOP consists of the following elements: i. an MIE), i.e., a direct interaction between a chemical and a biological molecule (e.g., a drug binding to a receptor or an enzyme); ii. the key events (KEs), which encompass the biologically meaningful changes that occur at the cellular, tissue, or organ level following the MIE (these are measurable processes such as changes in gene expression, proteomic changes, cellular stress, or inflammation), and finally, iii. the adverse outcome (AO), that is, the ultimate harmful biological effect, which could be organ toxicity, impaired reproduction, developmental defects, or death.[220,223] The incorporation of AOPs into drug discovery offers multiple advantages. Mainly, AOPs help in understanding how a drug's molecular interactions lead to toxicity, providing mechanistic insights into potential safety issues early in the discovery phase.[224] This allows to design compounds that avoid toxic pathways. AOP studies also help in the reduction in animal testing. Since AOPs are based on mechanistic understanding, they can be used in combination with alternative testing methods, such as in vitro models and computational simulations.[225] AOPs facilitate the prediction of potential toxicological outcomes based on early molecular events.[2] By screening compounds for interactions with key molecular targets, it is possible to predict whether a drug candidate is likely to cause an adverse outcome before it progresses into animal or human testing.[2] The AOP framework is compatible with data from various sources, including in vitro assays, omics technologies (e.g., genomics, proteomics), and computational models.[226] This integration enables a more comprehensive assessment of a drug's safety, incorporating different layers of biological data into the decision-making process.[216] For instance, Bakker et al. recently investigated the cross-integration of transcriptomic and proteomic data to improve the AOP context by focusing on *Folsomia candida*, a soil invertebrate, exposed to imidacloprid, which overstimulates the nicotinic acetylcholine receptor (nAChR), leading to neuronal disruption.[227] By collecting data at 12 h intervals up to 72 h, the researchers identified the most significant molecular changes at the 48-h mark. This time point was crucial for observing differential protein expression patterns, particularly in pathways related to neurodegeneration and synaptic signaling.[227] Another

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

case study of MS proteomics in AOPs characterization was well described by Liu et al., who discusses the effects of polystyrene nanoplastics on *Daphnia pulex*, a key species in aquatic ecosystems.[228] The group explored how nanoplastics, particularly spherical polystyrene, affect reproduction and growth in *Daphnia*. Using MS proteomics, the researchers identified changes in protein expression related to oxidative stress, energy metabolism, signaling pathways (such as mTOR and Jak-STAT), and cuticle development. This outcome allowed for refinement of the AOP for nAChR activation by adding new key events, such as GABA signaling and mitochondrial dysfunction. These insights also help identify biomarkers for imidacloprid exposure, contributing to more efficient environmental risk assessment.[228] Table 7 depicts the potentialities and limitations of MS proteomics studies of AOP events.

8. MS PROTEOMICS FOR BIOMARKER DISCOVERY, VERIFICATION, AND VALIDATION

Biomarkers are defined as biological molecules that indicate physiological or pathological states, playing a vital role in disease diagnosis, prognosis, and therapeutic monitoring.[229] MS proteomics has become a powerful tool also in accelerating the discovery pipeline and diminishing the costs of biomarker studies, due to its sensitivity and ability to analyze complex biological samples.[230] Biomarker discovery using MS-based proteomics involves a multiphase process that can be divided in three main phases: discovery, verification, and validation.[231] Each phase leverages different high-throughput MS technologies to explore the proteomic landscape and identify potential biomarkers, combined with orthogonal techniques.[175] In the discovery phase, MS-based proteomics is usually used to compare protein expression in biological samples (e.g., blood, tissue, and urine) from diseased and healthy individuals. This phase is performed on disease groups compared to control groups and often involve a validation of the corresponding gene expression levels.[232] Following discovery, candidate biomarkers (<10) undergo verification stage. Targeted MS approaches such as selected reaction monitoring (SRM) or parallel reaction monitoring (PRM) have extensively been used for this purpose and have replaced older optical techniques and immunoassays.[32,233] These techniques allow for the quantification of specific proteins with high precision and sensitivity but need to be combined with orthogonal assays including immunohistochemistry (ELISA, WB, etc) and kinetic assays, with the purpose of assessing sensitivity.[234] The final stage, validation, involves testing biomarkers in larger, independent cohorts to assess their clinical utility.[235] This phase often employs multiplexed immunohistochemistry assays, or MS to quantify several biomarkers simultaneously, ensuring the robustness of findings.[236] Finally, when biomarker is validated on a larger scale (n > 1000 patients), data can be submitted to national and supranational health authorities for revision.[237] A general scheme of the biomarker discovery process with MS proteomics integration is illustrated in Figure 14.

A notable platform used for early phase biomarkers investigation was The Clinical Proteomic Tumor Analysis Consortium (CPTAC), who, led by the National Cancer Institute (NCI), has integrated proteomic data with genomic and transcriptomic information from The Cancer Genome Atlas (TCGA), by giving birth to the so-called proteogenomic field.[238] The consortium uses advanced MS proteomics and other proteomic technologies to produce deep proteomic data sets linked to genomic data from projects like TCGA, facilitating a

Table 7. Potentiality of MS Proteomics Studies of AOP Events and Their Limitations in the Context of Drug Discovery

potential	limitations
high sensitivity and specificity: MS proteomics can detect and quantify low-abundance proteins, allowing the identification of early molecular changes in response to toxicants	**sample complexity and dynamic range**: biological samples often contain proteins with a wide range of abundances. High-abundance proteins (HAPs), like albumin in serum, can mask the detection of low-abundance proteins (LAPs), potentially overlooking critical components of AOPs
global proteome coverage: advanced MS techniques, such as shotgun proteomics, allow for the broad profiling of proteins, facilitating the mapping of entire signaling pathways involved in AOPs	**data analysis challenges**: the vast amount of data generated requires sophisticated bioinformatics tools and expertise. Misinterpretation can occur due to issues like peptide degeneracy and incomplete databases.
quantitative analysis: label-free and label-based quantification methods enable the measurement of protein expression changes over time, aiding in the temporal mapping of AOPs.	**ion suppression effects**: co-eluting compounds from complex matrices can suppress ionization efficiency, leading to inaccurate quantification and potential false negatives. This can be overcome only if standard reference compounds are available.
detection of PTMs: MS is adept at identifying PTMs, such as phosphorylation and oxidation, which are crucial in signaling events leading to AOPs	**incomplete PTM coverage**: some PTMs are labile or occur at low stoichiometry, making them difficult to detect and quantify accurately.

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

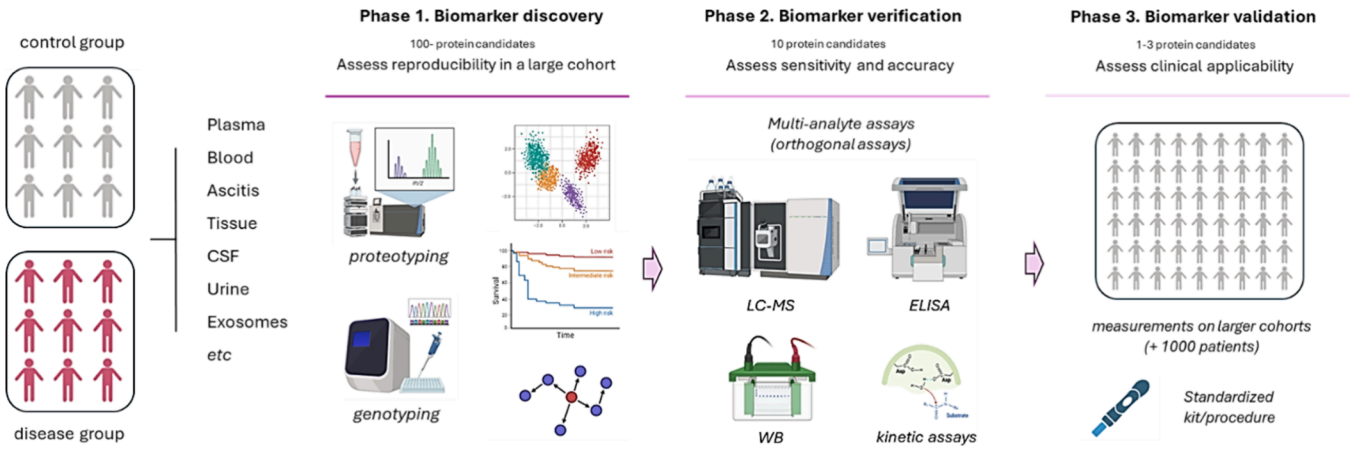


Figure 14. General protein biomarker discovery and validation pipeline. Phase 1 starts with the hypothesis generation of the involvement of a group of proteins/pathways in the progression of a disease. The total proteome from different biological samples (e.g., blood, urine, CSF) of the disease group is measured and compared to that of the control group. Multiparametric analysis (i.e., ANOVA) is then applied to verify which proteins/networks are significantly differentially regulated in the disease group. If possible, a primary validation of proteins is performed on their transcripts to ensure that the protein's differential expression has a genetic basis. Phase 1 assesses reproducibility in large cohorts and represents an explorative stage. Phase 2 restricts protein candidates $n < 10$, and consists of multianalyte assays with different techniques, including targeted MS, immunohistochemical assays or enzymatic/colorimetric detection. The aim of phase 2 is to assess sensitivity with different techniques and their accuracy. Phase 3, i.e. biomarker validation, is performed on larger cohorts (>1000 patients), in randomized, double-blind trials to assess clinical applicability. In this phase, the technique of standard measurement is chosen, which is often robust, poorly affected by false negatives, and affordable on a large scale (often kits are produced if the procedure is noninvasive). At this point, the biomarker data are submitted to health authorities for revision and qualification. The image was created with BioRender.

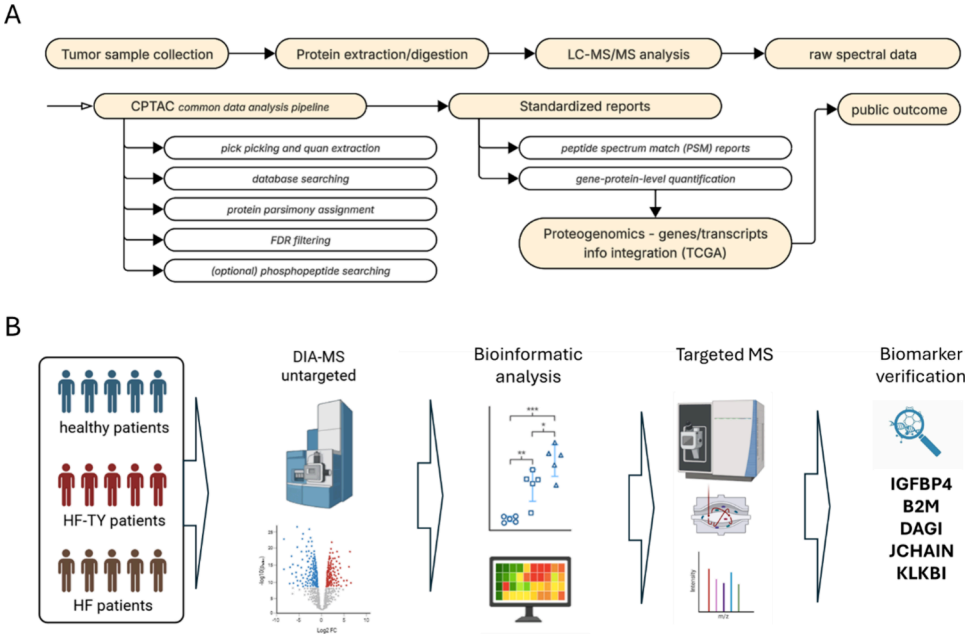


Figure 15. (A) Schematic representation of the CPTAC operative workflow, from sample collection to data repository. (B) Workflow representation of the whole serum MS proteomics experiment by Lan et al. to investigate potential biomarkers for heart failure patients with phlegm-blood stasis syndrome. A set of 84 DEPs with a DIA-untargeted experiment, 37 of which were chosen for a targeted analysis with the PRM method, and five validated proteins with high sensitivity and specificity, including insulin-like growth factor-binding protein 4 (IGFBP4), β-2-microglobulin (B2M), dystroglycan (DAG1), immunoglobulin J chain (JCHAIN), and kallikrein B1 (KLKB1). The image was created with BioRender.

unified understanding of cancer biology and enabling the development of targeted cancer treatments.[239] A schematic pipeline of CPTAC is represented in Figure 15A.

For example, a proteogenomic study assisted by the CPTAC helped identify S100A9 and GRN as potential combinatorial biomarkers for the early detection of hepatocellular carcinoma (HCC) from urine samples, offering two noninvasive diagnostic tools.[240,241] In lung adenocarcinoma, fusion proteins such as EML4-ALK and HMBOX1-ALK, discovered through RNA-seq and proteomics integration, have been identified as promising biomarkers, enhancing the understanding of tumor-specific mutations and guiding targeted therapies.[242] Similarly, in amyotrophic lateral sclerosis (ALS), ultrasensitive MS assays have been developed to measure isoform levels of C9ORF72, a mutation closely associated with FTD and ALS.[243] In C9ORF72 mutation carriers, the long isoform was found to be significantly

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

Table 8. Pros and Cons of the Use of MS Proteomics over Other Techniques in the Biomarker Verification Process

feature	pros	cons
diagnostic accuracy	high specificity and sensitivity can enhance diagnostic power compared to traditional methods	some MS platforms lack robust validation in diverse clinical populations
multiplexing capability	MS allows simultaneous quantification of multiple biomarkers from a single sample	interpretation of multiplex data can be complex and less transparent for clinicians
noninvasive sample use	compatible with various biofluids (*e.g.*, plasma, urine, saliva), supporting noninvasive testing	variability in sample collection methods and processing can affect reproducibility
speed and throughput	automation and high-throughput workflows can accelerate clinical decision-making	workflow is still slower than some immunoassay-based platforms in point-of-care settings
standardization potential	emerging protocols are improving interlaboratory reproducibility. Need to uniform SOPs (standard operating procedures)	lack of universally accepted clinical standards and reference materials for MS biomarkers
regulatory support	targeted MS assays (like MRM/SRM) are increasingly accepted in regulatory submissions. Nontargeted methods remain unaccepted due to variability problems associated with normalization methods	full validation and official approval are time-consuming and require expensive clinical trials with respect to other quantification techniques
cost-effectiveness	potential for cost-saving when replacing multiple single-analyte tests	high initial setup costs and maintenance make MS less accessible for smaller clinical laboratories.
clinical utility validation	enables retrospective and prospective cohort studies using archived samples	requires large, diverse patient cohorts to demonstrate clinical utility and generalizability
integration into clinical workflow	MS-based diagnostics can be integrated with electronic health records and lab systems	requires bioinformatics infrastructure and trained personnel not always available in clinics

reduced in brain tissue.[244] A recent study by Åkesson et al. utilized a proximity-extension MS proteomics assay combined with next-generation sequencing (Olink Explore HT[245]) to quantify 1,463 proteins in CSF and plasma from patients with early-stage multiple sclerosis.[230] Lower levels of neurofilament light chain (NfL) in CSF were found to be effective in predicting the absence of disease activity two years after sampling (AUC of survival rate = 0.77), and a combination of 11 proteins from CSF (e.g., CXCL13, LTA, NfL) were able to accurately predict the severity of disability worsening (AUC of survival rate = 0.90).[230] Another high-relevance exploratory study comes from Zertuche-Martínez et al., who employed label-free MS proteomics to discover candidate biomarkers from plasma-derived extracellular vesicles (EVs) of patients with cirrhosis and HCC.[246] EVs represent an attractive source of biomarkers due to their biomolecular cargo, and among the total 248 identified proteins, 5 DEPs were found relevant to cirrhosis, and 12 DEPs to HCC, with four (LCAT, SERPINF2, A2M, CRP) of both interest that should be investigated further.[246] In a whole-serum MS proteomics experiment, Lan et al. investigated potential biomarkers for heart failure patients with phlegm-blood stasis syndrome, by comparing the circulating proteome of 20 healthy and 40 heart failure patients.[247] They obtained a set of 84 DEPs with a DIA-untargeted experiment, 37 of which were chosen for a targeted analysis with the PRM method, and five validated proteins with high sensitivity and specificity, including insulin-like growth factor-binding protein 4 (IGFBP4), β-2-microglobulin (B2M), dystroglycan (DAG1), immunoglobulin J chain (JCHAIN), and kallikrein B1 (KLKB1), were considered potential biomarkers for heart failure patients with phlegm-blood stasis syndrome.[247] Their experimental design is represented in Figure 15B. The same untargeted MS proteomics-bioinformatic-targeted MS proteomics pipeline was exploited to validate putative biomarkers from a data set of proteins coming from an explorative analysis by Sanni et al. for the identification of candidate serum biomarkers in patients with narcolepsy Type 1.[248] Here the authors investigated 14 of the DEPs identified from the untargeted analysis, and eight of them maintained the same trend of fold change in the PRM analysis. Inter-α-inhibitor heavy chain 4 (ITIH4), fibronectin 1 (FN1), and complement component (C5) were proposed for further studies after cellular localization analysis.[248]

The recent research from Tapia et al. about HER2+ breast cancer (BR) resistance phenotype is a good example of the target verification stage.[211] Indeed, the authors compared trastuzumab-responding spheroids and trastuzumab-resistant spheroids and found that resistant 3D cells displayed a downregulation of carbohydrate metabolism and upregulation of mitochondria organization proteins, the tricarboxylic acid cycle, and oxidative phosphorylation.[211] A further biostatistical analysis demonstrated that up regulation of Complex I proteins NDUFA10 and NDUFS2 was associated with high clinical risk and worse survival for HER2+ BC patients, which was confirmed by a retrospective analysis of clinical metadata on mRNA expression of mitochondrial Complex I genes in BR patients.[211] Still exploring the field of target verification, researchers recently used a targeted proteomics approach to analyze plasma samples from individuals with early-stage Parkinson's disease, premotor individuals with isolated REM sleep behavior disorder (iRBD), and healthy controls.[231] Proteotypic peptides to be monitored were chosen from previous exploratory untargeted studies, like in Hällqvist et al.,[249] and from known literature data on Parkinson's and neural aging. Eight proteins, including granulin precursor, mannan-binding-lectin-serine-peptidase-2, and plasma-protease-C1-inhibitor, were identified in the targeted experiment as strong predictors for Parkinson's, and machine learning models classified 79% of premotor individuals as at risk for PD up to 7 years before motor symptoms, with 100% accuracy in distinguishing Parkinson's patients from healthy controls. This panel of blood biomarkers holds promise for early detection of Parkinson's and could aid in selecting participants for clinical trials aimed at preventing or delaying the onset of motor symptoms.[249]

Since secretome encompasses signaling molecules, cytokines, and growth factors, which play key roles in intercellular communication and disease progression, its analysis provides direct insights into drug mechanisms of action and novel biomarker identification.[250,251] Closely related to MS secretomics is the analysis of EVs, including exosomes, which are nanosized vesicles released by virtually all cell types.[252] Exosomes, the main components of secretome, carry a diverse cargo of proteins, nucleic acids and lipids, reflecting the physiological or pathological state of their cell of origin.[253] Their proteome is enriched in membrane proteins, signaling

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

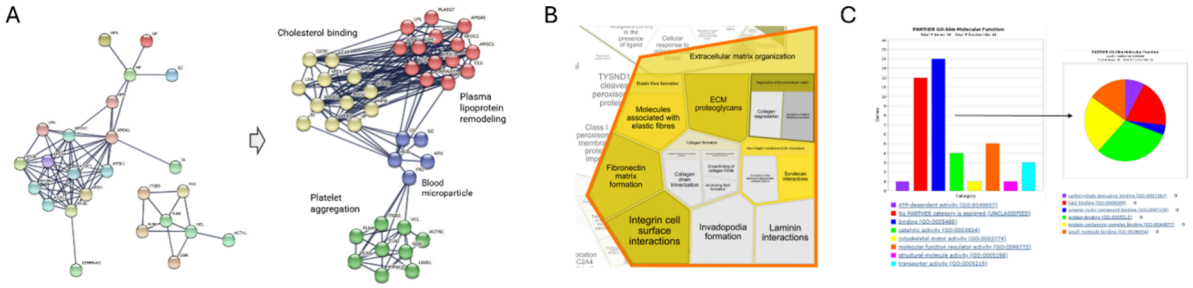


Figure 16. Some of the way to represent bioinformatics data from MS proteomics outcomes. (A) STRING nonenriched network generated with STRING vs its corresponding enriched network (right). (B) Voronoi plot generated with Reactome enrichment tool. (C) Histogram map obtained with GO-Panther encompassing the main biochemical networks identified in an MS proteomics experiment as functional GO (FC), with a further subclass representation (right).

mediators, and enzymes, many of which represent potential drug targets or biomarkers.[254] MS-based exosome proteomics typically involves isolation of vesicles from plasma, CSF, or urine using ultracentrifugation, size-exclusion chromatography, or affinity capture, followed by traditional bottom-up proteomic analysis.[255] Importantly, exosomes protect their cargo from degradation, making them a stable and rich source of novel biomarkers for pathology progression or ongoing therapy success. Many authors have studied these innovative sources of biomarkers coming from different biological tissues. This is the case of Li and colleagues, for instance, who profiled plasma circulating biomarkers in several cohorts of patients affected by lung and brain tumors and found that SELL and MUC5B protein levels could be used as diagnostic markers of brain tumor metastasis, while APOH, CD81, and CCT5 could help diagnose lung metastases in NSCLC.[256] In a similar manner but for a different disease, Zhang et al. used MS proteomics of CSF exosomes to profile the disease activity and long-term prognosis in patients with multiple sclerosis. They considered a cohort of 143 patients vs 43 controls and found that NfL in CSF is superior in predicting the absence of disease activity two years after sampling. Also, they demonstrated that NfL, along with other 10 proteins (XCL13, LTA, FCN2, ICAM3, LY9, SLAMF7, TYMP, CHI3L1, FYB1, and TNFRSF1B) can constitute a panel of biomarkers that can predict disease worsening when corrected by normalized age-related severity score.[230]

Other successful examples demonstrate the potential of MS proteomics in biomarker discovery. Nevertheless, it remains a secondary technique for biomarker investigation, even though it is believed to have huge potential to become the technique of choice. Its main pros and cons are illustrated in Table 8.

9. BIOINFORMATIC TOOLS AND AI

Data integration and bioinformatics are crucial in translating MS proteomic data into meaningful biological insights. By integration of protein concentration data with external biological knowledge, such as gene ontologies, interaction databases, and pathway maps, it is possible to draw comprehensive panels about cellular mechanisms, disease states, and drug effects.

9.1. PPI Networks. PPI networks map the reciprocal relationships between proteins based on experimental data (e.g., co-immunoprecipitation, yeast two-hybrid assays) or computational predictions.[257] The study of PPIs in the context of proteomics involves mapping out these interactions, visualizing the resulting networks, and performing computational analyses to identify key proteins or clusters of proteins involved in specific biological functions or diseases.[258,259] A variety of curated databases and repositories provide experimentally validated or computationally predicted PPIs.[260] One of the most widely used is STRING (Search Tool for the Retrieval of Interacting Genes/Proteins), which aggregates known and predicted PPIs from multiple sources, including experimental data, computational prediction methods, and public text-mining efforts.[261] STRING assigns a confidence score to each interaction, allowing users to filter interactions based on the strength of evidence.[262] A conceptual representation of STRING output is reported in Figure 16A. Other popular PPI databases include BioGRID (Biological General Repository for Interaction Data sets), which focuses on experimental data from high-throughput PPI studies, and IntAct, a database that provides detailed information about experimentally validated PPIs, including the methods used to determine the interactions and the organisms in which they were observed.[263−265] MINT (Molecular INTeraction Database) is another resource that specializes in experimentally determined interactions, with a focus on high-quality, curated interaction data.[265] In a PPI network, proteins are represented as nodes, while their interactions are represented as edges connecting these nodes.[257] The resulting networks can vary in complexity depending on the number of proteins identified and the degree of connectivity between them. In large-scale proteomics studies, where thousands of proteins may be identified, networks can become highly complex.[266] To manage this complexity, computational tools often apply filters to reduce the number of interactions by considering only high-confidence or experimentally validated PPIs.[267] Network construction is further enhanced by integrating additional data, such as gene expression levels, posttranslational modifications, or protein localization information.[85] This allows for the generation of context-specific networks that reflect the state of the proteome under particular experimental conditions, such as disease states or drug treatments.[268]

One of the fundamental properties of PPI networks is degree centrality, which measures the number of connections a protein has within the network.[269] Proteins with a high degree of centrality, often referred to as "hubs," are typically key regulators or scaffolding proteins that play major roles in organizing protein complexes or signaling pathways.[269] Hubs tend to be evolutionarily conserved and are often critical for maintaining cellular homeostasis. In disease contexts, hub proteins are frequently dysregulated, making them attractive targets for therapeutic intervention.[270] Another important topological measure is "betweenness centrality", which quantifies how often a protein lies on the shortest path between other proteins

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

Table 9. Advantages and Limitations of Bioinformatic Data Analysis of the MS Raw Data Output

advantages	limitations
high-throughput analysis: can identify and quantify thousands of proteins in a single experiment	missing values are common, especially for low-abundance proteins, complicating downstream statistical analysis
quantitative comparisons: allows differential expression analysis across multiple conditions/groups	data normalization and imputation choices significantly influence results
pathway and network analysis: integrates protein lists into biological context (e.g., GO, KEGG, Reactome)	functional annotation is sometimes incomplete or outdated for many proteins, especially nonmodel organisms
statistical rigor: supports significance testing (e.g., q-values, fold change) for robust conclusions	thresholding decisions should be decided (e.g., FC cutoffs, q-value) and the choice affect sensitivity and specificity
automation and reproducibility: same workflows can be applied to different experimental conditions (e.g., Perseus, MSstats, Proteome Discoverer)	tool selection variability: different pipelines may yield different results for the same data set
cross-omics integration: enables combination with transcriptomics, metabolomics, etc for multiomics purposes.	data integration challenges due to discordance between mRNA and protein levels
public databases: access to large reference data sets (e.g., PRIDE, PeptideAtlas, UniProt)	batch effects and technical variability can affect biological signals

in the network.[271] Proteins with high betweenness centrality act as bottlenecks or bridges in communication between different parts of the network, making them important for coordinating diverse cellular functions.[272] In addition to centrality measures, clustering coefficient is used to assess the degree to which proteins in a network tend to form clusters or modules.[273] Proteins that interact with each other frequently tend to form tightly connected subnetworks, which can represent protein complexes or functional modules involved in specific biological processes, such as DNA replication, cell division, or metabolic regulation.[274] By identifying these clusters, it is possible to gain insights into the functional organization of the proteome and identify key regulatory modules that may be perturbed in disease conditions.[268] Taking into account the practical steps to build an experimental analysis, the selection of proteins to build a PPI network and perform an enrichment analysis requires a combination of rigorous data processing and statistical criteria to ensure the biological validity of the results.[275] The workflow begins with data preprocessing, which includes the removal of contaminants, reverse and decoy sequences, and proteins with insufficient peptide evidence.[259] Quantitative data are normalized to correct for technical variation between samples, for instance, using median or total ion current normalization for label-free quantification, or reference channel normalization for TMT-based workflows.[276] Missing values are imputed according to the assumed mechanism of missingness: missing not at random (MNAR) values are typically imputed with a down-shifted normal distribution to simulate low-abundance features, whereas missing completely at random (MCAR) values may be replaced using nearest-neighbor or Bayesian methods.[277,278] To ensure data reliability, proteins supported by at least two unique peptides and detected in a minimum of 50 to 70% of biological replicates in at least one condition are typically retained.[279] Minimum coverage (usually set to 5% or above) can also be used as a filtering parameter.[23] Once the data set is quality-filtered, statistical analyses are applied to identify proteins that differ significantly between experimental groups. Moderated *t* tests or linear models are commonly used, as they account for technical and biological variance across replicates.[280] *p*-values are adjusted for multiple testing using the Benjamini-Hochberg procedure to control the false discovery rate (FDR).[281] Proteins are then selected for PPI and enrichment analysis according to both their statistical significance and magnitude of change.[282] A widely adopted threshold is an adjusted q-value <0.05 combined with an absolute $\log_2$ of the normalized abundance ratios ($|\log_2 FC| \geq 0.585$ (corresponding to a 1.5-fold overexpression, or 0.5-fold downregulation vs control), though stricter cutoffs ($|\log_2 FC| \geq 1.0$) may be used for more conservative analyses, sometimes combined with q-value <0.01 in the case of limited biological replicates.[283,284] Following selection, proteins are annotated with standardized identifiers such as UniProt accessions or official gene symbols to ensure compatibility with databases used for PPI network construction (e.g., STRING) and pathway enrichment.[285] The background population for enrichment tests should correspond to the set of all proteins reliably quantified in the experiment, rather than the entire theoretical proteome, to avoid enrichment bias due to MS detectability.[286] In cases in which the list of significant proteins is very large, further refinement can be achieved by prioritizing those with the highest statistical significance or those belonging to pathways of specific biological interest. This can be done, for instance, by ranking proteins by combined significance metrics, such as $-\log_{10}$(q-value), or focusing only on the top n proteins ranked by statistical or biological relevance.[287]

9.2. Gene Ontology (GO) and Pathway Enrichment Analysis. A major bioinformatic tool used to understand the functional implications of proteomics data is Gene Ontology (GO). GO terms describe proteins in terms of three categories: i. Molecular function (MF), that describe the biochemical activities of proteins (e.g., kinase activity, DNA binding). ii. Biological process (BP), which clusters the larger biological processes in which proteins are involved (e.g., cell division, apoptosis). Cellular component (CC), i.e., where proteins are located within the cell (e.g., nucleus, cytoplasm).[288] By statistically over-representing certain GO terms, it is possible to deduce which molecular functions or biological processes are significantly enriched in the data set (Figures 16B,C).[17] Tools like DAVID (Database for Annotation, Visualization, and Integrated Discovery), PANTHER, and GOrilla facilitate this type of analysis by linking the identified proteins with their corresponding GO annotations.[289−292] In addition to GO analysis, pathway enrichment analysis is used to identify metabolic or signaling pathways that are over-represented in the proteomic data. Tools like KEGG (Kyoto Encyclopaedia of Genes and Genomes), Reactome, and Ingenuity Pathway Analysis (IPA) are commonly employed to link proteins with known biological pathways.[289,293−296] For example, a study of cancer samples may reveal enrichment in pathways related to cell cycle regulation, DNA repair, or apoptotic signaling. Enrichment analysis is often conducted using statistical tests such as Fisher's exact test or hypergeometric distribution to determine the significance of pathway over-representation.[292] These analyses can reveal key signaling pathways that are dysregulated in diseases or altered in response to drug treatment, providing insights into the molecular underpinnings of the biological condition. Advantages and limitations of bioinfor-

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

matic data analysis of the MS raw data output are reported in Table 9.

9.3. AI in Managing MS Proteomics Data. The integration of AI into MS proteomics is still at its beginning. It significantly enhances the analysis and management of complex biological data, marking a transformative shift in the field. AI applications are particularly valuable in handling the high-dimensional data sets typical of proteomics, where traditional analytical methods often struggle.[297] For instance, deep learning models such as Prosit have been developed to predict fragmentation patterns in MS data, enabling investigators to achieve more accurate protein identifications, especially in highly multiplexed MS/MS spectra typical of DIA, by implementing classical database searches with intensity-based scores derived from AI predictions.[298,299] Another AI application in MS Proteomics is represented by the Chimerys peptide browser (Thermo Fisher) on the Proteome Discoverer platform. Chimeric spectra often arise when multiple peptides are co-isolated and fragmented in the ms/ms scans. Chimerys has been shown to double the number of peptide identifications compared to classical search algorithms like Sequest HT, achieving identification rates exceeding 80%.[300] Additionally, it increases the average number of identified peptides per protein by approximately 2.5-fold, translating to an average of two PSMs per spectrum.[301] Moreover, AI excels in noise reduction and pattern recognition within complex spectra, which is crucial for identifying low-abundance proteins that might otherwise be overlooked.[297] Machine learning algorithms can effectively denoise spectra, leading to enhanced accuracy in peptide and protein identification.[302] Predictive modeling is another significant application where AI can anticipate protein behavior based on historical data, thus providing insights into how proteins might respond to specific perturbations or treatments.[303] This capability extends to predicting protein−protein interactions and posttranslational modifications, which are essential for understanding cellular functions and disease mechanisms.[21] AI's role is not limited to data processing. Indeed, it also enhances biological interpretation by linking identified proteins with known biological processes and pathways.[304] Furthermore, AI facilitates the integration of proteomics data with other omics data (such as genomics and metabolomics), providing a comprehensive view of cellular processes and enabling more robust biological insights.[305]

In addition to these analytical capabilities, AI supports automated hypothesis generation by suggesting potential follow-up experiments based on identified patterns in the data.[306] For example, if certain proteins are consistently coregulated across different conditions, AI can prompt researchers to explore their functional relationships further.[307] The development of dynamic visualization tools powered by AI also aids researchers in interpreting complex data sets by highlighting significant areas of interest. Real-time analysis is another promising application of AI in proteomics.[308] As MS techniques evolve and data generation accelerates, the need for immediate feedback becomes crucial. AI can provide real-time insights during experiments, particularly in live-cell proteomics, where a timely analysis can influence experimental outcomes.

10. INTEGRATION OF MS PROTEOMICS WITH OTHER *OMICS* PLATFORMS—THE DREAM

The integration of MS proteomics data with other *omics* techniques holds the ability to accelerate the understanding of complex biological systems.[309] MS proteomics provides deep insights into the proteome, revealing relative protein abundances and their overall functions in relation to other proteins.[310] By integration of MS proteomics with other *omics* layers, such as genomics, transcriptomics, metabolomics, and epigenomics, investigators can develop a more holistic view of cellular functions and disease mechanisms. This approach is often referred to as "dual omics" when involving two omics disciplines, or "multiomics" for more, as represented in Figure 17.[311]

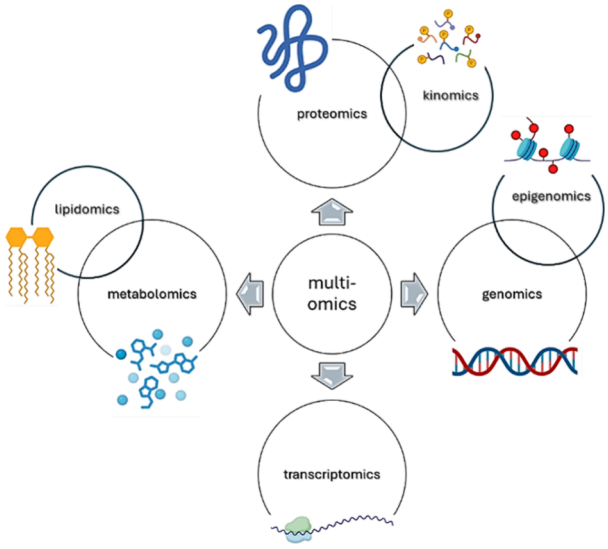


Figure 17. Depiction of the multiomics MS-based panorama, splitting in MS proteomics, genomics, transcriptomics, and metabolomics, and their respective subtechniques. Indeed, MS phosphoproteomics, or kinomics, is a phosphate-focused MS proteomics; epigenomics is the study of DNA methylation and other regulation mechanisms, and lipidomics is a form of MS metabolomics experiment focused on biological analytes with a lipid nature.

Genomics, which provides information about an organism's entire DNA sequence and gene variants, lays the basis for understanding the potential of biological systems. However, not all genetic variations lead to observable phenotypic changes, and gene expression does not always correlate with protein levels due to posttranscriptional and posttranslational modifications. By combining MS proteomics and genomics, it is possible to fill the gap between the genetic code and protein function.[312] This integration, often referred to as proteogenomic, is particularly useful for identifying novel protein-coding regions, alternative splicing events, and the effects of mutations on protein function in diseases such as cancer.[312] Transcriptomics, which analyses RNA expression levels, complements MS proteomics by showing which genes are actively being transcribed.[313] Yet, mRNA levels do not always correlate with protein abundance due to translational regulation and protein degradation. Integrating proteomics with transcriptomics enables a better understanding of the layers of gene regulation, revealing how variations in RNA expression translate into functional proteins.[314]

Metabolomics, the study of small molecules and metabolites within a biological system, offers another layer of valuable information.[315] Proteins, as enzymes, drive metabolic reactions, and the metabolome reflects the biochemical activity occurring in cells and tissues.[316] By integrating MS proteomics with metabolomics, it is possible to connect protein expression levels and modifications to specific metabolic pathways, revealing how

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

protein function impacts metabolic fluxes. This approach is essential in understanding disease metabolism, drug responses, and metabolic disorders.[317] Epigenomics, the study of chemical modifications on DNA and histones that affect gene expression, can also be integrated within the MS proteomics pipeline.[318] Epigenetic modifications influence chromatin structure and accessibility, which in turn affects which genes are transcribed into RNA and ultimately translated into proteins.[3] The integration of epigenomics and proteomics is helpful for studying how epigenetic changes alter protein expression in multifactorial diseases like cancer and neurological disorders.[319] After preprocessing and quality control assessment, the core of a multiomics workflow lies in the analytical strategies that allow data integration across distinct molecular layers. Since each omics data set (genomics, transcriptomics, proteomics, metabolomics, epigenomics, etc.) has unique characteristics in terms of dimensionality, measurement error, dynamic range, and biological interpretation, the design of the analysis pipeline strongly influences the type of insights that can be drawn. Numerous platforms were also designed to analyze in parallel MS Proteomics, RNA-seq data, and metabolomics data and combine the output to a single, unified multiomics biochemical network. These platforms include mixOmics (an R platform), which can deal with all the three techniques and uses multivariate analysis, MOFA (multi-omics factor analysis), iClusterPlus, and SNF (Similarity Network Fusion), ideal for cluster analysis, and others.[320−323] More user-friendly ecosystems, mostly web-based, include PaintOmics and Perseus, which are mainly proteomic-centered, but when provided with already analyzed RNA-seq or metabolomics data from other platforms, are able to merge the information and build correlation plots as well as integrated KEGG and GO pathway levels.[324,325] The integration of different omics layers is possible thanks to the creation of integrated databases like CPTAC, a compendium of a compendium of tissue samples from different cancer types, along with their respective healthy tissue samples.[326] Multiomics data analysis approaches are usually classified into three broad categories: early integration, intermediate integration, and late integration, that in practice are often combined or tailored depending on the biological question and the data available.[327,328] Early integration refers to approaches in which features from different omics layers are merged into a single data matrix before statistical or ML analyses. This strategy treats all omics features as comparable variables, often after normalization or transformation to bring them onto a common scale.[329] The simplest examples are concatenation-based methods, where transcript counts, protein abundances, and metabolite intensities are stacked together, and dimensionality reduction techniques such as PCA or clustering algorithms are applied to the combined data set. This is the case of MONIER, a novel multiomics early Integration framework addressed to disease diagnosis and biomarker discovery.[330] Early integration has the advantage of simplicity and provides a global view of the molecular system, to allow unsupervised methods to detect overarching patterns, patient subgroups, or molecular signatures that span multiple omics levels.[331] However, its substantial challenges are the differences in feature numbers (e.g., tens of thousands of transcripts versus a few hundred metabolites), and the missing values can bias analyses toward the most abundant or best-measured data type.[332] In addition, early integration does not explicitly model the hierarchical or causal relationships between omics layers, which may limit interpretability.[333] Intermediate integration methods attempt to overcome these limitations by jointly modeling omics data sets while preserving the structure and uniqueness of each layer. Instead of concatenating all features, these methods extract latent variables or factors that capture shared variance across data sets while also retaining data set-specific variation.[334] Canonical correlation analysis (CCA), as provided in Smccnet 2.0 toolm, is a classic example as it identifies linear combinations of variables that maximize correlations between omics layers.[335] More recent developments include partial least-squares (PLS)-based frameworks, as implemented in DIABLO (Data Integration Analysis for Biomarker discovery using Latent cOmponents), which enable supervised integration of multiomics data in relation to an outcome variable.[336] Similarly, MOFA provides an unsupervised Bayesian framework that decomposes variability into factors representing both shared and modality-specific signals.[323] Other matrix factorization, manifold learning, and deep-learning-based methods (such as variational autoencoders) have also been developed to capture nonlinear relationships between omics types.[337] The strength of intermediate integration is its ability to balance global structure with omics-specific insights, offering interpretable models that can highlight both cross-omics associations and unique biological contributions from each data set.[338] Late integration represents the most conservative strategy, in which each omics layer is analyzed separately, and the results are integrated at a higher level of biological interpretation.[339] This approach is particularly valuable when data sets are heterogeneous in size and/or quality, since it avoids forcing them into a single modeling framework.[339] Network-based late integration, as in the MOLI method proposed by Sharifi-Noghabi et al., is also common, where omics-specific networks (e.g., coexpression or PPI maps) are generated independently and then overlaid to reveal cross-layer regulatory modules.[340] The downside is that late integration may miss subtle, multilayer associations that only emerge when data sets are modeled together.[341] However, because it allows maximum flexibility and makes use of well-established single-omics pipelines, it still remains the most widely used in the scientific literature, especially in translational and clinical studies where robustness of each independent omics experiments and interpretability are prioritized.[333] In practice, the choice between early, intermediate, and late integration is not absolute but rather depends on the research context and the aims of the investigation. Early integration is often exploratory and well suited for discovering broad patient stratification or molecular subtypes. Intermediate integration offers a powerful balance of depth and interpretability, and it is particularly useful in biomarker discovery or mechanistic studies, where the interplay between omics layers is of interest. Late integration remains valuable when dealing with heterogeneous data sets, legacy data, or when interpretability and reproducibility take precedence.[342] Increasingly, hybrid strategies are being developed, for instance, combining late integration of pathway analyses with intermediate modeling of molecular networks (as in the directional P-value merging (DPM) method), to capture the complementary strengths of each approach.[85]

In conclusion, the integration of MS-based proteomics with complementary omics disciplines provides a systems-level framework that can accelerate drug discovery by connecting genomic variation, transcriptional regulation, protein dynamics, and metabolic fluxes into unified molecular networks. While challenges related to data dimensionality, missing values, and cross-platform normalization persist, the refinement of early, intermediate, and late integration algorithms, together with the

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

development of robust computational ecosystems, is progressively overcoming these limitations. As these approaches mature, MS proteomics embedded within multiomics integration will be the key for biomarker discovery, early-stage drug discovery, and mechanism-based therapeutic design, ultimately enhancing translational impact in medicinal chemistry and making more fruitful the collaboration between chemists, pharmacologists, and clinicians.

11. PERSPECTIVES

MS-based proteomics is quickly becoming an indispensable tool in modern drug discovery programs, delivering unparalleled information regarding protein expression, interactions, and posttranslational modifications. MS proteomics can be applied to all stages of the drug development pipeline, from target identification and validation to MoA clarification, biomarker identification, and clinical studies. The hallmark of one of the most significant strengths of MS proteomics lies in its ability to analyze proteins within their native biological environments. This is a possibility with chemoproteomics, e.g., with ABPP, HDX-MS, XL-MS, and other ad hoc experiments, which allow us to investigate protein functions and conformational changes with high sensitivity. These technologies provide a degree of biological information above that of traditional transcriptomic or genomic approaches. Recent technological developments in chemoproteomics methodologies, including TPP-MS and PAL, have significantly improved the capability to study drug−target interactions and off-target activities in complex biological systems. These approaches have been crucial in identifying direct targets of synthetic and natural product drugs, drug resistance mechanisms, and drug repurposing targets. Furthermore, the integration of MS proteomics into FBDD has enabled early-stage development through the capacity to rapidly, label-free confirm fragment binding and the stoichiometry of interactions.

At the preclinical stage, MS-based proteomics plays a central role in characterizing drug candidates' mechanisms of action, identifying off-target activities, and establishing their toxicity profiles. XL-MS, Co/IP-MS, HDX, LiP-MS, and ABPP enable the exploration of protein interaction with their ligands and with other proteins, providing a systems-level understanding of the drug MoA. Such data are essential for lead optimization and reduction of late-stage attrition. Moreover, MS proteomics also provides the possibility of identifying pharmacodynamic biomarkers that are important for drug efficacy and safety monitoring within clinical trials. By analyzing patient samples in therapeutic regimens at different time-lapses, translational investigators can simplify the evaluation of the molecular effect of therapeutics, thus making adaptive trial designs and personalized treatment regimens possible. As the field advances toward systems biology and individualized medicine, MS proteomics has stood the test of translational research. Clinical specimen analysis, from tissue biopsy to ascitic fluid and plasma, has aided in identifying disease-specific biomarkers and facilitated the stratification of patients. Protein information integration with clinical metadata is opening the gates for more precise, data-informed treatment choices and overcoming bench-to-bedside translational gaps.

Since 2024, single-cell proteomics has represented a rapidly advancing field at the intersection of MS and cellular biology, aimed at characterizing the proteome of individual cells rather than bulk populations. Traditional proteomic workflows average signals across thousands to millions of cells, thereby masking the heterogeneity that underlies critical biological processes such as differentiation, tumor evolution, and drug resistance. By contrast, single-cell proteomics leverages ultrasensitive sample preparation strategies (e.g., nanodroplet-based digestion, acoustic and microfluidic isolation) and state-of-the-art mass spectrometers (Orbitrap Astral from Thermo Fisher and timsTOF SCP from Bruker) to achieve reliable detection of proteins from a single cell. Quantification strategies include both label-free approaches and isobaric labeling with carrier proteome boosting (as pioneered in SCoPE-MS and extended in SCoPE2 and nPOP), which mitigate missing data. However, the field still faces technical and conceptual limitations. Sensitivity remains a major bottleneck, as protein copy numbers span several orders of magnitude, and many low-abundance proteins remain undetectable. Quantification accuracy is challenged by ion undersampling (i.e., still an important amount of signals is not detected by current MS proteomics experiments, due to low abundance), ratio compression in multiplexed analyses, and stochastic sampling in DDA, although DIA and PASEF (Parallel Accumulation Serial Fragmentation, by Bruker) have alleviated some of these issues.

MS proteomics has also allowed the development of the so-called "drug fingerprinting," the proteome-based analogue to gene expression profiling in transcriptomics (e.g., LINCS project), which captures functional and post-translational responses, i.e. it aims to translate the concept of transcriptomic perturbation profiling (as exemplified by the Connectivity Map, https://www.broadinstitute.org/connectivity-map-cmap) into the proteomic domain. The concept is to associate a chemical scaffold with whole cellular modifications. This technique may be implemented to drive drug discovery within a new, extended concept of SAR, by expanding the activity of the drug from a single target to the whole biochemical modifications induced by the considered chemical scaffold. A landmark example of MS proteomic drug fingerprinting is the ProTargetMiner resource, which systematically profiles the effects of ∼50 kinase inhibitors across multiple cancer cell lines using quantitative MS-based proteomics. This study demonstrated that proteomic responses could classify compounds according to their known mechanisms and uncover unanticipated off-target activities. More recent efforts, such as ProteomeHD, have extended this principle to broader chemical spaces and cellular systems, driving the research toward comprehensive proteome-level connectivity maps analogous to transcriptomic databases.

Furthermore, the exploitation of machine learning (ML) in MS proteomics workflows is a promising frontier. ML-based tools are being utilized more to address the high-dimensional data sets generated by modern mass spectrometers, enhancing sensitivity, reproducibility, and predictive modeling of drug toxicities and response. Programs such as MaxQuant, Skyline, Progenesis and Proteome Discoverer, as well as emerging AI-enabled platforms, have accelerated data interpretation and confidence, further bolstering MS proteomics' potential for high-throughput drug screening and mechanistic exploration. The integration of MS-based proteomics into drug discovery has led to the generation of enormous and intricate data sets, which call for stable data storage and management solutions. Traditional file-based systems are not able to handle the quantity and complexity of proteomic data, which leads to the retrieval, analysis, and sharing problems of the data. To address these challenges, various efforts have been put in place to standardize data formats and make data sharing easier and quicker. An example is the ProteomeXchange consortium,

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

which provides a common framework for data submission and distribution across several MS proteomics repositories, i.e., PRIDE, PeptideAtlas, and MassIVE. These repositories support data sharing and reanalysis through providing standardized formats and rich metadata annotations, which is also critical in ensuring the reproducibility and reusability of proteomics data. Thorough annotation of experimental conditions, procedures in sample treatment, and analysis parameters must be performed to facilitate meaningful interpretation and comparison of the data. Efforts like the Proteomics Standards Initiative (PSI) have aided in the development of standardized data formats and controlled vocabularies (e.g., the Minimum Information about a proteomics experiment (MIAPE) standard format) to ensure greater levels of metadata consistency. Furthermore, platforms such as iProX have embraced big data technologies in managing the growing size of proteomics data sets. By scalable design and fast data retrieval systems, iProX supports the storage and rapid data retrieval of large quantities of MS proteomic data, facilitating high-throughput analysis and retrospective investigations.

With the increase in MS proteomics workflows, coupling with other omics technologies such as genomics and metabolomics will help to bring a better understanding of disease mechanisms and therapeutic response. The development of advanced bioinformatics tools will again enhance the analysis and interpretation of high-complexity proteomic data and result in more efficient and targeted approaches to drug discovery. Also, its increasing levels of automatization from sample processing to data analysis and the always more consistent miniaturization of high-performance LC−MS/MS systems will allow high-quality data to be obtained and analyzed in large batches. On the other hand, limitations to the full exploitation of this technique still exist and include the consistent costs of MS Proteomics platforms (both as MS machines purchase and their maintenance, including specialized and dedicated personnel and consumables) and of data storage clouds and platforms, e.g., supercomputers. Also, omics sciences are intrinsically affected by interlaboratory reproducibility issues, as currently, few to no operative standard procedures are internationally recognized, especially for the label-free MS proteomics experiment. However, as for nearly any other technological applications and advancements of the most recent era, the costs of HRMS will become more and more affordable. With ongoing collaboration across academia, industry, and regulatory bodies, it is foreseeable that standardized, scalable, and cost-effective MS proteomics workflows will soon become an essential part of the pharmaceutical toolbox to accelerate the drug discovery process associated with biomedical investigations.

AUTHOR INFORMATION

Corresponding Authors

Lorenzo Tagliazucchi − *Department of Life Sciences, University of Modena and Reggio Emilia, 41125 Modena, Italy;* Present Address: Department of Food and Drug Sciences, University of Parma, Parco Area delle Scienze 27/A, 43124 Parma, Italy; Email: lorenzo.tagliazucchi@unipr.it

Maria Paola Costi − *Department of Life Sciences, University of Modena and Reggio Emilia, 41125 Modena, Italy;* orcid.org/0000-0002-0443-5402; Email: costimp@unimore.it

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jmedchem.5c01986

Notes

The authors declare no competing financial interest.

Biographies

Lorenzo Tagliazucchi is a tenured track researcher at the Drug Design and Discovery Group (https://saf.unipr.it/node/1323, Prof. Marco Mor) in the University of Parma, Italy. He is working in the fields of MS Chemical proteomics and chemical biology for the characterization of drug candidates, protein−protein interaction systems, and bioanalytic tools for target engagement studies. His recent work is on cancer drug resistance with MS omics techniques. He obtained a M.S. degree in Medicinal Chemistry with a thesis in toxicological chemistry at the Italian Minister of Agriculture (Milan). He spent 2 years as a predoc in the Drug discovery and biotechnology lab at Unimore, Italy. He owns a Ph.D. in Clinical and Experimental Medicine under the supervision of Prof. MP Costi. He was a visiting PhD at Dr. T. Katsila's Lab (NHRF, Athens, Greece).

Maria Paola Costi is a Professor of Medicinal Chemistry and a P.I. of the Drug Discovery and Biotechnology Lab at the University of Modena and Reggio Emilia. Her expertise includes medicinal chemistry, chemical biology, and protein engineering. Topics of her international collaborative research work are related to folate-dependent enzymes in cancer and parasitic diseases, and in the pharmacological modulation of the Hippo pathway. She has contributed to over 195 papers and book chapters in international publications and owns 20 patents. She is a member of the MITO oncology consortium, and other networks and foundations. She is the coordinator of several collaborative projects, such as European projects, two AIRC-funded projects, and a European COST Action.

ACKNOWLEDGMENTS

We apologize for the numerous excellent works that have not been included in this perspective, yet still contribute to the development and progression of the MS proteomics and chemical proteomics panorama. We want to thank the AIRC IG25785 to MPC and the COST Action CA2111 *OneHealth-Drugs* for inspiring this perspective and pursuing the necessity of integrating the omics platforms in the modern drug discovery pipeline as a sustainable approach.

ABBREVIATIONS

ABPP,activity-based protein profiling; AOP,adverse outcome pathway; CID,collision-induced dissociation; Co-IP-MS,co-immunoprecipitation followed by MS; DDA,data-dependent acquisition; DEP,differentially expressed protein; DIA,data-independent acquisition; FBDD,fragment-based drug design; FPLC,fast protein liquid chromatography; HCD,higher collision dissociation; HDX-MS,hydrogen−deuterium exchange mass spectrometry; HRMS,high-resolution mass spectrometry; HTS,high-throughput screening; IMAC,immobilized-metal affinity chromatography; LBDD,ligand-based drug design; LFQ,label-free quantification; MD,molecular dynamics; MRM,multiple reaction monitoring; MS,mass spectrometry; PAL,photoaffinity labeling; PIN,protein interaction network; PPI,protein−protein interaction; PRM,parallel reaction monitoring; PTM,posttranslational modification; SBDD,structure-based drug design; SRM,single reaction monitoring; TIC,total ion current; TMT,tandem mass tag; TPP-MS,thermal proteome profiling mass spectrometry; WB,Western blot; XIC,extracted ion chromatogram; XL-MS,cross-linking mass spectrometry

REFERENCES

(1) Chan, H. C. S.; Shan, H.; Dahoun, T.; Vogel, H.; Yuan, S. Advancing Drug Discovery via Artificial Intelligence. *Trends Pharmacol. Sci.* **2019**, *40* (8), 592−604.

(2) Berg, E. L. The Future of Phenotypic Drug Discovery. *Cell Chem. Biol.* **2021**, *28* (3), 424−430.

(3) Noberini, R.; Bonaldi, T. Proteomics Contributions to Epigenetic Drug Discovery. *Proteomics* **2023**, *23*, 23−24.

(4) Keane, R. E.; Tidy, R. J.; Parker, G. J.; Gummer, J. P. A.; Priddis, C. Mass Spectrometry Based Proteomics: Changing the Impact of Protein Analysis in Forensic Science. *WIREs Forensic Science* **2024**, *6* (4), No. e1516.

(5) Frantzi, M.; Latosinska, A.; Mischak, H. Proteomics in Drug Development: The Dawn of a New Era? *Proteomics Clin Appl.* **2019**, *13* (2), No. e1800087.

(6) Timp, W.; Timp, G. Beyond Mass Spectrometry, the next Step in Proteomics. *Sci. Adv.* **2020**, *6* (2), No. eaax8978.

(7) Pappireddi, N.; Martin, L.; Wühr, M. A Review on Quantitative Multiplexed Proteomics. *ChemBioChem.* **2019**, *20* (10), 1210−1224.

(8) Gao, Y.; Ma, M.; Li, W.; Lei, X. Chemoproteomics, A Broad Avenue to Target Deconvolution. *Adv. Sci. (Weinh)* **2024**, *11* (8), No. e2305608.

(9) Thomas, J. R.; Brittain, S. M.; Lipps, J.; Llamas, L.; Jain, R. K.; Schirle, M. A Photoaffinity Labeling-Based Chemoproteomics Strategy for Unbiased Target Deconvolution of Small Molecule Drug Candidates. *Methods Mol. Biol.* **2017**, *1647*, 1−18.

(10) Chan, W. C.; Sharifzadeh, S.; Buhrlage, S. J.; Marto, J. A. Chemoproteomic Methods for Covalent Drug Discovery. *Chem. Soc. Rev.* **2021**, *50* (15), 8361−8381.

(11) Smith, E.; Collins, I. Photoaffinity Labeling in Target- and Binding-Site Identification. *Future Med. Chem.* **2015**, *7* (2), 159−183.

(12) Pedro, L.; Quinn, R. Native Mass Spectrometry in Fragment-Based Drug Discovery. *Molecules* **2016**, *21* (8), 984.

(13) Agostini, M.; Traldi, P.; Hamdan, M. Mass Spectrometry-Based Proteomics: Analyses Related to Drug-Resistance and Disease Biomarkers. *Medicina (B Aires)* **2023**, *59* (10), 1722.

(14) Saboia-Vahia, L.; de Jesus, J. B.; Cuervo, P. *The Role of Proteomics in the Study of Drug Resistance*. In Drug Resistance in Leishmania Parasites; Springer International Publishing: Cham, 2018; pp 209−245..

(15) Fortuin, S.; Soares, N. C. The Integration of Proteomics and Metabolomics Data Paving the Way for a Better Understanding of the Mechanisms Underlying Microbial Acquired Drug Resistance. *Front Med. (Lausanne)* **2022**, *9*, No. 849838.

(16) Meng, F.; Zou, L.; Zhang, T.; Jiang, L.; Ding, Y.; Yu, P.; Peng, J. Using LC-MS/MS-Based Targeted Proteomics to Monitor the Pattern of ABC Transporters Expression in the Development of Drug Resistance. *Cancer Manag Res.* **2018**, *10*, 2859−2870.

(17) Duarte, T.; Spencer, C. Personalized Proteomics: The Future of Precision Medicine. *Proteomes* **2016**, *4* (4), 29.

(18) Meo, A. D.; Pasic, M. D.; Yousef, G. M. Proteomics and Peptidomics: Moving toward Precision Medicine in Urological Malignancies. *Oncotarget* **2016**, *7* (32), 52460−52474.

(19) He, T.; Liu, Y.; Zhou, Y.; Li, L.; Wang, H.; Chen, S.; Gao, J.; Jiang, W.; Yu, Y.; Ge, W.; Chang, H.-Y.; Fan, Z.; Nesvizhskii, A. I.; Guo, T.; Sun, Y. Comparative Evaluation of Proteome Discoverer and FragPipe for the TMT-Based Proteome Quantification. *J. Proteome Res.* **2022**, *21* (12), 3007−3015.

(20) Macklin, A.; Khan, S.; Kislinger, T. Recent Advances in Mass Spectrometry Based Clinical Proteomics: Applications to Cancer Research. *Clin Proteomics* **2020**, *17* (1), 17.

(21) Cui, M.; Cheng, C.; Zhang, L. High-Throughput Proteomics: A Methodological Mini-Review. *Laboratory Investigation* **2022**, *102* (11), 1170−1181.

(22) Mann, M.; Kumar, C.; Zeng, W.-F.; Strauss, M. T. Artificial Intelligence for Proteomics and Biomarker Discovery. *Cell Syst* **2021**, *12* (8), 759−770.

(23) Frejno, M.; Berger, M. T.; Tüshaus, J.; Hogrebe, A.; Seefried, F.; Graber, M.; Samaras, P.; Ben Fredj, S.; Sukumar, V.; Eljagh, L.; Bronshtein, I.; Mamisashvili, L.; Schneider, M.; Gessulat, S.; Schmidt, T.; Kuster, B.; Zolg, D. P.; Wilhelm, M. Unifying the Analysis of Bottom-up Proteomics Data with CHIMERYS. *Nat. Methods* **2025**, *22* (5), 1017−1027.

(24) Cassidy, L.; Kaulich, P. T.; Maaß, S.; Bartel, J.; Becher, D.; Tholey, A. Bottom-up and Top-down Proteomic Approaches for the Identification, Characterization, and Quantification of the Low Molecular Weight Proteome with Focus on Short Open Reading Frame-encoded Peptides. *Proteomics* **2021**, *21*, 23−24.

(25) Vidova, V.; Spacil, Z. A Review on Mass Spectrometry-Based Quantitative Proteomics: Targeted and Data Independent Acquisition. *Anal. Chim. Acta* **2017**, *964*, 7−23.

(26) Borràs, E.; Sabidó, E. What Is Targeted Proteomics? A Concise Revision of Targeted Acquisition and Targeted Data Analysis in Mass Spectrometry. *Proteomics* **2017**, *17* (17−18), 1700180.

(27) Nguyen, C. D. L.; Malchow, S.; Reich, S.; Steltgens, S.; Shuvaev, K. V.; Loroch, S.; Lorenz, C.; Sickmann, A.; Knobbe-Thomsen, C. B.; Tews, B.; Medenbach, J.; Ahrends, R. A Sensitive and Simple Targeted Proteomics Approach to Quantify Transcription Factor and Membrane Proteins of the Unfolded Protein Response Pathway in Glioblastoma Cells. *Sci. Rep* **2019**, *9* (1), 8836.

(28) Yu, Q.; Liu, X.; Keller, M. P.; Navarrete-Perea, J.; Zhang, T.; Fu, S.; Vaites, L. P.; Shuken, S. R.; Schmid, E.; Keele, G. R.; Li, J.; Huttlin, E. L.; Rashan, E. H.; Simcox, J.; Churchill, G. A.; Schweppe, D. K.; Attie, A. D.; Paulo, J. A.; Gygi, S. P. Sample Multiplexing-Based Targeted Pathway Proteomics with Real-Time Analytics Reveals the Impact of Genetic Variation on Protein Expression. *Nat. Commun.* **2023**, *14* (1), 555.

(29) Chalikiopoulou, C.; Gómez-Tamayo, J. C.; Katsila, T. Untargeted Metabolomics for Disease-Specific Signatures. *Methods Mol. Biol.* **2023**, *2571*, 71−81.

(30) Nakayasu, E. S.; Gritsenko, M.; Piehowski, P. D.; Gao, Y.; Orton, D. J.; Schepmoes, A. A.; Fillmore, T. L.; Frohnert, B. I.; Rewers, M.; Krischer, J. P.; Ansong, C.; Suchy-Dicey, A. M.; Evans-Molina, C.; Qian, W.-J.; Webb-Robertson, B.-J. M.; Metz, T. O. Tutorial: Best Practices and Considerations for Mass-Spectrometry-Based Protein Biomarker Discovery and Validation. *Nat. Protoc* **2021**, *16* (8), 3737−3760.

(31) Zanon Agapito-Tenfen, S.; Guerra, M. P.; Nodari, R. O.; Wikmark, O.-G. Untargeted Proteomics-Based Approach to Investigate Unintended Changes in Genetically Modified Maize for Environmental Risk Assessment Purpose. *Frontiers in Toxicology* **2021**, *3*, No. 655968.

(32) Sobsey, C. A.; Ibrahim, S.; Richard, V. R.; Gaspar, V.; Mitsa, G.; Lacasse, V.; Zahedi, R. P.; Batist, G.; Borchers, C. H. Targeted and Untargeted Proteomics Approaches in Biomarker Development. *Proteomics* **2020**, *20* (9), No. e1900029.

(33) Tsai, T.-H.; Wang, M.; Ressom, H. W. Preprocessing and Analysis of LC-MS-Based Proteomic Data. *Methods Mol. Biol.* **2016**, *1362*, 63−76.

(34) Schubert, O. T.; Röst, H. L.; Collins, B. C.; Rosenberger, G.; Aebersold, R. Quantitative Proteomics: Challenges and Opportunities in Basic and Applied Research. *Nat. Protoc* **2017**, *12* (7), 1289−1294.

(35) Zhang, Y.; Fonslow, B. R.; Shan, B.; Baek, M.-C.; Yates, J. R., 3rd Protein Analysis by Shotgun/Bottom-up Proteomics. *Chem. Rev.* **2013**, *113* (4), 2343−2394.

(36) Santos, M. D. M.; Lima, D. B.; Fischer, J. S. G.; Clasen, M. A.; Kurt, L. U.; Camillo-Andrade, A. C.; Monteiro, L. C.; de Aquino, P. F.; Neves-Ferreira, A. G. C.; Valente, R. H.; Trugilho, M. R. O.; Brunoro, G. V. F.; Souza, T. A. C. B.; Santos, R. M.; Batista, M.; Gozzo, F. C.; Durán, R.; Yates, J. R.; Barbosa, V. C.; Carvalho, P. C. Simple, Efficient and Thorough Shotgun Proteomic Analysis with PatternLab V. *Nat. Protoc* **2022**, *17* (7), 1553−1578.

(37) Ma, M.; Huo, S.; Zhang, M.; Qian, S.; Zhu, X.; Pu, J.; Rasam, S.; Xue, C.; Shen, S.; An, B.; Wang, J.; Qu, J. In-Depth Mapping of Protein Localizations in Whole Tissue by Micro-Scaffold Assisted Spatial Proteomics (MASP). *Nat. Commun.* **2022**, *13* (1), 7736.

(38) Lee, Y. H.; Tan, H. T.; Chung, M. C. M. Subcellular Fractionation Methods and Strategies for Proteomics. *Proteomics* **2010**, *10* (22), 3935−3956.

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

(39) Geladaki, A.; Kočevar Britovšek, N.; Breckels, L. M.; Smith, T. S.; Vennard, O. L.; Mulvey, C. M.; Crook, O. M.; Gatto, L.; Lilley, K. S. Combining LOPIT with Differential Ultracentrifugation for High-Resolution Spatial Proteomics. *Nat. Commun.* **2019**, *10* (1), 331.
(40) Lee, H.-J.; Kim, H.-J.; Liebler, D. C. Efficient Microscale Basic Reverse Phase Peptide Fractionation for Global and Targeted Proteomics. *J. Proteome Res.* **2016**, *15* (7), 2346−2354.
(41) Mant, C. T.; Chen, Y.; Yan, Z.; Popa, T. V; Kovacs, J. M.; Mills, J. B.; Tripet, B. P.; Hodges, R. S. *HPLC Analysis and Purification of Peptides*. In Peptide Characterization and Application Protocols; Methods in molecular biology; Humana Press: Clifton, Totowa, NJ, 2007; pp 3−55..
(42) Wang, F.; Veth, T.; Kuipers, M.; Altelaar, M.; Stecker, K. E. Optimized Suspension Trapping Method for Phosphoproteomics Sample Preparation. *Anal. Chem.* **2023**, *95* (25), 9471−9479.
(43) Chen, C.-W.; Tsai, C.-F.; Lin, M.-H.; Lin, S.-Y.; Hsu, C.-C. Suspension Trapping-Based Sample Preparation Workflow for in-Depth Plant Phosphoproteomics. *Anal. Chem.* **2023**, *95* (33), 12232−12239.
(44) Distler, U.; Kuharev, J.; Navarro, P.; Tenzer, S. Label-Free Quantification in Ion Mobility−Enhanced Data-Independent Acquisition Proteomics. *Nat. Protoc* **2016**, *11* (4), 795−812.
(45) Anand, S.; Samuel, M.; Ang, C.-S.; Keerthikumar, S.; Mathivanan, S. Label-Based and Label-Free Strategies for Protein Quantitation. *Methods Mol. Biol.* **2017**, *1549*, 31−43.
(46) Fu, J.; Yang, Q.; Luo, Y.; Zhang, S.; Tang, J.; Zhang, Y.; Zhang, H.; Xu, H.; Zhu, F. Label-Free Proteome Quantification and Evaluation. *Brief Bioinform* **2023**, *24* (1), No. bbac477.
(47) Deracinois, B.; Flahaut, C.; Duban-Deweer, S.; Karamanos, Y. Comparative and Quantitative Global Proteomics Approaches: An Overview. *Proteomes* **2013**, *1* (3), 180−218.
(48) Lee, H. Y.; Kim, E. G.; Jung, H. R.; Jung, J. W.; Kim, H. B.; Cho, J. W.; Kim, K. M.; Yi, E. C. Refinements of LC-MS/MS Spectral Counting Statistics Improve Quantification of Low Abundance Proteins. *Sci. Rep* **2019**, *9* (1), 13653.
(49) Webel, H.; Niu, L.; Nielsen, A. B.; Locard-Paulet, M.; Mann, M.; Jensen, L. J.; Rasmussen, S. Imputation of Label-Free Quantitative Mass Spectrometry-Based Proteomics Data Using Self-Supervised Deep Learning. *Nat. Commun.* **2024**, *15* (1), 5405.
(50) Zhu, W.; Smith, J. W.; Huang, C.-M. Mass Spectrometry-Based Label-Free Quantitative Proteomics. *J. Biomed Biotechnol* **2010**, *2010*, 1−6.
(51) Arike, L.; Valgepea, K.; Peil, L.; Nahku, R.; Adamberg, K.; Vilu, R. Comparison and Applications of Label-Free Absolute Proteome Quantification Methods on Escherichia Coli. *J. Proteomics* **2012**, *75* (17), 5437−5448.
(52) Kettenbach, A. N.; Rush, J.; Gerber, S. A. Absolute Quantification of Protein and Post-Translational Modification Abundance with Stable Isotope−Labeled Synthetic Peptides. *Nat. Protoc* **2011**, *6* (2), 175−186.
(53) Cox, J.; Hein, M. Y.; Luber, C. A.; Paron, I.; Nagaraj, N.; Mann, M. Accurate Proteome-Wide Label-Free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction. *Termed MaxLFQ. Mol. Cell Proteomics* **2014**, *13* (9), 2513−2526.
(54) Millán-Oropeza, A.; Blein-Nicolas, M.; Monnet, V.; Zivy, M.; Henry, C. Comparison of Different Label-Free Techniques for the Semi-Absolute Quantification of Protein Abundance. *Proteomes* **2022**, *10* (1), 2.
(55) Wiese, S.; Reidegeld, K. A.; Meyer, H. E.; Warscheid, B. Protein Labeling by ITRAQ: A New Tool for Quantitative Mass Spectrometry in Proteome Research. *Proteomics* **2007**, *7* (3), 340−350.
(56) Chen, X.; Wei, S.; Ji, Y.; Guo, X.; Yang, F. Quantitative Proteomics Using SILAC: Principles, Applications, and Developments. *Proteomics* **2015**, *15* (18), 3175−3192.
(57) Zecha, J.; Satpathy, S.; Kanashova, T.; Avanessian, S. C.; Kane, M. H.; Clauser, K. R.; Mertins, P.; Carr, S. A.; Kuster, B. TMT Labeling for the Masses: A Robust and Cost-Efficient, In-Solution Labeling Approach. *Molecular & Cellular Proteomics* **2019**, *18* (7), 1468−1478.
(58) Geiger, T.; Wisniewski, J. R.; Cox, J.; Zanivan, S.; Kruger, M.; Ishihama, Y.; Mann, M. Use of Stable Isotope Labeling by Amino Acids in Cell Culture as a Spike-in Standard in Quantitative Proteomics. *Nat. Protoc* **2011**, *6* (2), 147−157.
(59) Tian, X.; Permentier, H. P.; Bischoff, R. Chemical Isotope Labeling for Quantitative Proteomics. *Mass Spectrom Rev.* **2023**, *42* (2), 546−576.
(60) Ye, X.; Luke, B.; Andresson, T.; Blonder, J. 18O Stable Isotope Labeling in MS-Based Proteomics. *Brief Funct Genomic Proteomic* **2009**, *8* (2), 136−144.
(61) Hsu, J.-L.; Chen, S.-H. Stable Isotope Dimethyl Labelling for Quantitative Proteomics and Beyond. *Philos. Trans A Math Phys. Eng. Sci.* **2016**, *374* (2079), 20150364.
(62) Bateman, N. W.; Goulding, S. P.; Shulman, N. J.; Gadok, A. K.; Szumlinski, K. K.; MacCoss, M. J.; Wu, C. C. Maximizing Peptide Identification Events in Proteomic Workflows Using Data-Dependent Acquisition (DDA). *Molecular & Cellular Proteomics* **2014**, *13* (1), 329−338.
(63) Yu, F.; Deng, Y.; Nesvizhskii, A. I. MSFragger-DDA+ Enhances Peptide Identification Sensitivity with Full Isolation Window Search. *Nat. Commun.* **2025**, *16* (1), 3329.
(64) Fröhlich, K.; Fahrner, M.; Brombacher, E.; Seredynska, A.; Maldacker, M.; Kreutz, C.; Schmidt, A.; Schilling, O. Data-Independent Acquisition: A Milestone and Prospect in Clinical Mass Spectrometry−Based Proteomics. *Molecular & Cellular Proteomics* **2024**, *23* (8), No. 100800.
(65) Guzman, U. H.; Martinez-Val, A.; Ye, Z.; Damoc, E.; Arrey, T. N.; Pashkova, A.; Renuse, S.; Denisov, E.; Petzoldt, J.; Peterson, A. C.; Harking, F.; Østergaard, O.; Rydbirk, R.; Aznar, S.; Stewart, H.; Xuan, Y.; Hermanson, D.; Horning, S.; Hock, C.; Makarov, A.; Zabrouskov, V.; Olsen, J. V. Ultra-Fast Label-Free Quantification and Comprehensive Proteome Coverage with Narrow-Window Data-Independent Acquisition. *Nat. Biotechnol.* **2024**, *42* (12), 1855−1866.
(66) Ludwig, C.; Gillet, L.; Rosenberger, G.; Amon, S.; Collins, B. C.; Aebersold, R. Data-independent Acquisition-based SWATH - MS for Quantitative Proteomics: A Tutorial. *Mol. Syst. Biol.* **2018**, *14* (8), No. e8126.
(67) Huang, P.; Liu, C.; Gao, W.; Chu, B.; Cai, Z.; Tian, R. Synergistic Optimization of Liquid Chromatography and Mass Spectrometry Parameters on Orbitrap Tribrid Mass Spectrometer for High Efficient Data-Dependent Proteomics. *Journal of mass spectrometry* **2021**, *56* (4), No. e4653.
(68) Lancaster, N. M.; Sinitcyn, P.; Forny, P.; Peters-Clarke, T. M.; Fecher, C.; Smith, A. J.; Shishkova, E.; Arrey, T. N.; Pashkova, A.; Robinson, M. L.; Arp, N.; Fan, J.; Hansen, J.; Galmozzi, A.; Serrano, L. R.; Rojas, J.; Gasch, A. P.; Westphall, M. S.; Stewart, H.; Hock, C.; Damoc, E.; Pagliarini, D. J.; Zabrouskov, V.; Coon, J. J. Fast and Deep Phosphoproteome Analysis with the Orbitrap Astral Mass Spectrometer. *Nat. Commun.* **2024**, *15* (1), 7016.
(69) Grgic, A.; Cuypers, E.; Dubois, L. J.; Ellis, S. R.; Heeren, R. M. A. MALDI MSI Protocol for Spatial Bottom-Up Proteomics at Single-Cell Resolution. *J. Proteome Res.* **2024**, *23* (12), 5372−5379.
(70) Houfani, A. A.; Foster, L. J. Review of the Real and Sometimes Hidden Costs in Proteomics Experimental Workflows. *Methods Mol. Biol.* **2022**, *2456*, 1−14.
(71) Guo, T.; Steen, J. A.; Mann, M. Mass-Spectrometry-Based Proteomics: From Single Cells to Clinical Applications. *Nature* **2025**, *638* (8052), 901−911.
(72) Meissner, F.; Geddes-McAlister, J.; Mann, M.; Bantscheff, M. The Emerging Role of Mass Spectrometry-Based Proteomics in Drug Discovery. *Nat. Rev. Drug Discov* **2022**, *21* (9), 637−654.
(73) Varshney, G. K.; Pei, W.; LaFave, M. C.; Idol, J.; Xu, L.; Gallardo, V.; Carrington, B.; Bishop, K.; Jones, M.; Li, M.; Harper, U.; Huang, S. C.; Prakash, A.; Chen, W.; Sood, R.; Ledin, J.; Burgess, S. M. High-Throughput Gene Targeting and Phenotyping in Zebrafish Using CRISPR/Cas9. *Genome Res.* **2015**, *25* (7), 1030−1042.
(74) Schenone, M.; Dančík, V.; Wagner, B. K.; Clemons, P. A. Target Identification and Mechanism of Action in Chemical Biology and Drug Discovery. *Nat. Chem. Biol.* **2013**, *9* (4), 232−240.

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

(75) Bochkov, Y. A.; Palmenberg, A. C.; Lee, W.-M.; Rathe, J. A.; Amineva, S. P.; Sun, X.; Pasic, T. R.; Jarjour, N. N.; Liggett, S. B.; Gern, J. E. Molecular Modeling, Organ Culture and Reverse Genetics for a Newly Identified Human Rhinovirus C. *Nat. Med.* **2011**, *17* (5), 627−632.
(76) Berdigaliyev, N.; Aljofan, M. An Overview of Drug Discovery and Development. *Future Med. Chem.* **2020**, *12* (10), 939−947.
(77) Emmerich, C. H.; Gamboa, L. M.; Hofmann, M. C. J.; Bonin-Andresen, M.; Arbach, O.; Schendel, P.; Gerlach, B.; Hempel, K.; Bespalov, A.; Dirnagl, U.; Parnham, M. J. Improving Target Assessment in Biomedical Research: The GOT-IT Recommendations. *Nat. Rev. Drug Discov* **2021**, *20* (1), 64−81.
(78) Amiri-Dashatan, N.; Koushki, M.; Abbaszadeh, H.-A.; Rostami-Nejad, M.; Rezaei-Tavirani, M. Proteomics Applications in Health: Biomarker and Drug Discovery and Food Industry. *Iran J. Pharm. Res.* **2018**, *17* (4), 1523−1536.
(79) Dueñas, M. E.; Peltier-Heap, R. E.; Leveridge, M.; Annan, R. S.; Büttner, F. H.; Trost, M. Advances in High-Throughput Mass Spectrometry in Drug Discovery. *EMBO Mol. Med.* **2023**, *15* (1), No. e14850.
(80) Mateus, A.; Määttä, T. A.; Savitski, M. M. Thermal Proteome Profiling: Unbiased Assessment of Protein State through Heat-Induced Stability Changes. *Proteome Sci.* **2016**, *15* (1), 13.
(81) Gaetani, M.; Sabatier, P.; Saei, A. A.; Beusch, C. M.; Yang, Z.; Lundström, S. L.; Zubarev, R. A. Proteome Integral Solubility Alteration: A High-Throughput Proteomics Assay for Target Deconvolution. *J. Proteome Res.* **2019**, *18* (11), 4027−4037.
(82) Fedorov, I. I.; Bubis, J. A.; Kazakova, E. M.; Lobas, A. A.; Ivanov, M. V.; Emekeeva, D. D.; Tarasova, I. A.; Nazarov, A. A.; Gorshkov, M. V. On the Utility of Ultrafast MS1-Only Proteomics in Drug Target Discovery Studies Based on Thermal Proteome Profiling Method. *Anal Bioanal Chem.* **2024**, *416* (18), 4083−4089.
(83) Figueroa-Navedo, A. M.; Ivanov, A. R. Experimental and Data Analysis Advances in Thermal Proteome Profiling. *Cell Reports Methods* **2024**, *4* (2), No. 100717.
(84) Aebersold, R.; Mann, M. Mass-Spectrometric Exploration of Proteome Structure and Function. *Nature* **2016**, *537* (7620), 347−355.
(85) Slobodyanyuk, M.; Bahcheli, A. T.; Klein, Z. P.; Bayati, M.; Strug, L. J.; Reimand, J. Directional Integration and Pathway Enrichment Analysis for Multi-Omics Data. *Nat. Commun.* **2024**, *15* (1), 5690.
(86) Sauer, P.; Bantscheff, M. Thermal Proteome Profiling for Drug Target Identification and Probing of Protein States. *Methods Mol. Biol.* **2023**, *2718*, 73−98.
(87) McCracken, N. A.; Peck Justice, S. A.; Wijeratne, A. B.; Mosley, A. L. Inflect: Optimizing Computational Workflows for Thermal Proteome Profiling Data Analysis. *J. Proteome Res.* **2021**, *20* (4), 1874−1888.
(88) Saei, A. A.; Gullberg, H.; Sabatier, P.; Beusch, C. M.; Johansson, K.; Lundgren, B.; Arvidsson, P. I.; Arnér, E. S. J.; Zubarev, R. A. Comprehensive Chemical Proteomics for Target Deconvolution of the Redox Active Drug Auranofin. *Redox Biol.* **2020**, *32*, No. 101491.
(89) Franken, H.; Mathieson, T.; Childs, D.; Sweetman, G. M. A.; Werner, T.; Tögel, I.; Doce, C.; Gade, S.; Bantscheff, M.; Drewes, G.; Reinhard, F. B. M.; Huber, W.; Savitski, M. M. Thermal Proteome Profiling for Unbiased Identification of Direct and Indirect Drug Targets Using Multiplexed Quantitative Mass Spectrometry. *Nat. Protoc* **2015**, *10* (10), 1567−1593.
(90) George, A. L.; Sidgwick, F. R.; Watt, J. E.; Martin, M. P.; Trost, M.; Marín-Rubio, J. L.; Dueñas, M. E. Comparison of Quantitative Mass Spectrometric Methods for Drug Target Identification by Thermal Proteome Profiling. *J. Proteome Res.* **2023**, *22* (8), 2629−2640.
(91) Marín-Rubio, J. L.; Peltier-Heap, R. E.; Dueñas, M. E.; Heunis, T.; Dannoura, A.; Inns, J.; Scott, J.; Simpson, A. J.; Blair, H. J.; Heidenreich, O.; Allan, J. M.; Watt, J. E.; Martin, M. P.; Saxty, B.; Trost, M. A Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Assay Identifies Nilotinib as an Inhibitor of Inflammation in Acute Myeloid Leukemia. *J. Med. Chem.* **2022**, *65* (18), 12014−12030.
(92) Deng, H.; Lei, Q.; Wu, Y.; He, Y.; Li, W. Activity-Based Protein Profiling: Recent Advances in Medicinal Chemistry. *Eur. J. Med. Chem.* **2020**, *191*, No. 112151.
(93) Xu, J.; Li, X.; Ding, K.; Li, Z. Applications of Activity-Based Protein Profiling (ABPP) and Bioimaging in Drug Discovery. *Chem. Asian J.* **2020**, *15* (1), 34−41.
(94) Jones, H. B. L.; Heilig, R.; Davis, S.; Fischer, R.; Kessler, B. M.; Pinto-Fernández, A. ABPP-HT*—Deep Meets Fast for Activity-Based Profiling of Deubiquitylating Enzymes Using Advanced DIA Mass Spectrometry Methods. *Int. J. Mol. Sci.* **2022**, *23* (6), 3263.
(95) Porta, E. O. J.; Steel, P. G. Activity-Based Protein Profiling: A Graphical Review. *Current Research in Pharmacology and Drug Discovery* **2023**, *5*, No. 100164.
(96) Speers, A. E.; Cravatt, B. F. Profiling Enzyme Activities In Vivo Using Click Chemistry Methods. *Chem. Biol.* **2004**, *11* (4), 535−546.
(97) Krysiak, J. M.; Kreuzer, J.; Macheroux, P.; Hermetter, A.; Sieber, S. A.; Breinbauer, R. Activity-Based Probes for Studying the Activity of Flavin-Dependent Oxidases and for the Protein Target Profiling of Monoamine Oxidase Inhibitors. *Angew. Chem., Int. Ed.* **2012**, *51* (28), 7035−7040.
(98) McCloud, R. L.; Yuan, K.; Mahoney, K. E.; Bai, D. L.; Shabanowitz, J.; Ross, M. M.; Hunt, D. F.; Hsu, K.-L. Direct Target Site Identification of a Sulfonyl−Triazole Covalent Kinase Probe by LC-MS Chemical Proteomics. *Anal. Chem.* **2021**, *93* (35), 11946−11955.
(99) Zhao, Q.; Ouyang, X.; Wan, X.; Gajiwala, K. S.; Kath, J. C.; Jones, L. H.; Burlingame, A. L.; Taunton, J. Broad-Spectrum Kinase Profiling in Live Cells with Lysine-Targeted Sulfonyl Fluoride Probes. *J. Am. Chem. Soc.* **2017**, *139* (2), 680−685.
(100) Tan, M. S. Y.; Davison, D.; Sanchez, M. I.; Anderson, B. M.; Howell, S.; Snijders, A. P.; Edgington-Mitchell, L. E.; Deu, E. Correction: Novel Broad-Spectrum Activity-Based Probes to Profile Malarial Cysteine Proteases. *PLoS One* **2020**, *15* (3), No. e0231231.
(101) Porta, E. O. J.; Isern, J. A.; Kalesh, K.; Steel, P. G. Discovery of Leishmania Druggable Serine Proteases by Activity-Based Protein Profiling. *Front Pharmacol* **2022**, *13*, No. 929493.
(102) Choi, K.-M.; Kim, S.-J.; Ji, M.-J.; Kim, E.; Kim, J.-S.; Park, H.; Kim, J.-Y. Activity-Based Protein Profiling and Global Proteome Analysis Reveal MASTL as a Potential Therapeutic Target in Gastric Cancer. *Cell Communication and Signaling* **2024**, *22* (1), 397.
(103) Burton, N. R.; Kim, P.; Backus, K. M. Photoaffinity Labelling Strategies for Mapping the Small Molecule−Protein Interactome. *Org. Biomol Chem.* **2021**, *19* (36), 7792−7809.
(104) Homan, R. A.; Lapek, J. D.; Woo, C. M.; Niessen, S.; Jones, L. H.; Parker, C. G. Photoaffinity Labelling with Small Molecules. *Nature Reviews Methods Primers* **2024**, *4* (1), 30.
(105) Lee, W.; Huang, Z.; am Ende, C. W.; Seneviratne, U. Protocol for Clickable Photoaffinity Labeling and Quantitative Chemical Proteomics. *STAR Protoc* **2021**, *2* (2), No. 100593.
(106) Murale, D. P.; Hong, S. C.; Haque, Md. M.; Lee, J.-S. Photo-Affinity Labeling (PAL) in Chemical Proteomics: A Handy Tool to Investigate Protein-Protein Interactions (PPIs). *Proteome Sci.* **2016**, *15* (1), 14.
(107) Chuang, V.; Otagiri, M. Photoaffinity Labeling of Plasma Proteins. *Molecules* **2013**, *18* (11), 13831−13859.
(108) Ruoho, A. E.; Kiefer, H.; Roeder, P. E.; Singer, S. J. The Mechanism of Photoaffinity Labeling. *Proc. Natl. Acad. Sci. U. S. A.* **1973**, *70* (9), 2567−2571.
(109) Tay, N. E. S.; Ryu, K. A.; Weber, J. L.; Olow, A. K.; Cabanero, D. C.; Reichman, D. R.; Oslund, R. C.; Fadeyi, O. O.; Rovis, T. Targeted Activation in Localized Protein Environments via Deep Red Photoredox Catalysis. *Nat. Chem.* **2023**, *15* (1), 101−109.
(110) Sumranjit, J.; Chung, S. Recent Advances in Target Characterization and Identification by Photoaffinity Probes. *Molecules* **2013**, *18* (9), 10425−10451.
(111) Karaj, E.; Sindi, S. H.; Viranga Tillekeratne, L. M. Photoaffinity Labeling and Bioorthogonal Ligation: Two Critical Tools for Designing "Fish Hooks" to Scout for Target Proteins. *Bioorg. Med. Chem.* **2022**, *62*, No. 116721.

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

(112) Belanger, J. M.; Raviv, Y.; Viard, M.; Jason de la Cruz, M.; Nagashima, K.; Blumenthal, R. Characterization of the Effects of Aryl-azido Compounds and UVA Irradiation on the Viral Proteins and Infectivity of Human Immunodeficiency Virus Type 1. *Photochem. Photobiol.* **2010**, *86* (5), 1099−1108.
(113) Sakurai, K.; Ozawa, S.; Yamada, R.; Yasui, T.; Mizuno, S. Comparison of the Reactivity of Carbohydrate Photoaffinity Probes with Different Photoreactive Groups. *ChemBioChem.* **2014**, *15* (10), 1399−1403.
(114) Vervacke, J. S.; Funk, A. L.; Wang, Y.-C.; Strom, M.; Hrycyna, C. A.; Distefano, M. D. Diazirine-Containing Photoactivatable Isoprenoid: Synthesis and Application in Studies with Isoprenylcysteine Carboxyl Methyltransferase. *J. Org. Chem.* **2014**, *79* (5), 1971−1978.
(115) Tabana, Y.; Babu, D.; Fahlman, R.; Siraki, A. G.; Barakat, K. Target Identification of Small Molecules: An Overview of the Current Applications in Drug Discovery. *BMC Biotechnol* **2023**, *23* (1), 44.
(116) Dziuba, D.; Hoffmann, J.; Hentze, M. W.; Schultz, C. A Genetically Encoded Diazirine Analogue for RNA−Protein Photo-crosslinking. *ChemBioChem.* **2020**, *21* (1−2), 88−93.
(117) Kotnik, K.; Kosjek, T.; Žegura, B.; Filipič, M.; Heath, E. Photolytic Fate and Genotoxicity of Benzophenone-Derived Compounds and Their Photodegradation Mixtures in the Aqueous Environment. *Chemosphere* **2016**, *147*, 114−123.
(118) Imamović, B.; Trebše, P.; Omeragić, E.; Bečić, E.; Pečet, A.; Dedić, M. Stability and Removal of Benzophenone-Type UV Filters from Water Matrices by Advanced Oxidation Processes. *Molecules* **2022**, *27* (6), 1874.
(119) West, A. V.; Muncipinto, G.; Wu, H.-Y.; Huang, A. C.; Labenski, M. T.; Jones, L. H.; Woo, C. M. Labeling Preferences of Diazirines with Protein Biomolecules. *J. Am. Chem. Soc.* **2021**, *143* (17), 6691−6700.
(120) Muranaka, H.; Momose, T.; Handa, C.; Ozawa, T. Photoaffinity Labeling of the Human A_{2A} Adenosine Receptor and Cross-Link Position Analysis by Mass Spectrometry. *ACS Med. Chem. Lett.* **2017**, *8* (6), 660−665.
(121) Bon, C.; Goretzki, B.; Flamme, M.; Shelton, C.; Davis, H.; Lima, F.; Garcia, F.; Brittain, S.; Brocklehurst, C. E. Oxadiazolines as Photoreleasable Labels for Drug Target Identification. *J. Am. Chem. Soc.* **2024**, *146* (39), 26759−26765.
(122) Takamura, A.; Thuy-Boun, P. S.; Kitamura, S.; Han, Z.; Wolan, D. W. A Photoaffinity Probe That Targets Folate-Binding Proteins. *Bioorg. Med. Chem. Lett.* **2021**, *40*, No. 127903.
(123) Kaneda, M.; Masuda, S.; Tomohiro, T.; Hatanaka, Y. A Simple and Efficient Photoaffinity Method for Proteomics of GTP-Binding Proteins. *ChemBioChem.* **2007**, *8* (6), 595−598.
(124) George Cisar, E. A.; Nguyen, N.; Rosen, H. A GTP Affinity Probe for Proteomics Highlights Flexibility in Purine Nucleotide Selectivity. *J. Am. Chem. Soc.* **2013**, *135* (12), 4676−4679.
(125) Zhang, L.; Zhang, Y.; Dong, J.; Liu, J.; Zhang, L.; Sun, H. Design and Synthesis of Novel Photoaffinity Probes for Study of the Target Proteins of Oleanolic Acid. *Bioorg. Med. Chem. Lett.* **2012**, *22* (2), 1036−1039.
(126) Tulloch, L. B.; Menzies, S. K.; Fraser, A. L.; Gould, E. R.; King, E. F.; Zacharova, M. K.; Florence, G. J.; Smith, T. K. Photo-Affinity Labelling and Biochemical Analyses Identify the Target of Trypanocidal Simplified Natural Product Analogues. *PLoS Negl Trop Dis* **2017**, *11* (9), No. e0005886.
(127) Ábrányi-Balogh, P.; Bajusz, D.; Orgován, Z.; Keeley, A. B.; Petri, L.; Péczka, N.; Szalai, T. V.; Pálfy, G.; Gadanecz, M.; Grant, E. K.; Imre, T.; Takács, T.; Ranđelović, I.; Baranyi, M.; Marton, A.; Schlosser, G.; Ashraf, Q. F.; de Araujo, E. D.; Karancsi, T.; Buday, L.; Tóvári, J.; Perczel, A.; Bush, J. T.; Keserű, G. M. Mapping Protein Binding Sites by Photoreactive Fragment Pharmacophores. *Commun. Chem.* **2024**, *7* (1), 168.
(128) Reber, V.; Gstaiger, M. *Target Deconvolution by Limited Proteolysis Coupled to Mass Spectrometry* **2023**, *2706*, 177−190.
(129) Piazza, I.; Beaton, N.; Bruderer, R.; Knobloch, T.; Barbisan, C.; Chandat, L.; Sudau, A.; Siepe, I.; Rinner, O.; de Souza, N.; Picotti, P.; Reiter, L. A Machine Learning-Based Chemoproteomic Approach to Identify Drug Targets and Binding Sites in Complex Proteomes. *Nat. Commun.* **2020**, *11* (1), 4200.
(130) Malinovska, L.; Cappelletti, V.; Kohler, D.; Piazza, I.; Tsai, T.-H.; Pepelnjak, M.; Stalder, P.; Dörig, C.; Sesterhenn, F.; Elsässer, F.; Kralickova, L.; Beaton, N.; Reiter, L.; de Souza, N.; Vitek, O.; Picotti, P. Proteome-Wide Structural Changes Measured with Limited Proteolysis-Mass Spectrometry: An Advanced Protocol for High-Throughput Applications. *Nat. Protoc* **2023**, *18* (3), 659−682.
(131) Holfeld, A.; Quast, J.-P.; Bruderer, R.; Reiter, L.; de Souza, N.; Picotti, P. Limited Proteolysis−Mass Spectrometry to Identify Metabolite−Protein Interactions. *Methods Mol. Biol.* **2023**, *2554*, 69−89.
(132) Shuken, S. R.; Rutledge, J.; Iram, T.; Losada, P. M.; Wilson, E. N.; Andreasson, K. I.; Leib, R. D.; Wyss-Coray, T. Limited Proteolysis−Mass Spectrometry Reveals Aging-Associated Changes in Cerebrospinal Fluid Protein Abundances and Structures. *Nat. Aging* **2022**, *2* (5), 379−388.
(133) Zilu, Z.; Zhang, X.; Dong, X.; Hong, Z. Decoding Protein Dynamics with Limited Proteolysis Coupled to Mass Spectrometry: A Comprehensive Review. *J. Pharm. Anal* **2025**, No. 101319.
(134) Van Leene, C.; Van Moortel, L.; De Bosscher, K.; Gevaert, K. Exploring Protein Conformations with Limited Proteolysis Coupled to Mass Spectrometry. *Trends Biochem. Sci.* **2025**, *50* (2), 143−155.
(135) Nagel, L.; Grossbach, J.; Cappelletti, V.; Dörig, C.; Picotti, P.; Beyer, A. Analysis of Limited Proteolysis-Coupled Mass Spectrometry Data. *Molecular & Cellular Proteomics* **2025**, *24* (4), No. 100934.
(136) Koudelka, T.; Bassot, C.; Piazza, I. Benchmarking of Quantitative Proteomics Workflows for Limited Proteolysis Mass Spectrometry. *Molecular & Cellular Proteomics* **2025**, *24* (4), No. 100945.
(137) Yin, Z.; Li, G.; Zhong, Q.; Zhang, X.; Ou, R.; Shen, H.; Huang, J.; Mahati, S.; Liu, Z.; Zhu, Y.; Zhang, Q.; Liu, S. Cinobufagin Overcomes Bortezomib Resistance in Multiple Myeloma Strains by Targeting SEC62/TRPM4-Mediated NECSO. *Phytomedicine* **2025**, *147*, No. 157171.
(138) Peng, Y.-C.; He, Z.-J.; Yin, L.-C.; Pi, H.-F.; Jiang, Y.; Li, K.-Y.; Tian, L.; Xie, J.; Zhang, J.-B.; Li, C.-Y.; Feng, G.-Y.; Wang, K.; Zhou, D.-Z.; Xie, X.-W.; Zhang, Z.-Y.; Fan, T.-F. Sanguinarine Suppresses Oral Squamous Cell Carcinoma Progression by Targeting the PKM2/TFEB Aix to Inhibit Autophagic Flux. *Phytomedicine* **2025**, *136*, No. 156337.
(139) Capuano, A.; D'Urso, G.; Aliberti, M.; Ruggiero, D.; Terracciano, S.; Festa, C.; Tosco, A.; Chini, M. G.; Lauro, G.; Bifulco, G.; Casapullo, A. Chemoproteomics Reveals USP5 (Ubiquitin Carboxyl-Terminal Hydrolase 5) as Promising Target of the Marine Polyketide Gracilioether A. *Mar Drugs* **2024**, *22* (1), 41.
(140) Morretta, E.; Capuano, A.; D'Urso, G.; Voli, A.; Mozzicafreddo, M.; Di Gaetano, S.; Capasso, D.; Sala, M.; Scala, M. C.; Campiglia, P.; Piccialli, V.; Casapullo, A. Identification of Mortalin as the Main Interactor of Mycalin A, a Poly-Brominated C-15 Acetogenin Sponge Metabolite, by MS-Based Proteomics. *Mar Drugs* **2024**, *22* (2), 52.
(141) Bader, J. M.; Albrecht, V.; Mann, M. MS-Based Proteomics of Body Fluids: The End of the Beginning. *Molecular & Cellular Proteomics* **2023**, *22* (7), No. 100577.
(142) Beimers, W. F.; Overmyer, K. A.; Sinitcyn, P.; Lancaster, N. M.; Quarmby, S. T.; Coon, J. J. Technical Evaluation of Plasma Proteomics Technologies. *J. Proteome Res.* **2025**, *24* (6), 3074−3087.
(143) Zhu, Y. Plasma/Serum Proteomics Based on Mass Spectrometry. *Protein Pept Lett.* **2024**, *31* (3), 192−208.
(144) Peng, Z.; Munige, S. W.; Bhusal, D.; Yang, I. L.; Yang, Z.; Ahsan, N. A Cross-Species Proteomic Assessment of Cost-Effective Platforms for Depleting High-Abundant Proteins from Blood Serum. *BMC Methods* **2025**, *2* (1), 21.
(145) Cao, X.; Sandberg, A.; Araújo, J. E.; Cvetkovski, F.; Berglund, E.; Eriksson, L. E.; Pernemalm, M. Evaluation of Spin Columns for Human Plasma Depletion to Facilitate MS-Based Proteomics Analysis of Plasma. *J. Proteome Res.* **2021**, *20* (9), 4610−4620.
(146) Zhou, Y.; Zheng, H.; Tan, Z.; Kang, E.; Xue, P.; Li, X.; Guan, F. Optimizing and Integrating Depletion and Precipitation Methods for

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

Plasma Proteomics through Data-Independent Acquisition-Mass Spectrometry. *Journal of Chromatography B* **2024**, *1235*, No. 124046.
(147) Mazidi, M.; Wright, N.; Yao, P.; Kartsonaki, C.; Millwood, I. Y.; Fry, H.; Said, S.; Pozarickij, A.; Pei, P.; Chen, Y.; Avery, D.; Du, H.; Schmidt, D. V.; Yang, L.; Lv, J.; Yu, C.; Chen, J.; Hill, M.; Holmes, M. V.; Howson, J. M. M.; Peto, R.; Collins, R.; Bennett, D. A.; Walters, R. G.; Li, L.; Clarke, R.; Chen, Z.; Chen, J.; Chen, Z.; Clarke, R.; Collins, R.; Li, L.; Wang, C.; Lv, J.; Peto, R.; Walters, R.; Avery, D.; Barnard, M.; Bennett, D.; Boxall, R.; Burgess, S.; Chan, K. H.; Chen, Y.; Chen, Z.; Clarke, J.; Clarke, R.; Du, H.; Mohamed, A. E.; Fry, H.; Gilbert, S.; Im, P. K.; Iona, A.; Kakkoura, M.; Kartsonaki, C.; Lam, H.; Lin, K.; Liu, J.; Mazidi, M.; Millwood, I.; Morris, S.; Nie, Q.; Pozarickij, A.; Ryder, P.; Said, S.; Schmidt, D.; Stevens, B.; Turnbull, I.; Walters, R.; Wang, B.; Wang, L.; Wright, N.; Yang, L.; Yang, X.; Yao, P.; Han, X.; Hou, C.; Xia, Q.; Liu, C.; Lv, J.; Pei; Sun, D.; Yu, C.; Chen, N.; Liu, D.; Tang, Z.; Chen, N.; Jiang, Q.; Lan, J.; Li, M.; Liu, Y.; Meng, F.; Meng, J.; Pan, R.; Qin, Y.; Wang, P.; Wang, S.; Wei, L.; Zhou, L.; Dong, C.; Ge, P.; Ren, X.; Li, Z.; Mao, E.; Wang, T.; Zhang, H.; Zhang, X.; Chen, J.; Hu, X.; Wang, X.; Guo, Z.; Li, H.; Li, Y.; Weng, M.; Wu, S.; Yan, S.; Zou, M.; Zhou, X.; Guo, Z.; Kang, Q.; Li, Y.; Yu, B.; Xu, Q.; Chang, L.; Fan, L.; Feng, S.; Zhang, D.; Zhou, G.; Gao, Y.; He, T.; He, P.; Hu, C.; Sun, H.; Zhang, X.; Chen, B.; Fu, Z.; Huang, Y.; Liu, H.; Xu, Q.; Yin, L.; Long, H.; Xu, X.; Zhang, H.; Zhang, L.; Su, J.; Tao, R.; Wu, M.; Yang, J.; Zhou, J.; Zhou, Y.; Hu, Y.; Hua, Y.; Jin, J.; Liu, F.; Liu, J.; Lu, Y.; Ma, L.; Tang, A.; Zhang, J.; Cheng, L.; Du, R.; Gao, R.; Li, F.; Li, S.; Liu, Y.; Ning, F.; Pang, Z.; Sun, X.; Tian, X.; Wang, S.; Zhai, Y.; Zhang, H.; Hou, W.; Lv, S.; Wang, J.; Chen, X.; Wu, X.; Zhang, N.; Zhou, W.; Chen, X.; Li, J.; Liu, J.; Luo, G.; Sun, Q.; Zhong, X.; Gong, W.; Hu, R.; Wang, H.; Wang, M.; Yu, M.; Chen, L.; Gu, Q.; Pan, D.; Wang, C.; Xie, K.; Zhang, X. Plasma Proteomics to Identify Drug Targets for Ischemic Heart Disease. *J. Am. Coll Cardiol* **2023**, *82* (20), 1906−1920.
(148) Rabin, A.; Bello, E.; Kumar, S.; Zeki, D. A.; Afshari, K.; Deshpande, M.; Francis, N.; Khalighinejad, F.; Umeton, R.; Radu, I.; Qutab, F.; Kwong, D.; Kurban, M.; Hemond, C.; Richmond, J. M.; Ionete, C. Targeted Proteomics of Cerebrospinal Fluid in Treatment Naïve Multiple Sclerosis Patients Identifies Immune Biomarkers of Clinical Phenotypes. *Sci. Rep* **2024**, *14* (1), 21793.
(149) Fernández-Irigoyen, J.; Santamaría, E. Recent Advances in Human Cerebrospinal Fluid Proteomics. *Methods Mol. Biol.* **2025**, *2914*, 3−12.
(150) Imam, F.; Saloner, R.; Vogel, J. W.; Krish, V.; Abdel-Azim, G.; Ali, M.; An, L.; Anastasi, F.; Bennett, D.; Pichet Binette, A.; Boxer, A. L.; Bringmann, M.; Burns, J. M.; Cruchaga, C.; Dage, J. L.; Farinas, A.; Ferrucci, L.; Finney, C. A.; Frasier, M.; Hansson, O.; Hohman, T. J.; Johnson, E. C. B.; Kivimaki, M.; Korologou-Linden, R.; Ruiz Laza, A.; Levey, A. I.; Liepelt-Scarfone, I.; Lu, L.; Mattsson-Carlgren, N.; Middleton, L. T.; Nho, K.; Oh, H. S.-H.; Petersen, R. C.; Reiman, E. M.; Robinson, O.; Rothstein, J. D.; Saykin, A. J.; Shvetcov, A.; Slawson, C.; Smets, B.; Suárez-Calvet, M.; Tijms, B. M.; Timmers, M.; Vieira, F.; Vilor-Tejedor, N.; Visser, P. J.; Walker, K. A.; Winchester, L. M.; Wyss-Coray, T.; Yang, C.; Bose, N.; Lovestone, S. The Global Neurodegeneration Proteomics Consortium: Biomarker and Drug Target Discovery for Common Neurodegenerative Diseases and Aging. *Nat. Med.* **2025**, *31* (8), 2556−2566.
(151) Mavrogeorgis, E.; Mischak, H.; Beige, J.; Latosinska, A.; Siwy, J. Understanding Glomerular Diseases through Proteomics. *Expert Rev. Proteomics* **2021**, *18* (2), 137−157.
(152) Jiao, L.; Dong, H.; Qin, Y.; Zhu, J.; Wu, P.; Liu, J.; Cao, Y.; Wu, C.; Zhang, Y.; Cao, F.; Li, F.; Zhu, H. Comparisons of Whole Saliva and Cell Free Saliva by DIA-Based Proteome Profiling. *Proteomics Clin Appl.* **2025**, *19* (1), No. e202400031.
(153) Scannell, J. W.; Bosley, J.; Hickman, J. A.; Dawson, G. R.; Truebel, H.; Ferreira, G. S.; Richards, D.; Treherne, J. M. Predictive Validity in Drug Discovery: What It Is, Why It Matters and How to Improve It. *Nat. Rev. Drug Discov* **2022**, *21* (12), 915−931.
(154) Coyle, J.; Walser, R. Applied Biophysical Methods in Fragment-Based Drug Discovery. *SLAS Discovery* **2020**, *25* (5), 471−490.
(155) Prudent, R.; Annis, D. A.; Dandliker, P. J.; Ortholand, J.-Y.; Roche, D. Exploring New Targets and Chemical Space with Affinity Selection-Mass Spectrometry. *Nat. Rev. Chem.* **2021**, *5* (1), 62−71.
(156) Hevener, K. E.; Pesavento, R.; Ren, J.; Lee, H.; Ratia, K.; Johnson, M. E. Hit-to-Lead: Hit Validation and Assessment. *Methods Enzymol.* **2018**, *610*, 265−309.
(157) Kirsch, P.; Hartman, A. M.; Hirsch, A. K. H.; Empting, M. Concepts and Core Principles of Fragment-Based Drug Design. *Molecules* **2019**, *24* (23), 4309.
(158) Doak, B. C.; Norton, R. S.; Scanlon, M. J. The Ways and Means of Fragment-Based Drug Design. *Pharmacol Ther* **2016**, *167*, 28−37.
(159) Jacquemard, C.; Kellenberger, E. A Bright Future for Fragment-Based Drug Discovery: What Does It Hold? *Expert Opin Drug Discov* **2019**, *14* (5), 413−416.
(160) Mureddu, L. G.; Vuister, G. W. Fragment-Based Drug Discovery by NMR. Where Are the Successes and Where Can It Be Improved? *Front Mol. Biosci* **2022**, *9*, No. 834453.
(161) Gavriilidou, A. F. M.; Holding, F. P.; Coyle, J. E.; Zenobi, R. Application of Native ESI-MS to Characterize Interactions between Compounds Derived from Fragment-Based Discovery Campaigns and Two Pharmaceutically Relevant Proteins. *SLAS Discovery* **2018**, *23* (9), 951−959.
(162) Vaaltyn, M. C.; Mateos-Jimenez, M.; Müller, R.; Mackay, C. L.; Edkins, A. L.; Clarke, D. J.; Veale, C. G. L. Native Mass Spectrometry-Guided Screening Identifies Hit Fragments for HOP-HSP90 PPI Inhibition**. *ChemBioChem.* **2022**, *23* (21), No. e202200322.
(163) Bissaro, M.; Sturlese, M.; Moro, S. The Rise of Molecular Simulations in Fragment-Based Drug Design (FBDD): An Overview. *Drug Discov Today* **2020**, *25* (9), 1693−1701.
(164) Bon, M.; Bilsland, A.; Bower, J.; McAulay, K. Fragment-based Drug Discovery—the Importance of High-quality Molecule Libraries. *Mol. Oncol* **2022**, *16* (21), 3761−3777.
(165) Ohlson, S.; Duong-Thi, M.-D. Fragment Screening for Drug Leads by Weak Affinity Chromatography (WAC-MS). *Methods* **2018**, *146*, 26−38.
(166) Kavanagh, M. E.; Coyne, A. G.; McLean, K. J.; James, G. G.; Levy, C. W.; Marino, L. B.; de Carvalho, L. P. S.; Chan, D. S. H.; Hudson, S. A.; Surade, S.; Leys, D.; Munro, A. W.; Abell, C. Fragment-Based Approaches to the Development of *Mycobacterium Tuberculosis* CYP121 Inhibitors. *J. Med. Chem.* **2016**, *59* (7), 3272−3302.
(167) Götze, M.; Iacobucci, C.; Ihling, C. H.; Sinz, A. A Simple Cross-Linking/Mass Spectrometry Workflow for Studying System-Wide Protein Interactions. *Anal. Chem.* **2019**, *91* (15), 10236−10244.
(168) Mendes, M. L.; Fischer, L.; Chen, Z. A.; Barbon, M.; O'Reilly, F. J.; Giese, S. H.; Bohlke-Schneider, M.; Belsom, A.; Dau, T.; Combe, C. W.; Graham, M.; Eisele, M. R.; Baumeister, W.; Speck, C.; Rappsilber, J. An Integrated Workflow for Crosslinking Mass Spectrometry. *Mol. Syst. Biol.* **2019**, *15* (9), No. e8994.
(169) O'Reilly, F. J.; Rappsilber, J. Cross-Linking Mass Spectrometry: Methods and Applications in Structural, Molecular and Systems Biology. *Nat. Struct Mol. Biol.* **2018**, *25* (11), 1000−1008.
(170) Piersimoni, L.; Kastritis, P. L.; Arlt, C.; Sinz, A. Cross-Linking Mass Spectrometry for Investigating Protein Conformations and Protein−Protein Interactions—A Method for All Seasons. *Chem. Rev.* **2022**, *122* (8), 7500−7531.
(171) Gavriilidou, A. F. M.; Sokratous, K.; Yen, H.-Y.; De Colibus, L. High-Throughput Native Mass Spectrometry Screening in Drug Discovery. *Front Mol. Biosci* **2022**, *9*, No. 837901.
(172) Yu, C.; Huang, L. New Advances in Cross-Linking Mass Spectrometry toward Structural Systems Biology. *Curr. Opin Chem. Biol.* **2023**, *76*, No. 102357.
(173) Costantino, L.; Ferrari, S.; Santucci, M.; Salo-Ahen, O. M. H.; Carosati, E.; Franchini, S.; Lauriola, A.; Pozzi, C.; Trande, M.; Gozzi, G.; Saxena, P.; Cannazza, G.; Losi, L.; Cardinale, D.; Venturelli, A.; Quotadamo, A.; Linciano, P.; Tagliazucchi, L.; Moschella, M. G.; Guerrini, R.; Pacifico, S.; Luciani, R.; Genovese, F.; Henrich, S.; Alboni, S.; Santarem, N.; da Silva Cordeiro, A.; Giovannetti, E.; Peters, G. J.; Pinton, P.; Rimessi, A.; Cruciani, G.; Stroud, R. M.; Wade, R. C.; Mangani, S.; Marverti, G.; D'Arca, D.; Ponterini, G.; Costi, M. P.

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

Destabilizers of the Thymidylate Synthase Homodimer Accelerate Its Proteasomal Degradation and Inhibit Cancer Growth. *Elife* **2022**, *11*, No. e73862.
(174) Fan, M.; Lu, W.; Che, J.; Kwiatkowski, N. P.; Gao, Y.; Seo, H.-S.; Ficarro, S. B.; Gokhale, P. C.; Liu, Y.; Geffken, E. A.; Lakhani, J.; Song, K.; Kuljanin, M.; Ji, W.; Jiang, J.; He, Z.; Tse, J.; Boghossian, A. S.; Rees, M. G.; Ronan, M. M.; Roth, J. A.; Mancias, J. D.; Marto, J. A.; Dhe-Paganon, S.; Zhang, T.; Gray, N. S. Covalent Disruptor of YAP-TEAD Association Suppresses Defective Hippo Signaling. *Elife* **2022**, *11*, No. e78810.
(175) Birhanu, A. G. Mass Spectrometry-Based Proteomics as an Emerging Tool in Clinical Laboratories. *Clin Proteomics* **2023**, *20* (1), 32.
(176) Narang, D.; Lento, C.; Wilson, D. J. HDX-MS: An Analytical Tool to Capture Protein Motion in Action. *Biomedicines* **2020**, *8* (7), 224.
(177) Burckhardt, C. J.; Minna, J. D.; Danuser, G. Co-Immunoprecipitation and Semi-Quantitative Immunoblotting for the Analysis of Protein-Protein Interactions. *STAR Protoc* **2021**, *2* (3), No. 100644.
(178) Valdés, A.; Bergström Lind, S. Mass Spectrometry-Based Analysis of Time-Resolved Proteome Quantification. *Proteomics* **2020**, *20* (9), No. e1800425.
(179) Zhai, L.; Chen, K.; Hao, B.; Tan, M. Proteomic Characterization of Post-Translational Modifications in Drug Discovery. *Acta Pharmacol Sin* **2022**, *43* (12), 3112−3129.
(180) Chang, Y.-C.; Gnann, C.; Steimbach, R. R.; Bayer, F. P.; Lechner, S.; Sakhteman, A.; Abele, M.; Zecha, J.; Trendel, J.; The, M.; Lundberg, E.; Miller, A. K.; Kuster, B. Decrypting Lysine Deacetylase Inhibitor Action and Protein Modifications by Dose-Resolved Proteomics. *Cell Rep* **2024**, *43* (6), No. 114272.
(181) Saei, A. A.; Beusch, C. M.; Chernobrovkin, A.; Sabatier, P.; Zhang, B.; Tokat, Ü. G.; Stergiou, E.; Gaetani, M.; Végvári, Á.; Zubarev, R. A. ProTargetMiner as a Proteome Signature Library of Anticancer Molecules for Functional Discovery. *Nat. Commun.* **2019**, *10* (1), 5715.
(182) Ruprecht, B.; Di Bernardo, J.; Wang, Z.; Mo, X.; Ursu, O.; Christopher, M.; Fernandez, R. B.; Zheng, L.; Dill, B. D.; Wang, H.; Xu, Y.; Liaw, A.; Mortison, J. D.; Siriwardana, N.; Andresen, B.; Glick, M.; Tata, J. R.; Kutilek, V.; Cornella-Taracido, I.; Chi, A. A Mass Spectrometry-Based Proteome Map of Drug Action in Lung Cancer Cell Lines. *Nat. Chem. Biol.* **2020**, *16* (10), 1111−1119.
(183) El-Baba, C.; Ayache, Z.; Goli, M.; Hayar, B.; Kawtharani, Z.; Pisano, C.; Kobeissy, F.; Mechref, Y.; Darwiche, N. The Antitumor Effect of the DNA Polymerase Alpha Inhibitor ST1926 in Glioblastoma: A Proteomics Approach. *Int. J. Mol. Sci.* **2023**, *24* (18), No. 14069.
(184) Eckert, S.; Berner, N.; Kramer, K.; Schneider, A.; Müller, J.; Lechner, S.; Brajkovic, S.; Sakhteman, A.; Graetz, C.; Fackler, J.; Dudek, M.; Pfaffl, M. W.; Knolle, P.; Wilhelm, S.; Kuster, B. Decrypting the Molecular Basis of Cellular Drug Phenotypes by Dose-Resolved Expression Proteomics. *Nat. Biotechnol.* **2025**, *43*, 406.
(185) Masson, G. R.; Burke, J. E.; Ahn, N. G.; Anand, G. S.; Borchers, C.; Brier, S.; Bou-Assaf, G. M.; Engen, J. R.; Englander, S. W.; Faber, J.; Garlish, R.; Griffin, P. R.; Gross, M. L.; Guttman, M.; Hamuro, Y.; Heck, A. J. R.; Houde, D.; Iacob, R. E.; Jørgensen, T. J. D.; Kaltashov, I. A.; Klinman, J. P.; Konermann, L.; Man, P.; Mayne, L.; Pascal, B. D.; Reichmann, D.; Skehel, M.; Snijder, J.; Strutzenberg, T. S.; Underbakke, E. S.; Wagner, C.; Wales, T. E.; Walters, B. T.; Weis, D. D.; Wilson, D. J.; Wintrode, P. L.; Zhang, Z.; Zheng, J.; Schriemer, D. C.; Rand, K. D. Recommendations for Performing, Interpreting and Reporting Hydrogen Deuterium Exchange Mass Spectrometry (HDX-MS) Experiments. *Nat. Methods* **2019**, *16* (7), 595−602.
(186) Hamuro, Y. Tutorial: Chemistry of Hydrogen/Deuterium Exchange Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2021**, *32* (1), 133−151.
(187) Oganesyan, I.; Lento, C.; Wilson, D. J. Contemporary Hydrogen Deuterium Exchange Mass Spectrometry. *Methods* **2018**, *144*, 27−42.
(188) Sun, H.; Ma, L.; Wang, L.; Xiao, P.; Li, H.; Zhou, M.; Song, D. Research Advances in Hydrogen−Deuterium Exchange Mass Spectrometry for Protein Epitope Mapping. *Anal Bioanal Chem.* **2021**, *413* (9), 2345−2359.
(189) Kostyukevich, Y.; Acter, T.; Zherebker, A.; Ahmed, A.; Kim, S.; Nikolaev, E. Hydrogen/Deuterium Exchange in Mass Spectrometry. *Mass Spectrom Rev.* **2018**, *37* (6), 811−853.
(190) Filandr, F.; Sarpe, V.; Raval, S.; Crowder, D. A.; Khan, M. F.; Douglas, P.; Coales, S.; Viner, R.; Syed, A.; Tainer, J. A.; Lees-Miller, S. P.; Schriemer, D. C. Automating Data Analysis for Hydrogen/Deuterium Exchange Mass Spectrometry Using Data-Independent Acquisition Methodology. *Nat. Commun.* **2024**, *15* (1), 2200.
(191) Liu, T.; Limpikirati, P.; Vachet, R. W. Synergistic Structural Information from Covalent Labeling and Hydrogen-Deuterium Exchange Mass Spectrometry for Protein-Ligand Interactions. *Anal. Chem.* **2019**, *91* (23), 15248−15254.
(192) Marciano, D. P.; Dharmarajan, V.; Griffin, P. R. HDX-MS Guided Drug Discovery: Small Molecules and Biopharmaceuticals. *Curr. Opin Struct Biol.* **2014**, *28*, 105−111.
(193) Ow, S. Y.; Kapp, E. A.; Tomasetig, V.; Zalewski, A.; Simmonds, J.; Panousis, C.; Wilson, M. J.; Nash, A. D.; Pelzing, M. HDX-MS Study on Garadacimab Binding to Activated FXII Reveals Potential Binding Interfaces through Differential Solvent Exposure. *MAbs* **2023**, *15* (1), No. 2163459.
(194) Woods, V. A.; Abzalimov, R. R.; Keedy, D. A. Native Dynamics and Allosteric Responses in PTP1B Probed by High-Resolution HDX-MS. *Protein Sci.* **2024**, *33* (6), No. e5024.
(195) Avila, J. R.; Lee, J. S.; Toriia, K. U. Co-Immunoprecipitation of Membrane-Bound Receptors. *Arabidopsis Book* **2015**, *13*, No. e0180.
(196) Sommer, F.; Mühlhaus, T.; Hemme, D.; Veyel, D.; Schroda, M. Identification and Validation of Protein-Protein Interactions by Combining Co-Immunoprecipitation, Antigen Competition, and Stable Isotope Labeling. *Methods Mol. Biol.* **2014**, *1188*, 245−261.
(197) Medina-Puche, L.; Lozano-Durán, R. Immunoprecipitation Followed by Mass Spectrometry: An Approach for Identifying Host−Viral Protein−Protein Interactions. *Methods Mol. Biol.* **2024**, *2724*, 289−305.
(198) Cai, Z.; Wang, D.; Li, Z.; Gu, M.; You, Q.; Wang, L. The Value of Coimmunoprecipitation (Co-IP) Assays in Drug Discovery. *Expert Opin. Drug Discov.* **2025**, *20*, 859−872.
(199) Baytek, G.; Popp, O.; Mertins, P.; Tursun, B. Robust Co-Immunoprecipitation with Mass Spectrometry for *Caenorhabditis Elegans* Using Solid-Phase Enhanced Sample Preparation. *Biotechniques* **2022**, *72* (5), 175−184.
(200) Liu, X.; Zhang, X.; Zeng, T.; Chen, Y.; Ye, L.; Wang, S.; Li, Y. FOSL1 Drives the Malignant Progression of Pancreatic Cancer Cells by Regulating Cell Stemness, Metastasis and Multidrug Efflux System. *J. Transl Med.* **2025**, *23* (1), 268.
(201) Maccarrone, G.; Bonfiglio, J. J.; Silberstein, S.; Turck, C. W.; Martins-de-Souza, D. Characterization of a Protein Interactome by Co-Immunoprecipitation and Shotgun Mass Spectrometry. *Methods Mol. Biol.* **2017**, *1546*, 223−234.
(202) Free, R. B.; Hazelwood, L. A.; Sibley, D. R. Identifying Novel Protein-Protein Interactions Using Co-Immunoprecipitation and Mass Spectroscopy. *Curr. Protoc Neurosci* **2009**, *46* (1), No. ns0528s46.
(203) Lagundžin, D.; Krieger, K. L.; Law, H. C.-H.; Woods, N. T. An Optimized Co-Immunoprecipitation Protocol for the Analysis of Endogenous Protein-Protein Interactions in Cell Lines Using Mass Spectrometry. *STAR Protoc* **2022**, *3* (1), No. 101234.
(204) Yang, L.; Nie, H.; Du, Y.; Liu, X.; Cai, B.; Li, J. Isoliquiritigenin Exhibits Anti-Inflammatory Responses in Acute Lung Injury by Covalently Binding to the Myeloid Differentiation Protein-2 Domain. *Phytotherapy Research* **2025**, *39* (2), 922−937.
(205) Wang, Z.; Zhang, Y.; Shen, Y.; Zhu, C.; Qin, X.; Gao, Y. Liquidambaric Acid Inhibits Cholangiocarcinoma Progression by Disrupting the STAMBPL1/NRF2 Positive Feedback Loop. *Phytomedicine* **2025**, *136*, No. 156303.
(206) Gorini, G.; Ponomareva, O.; Shores, K. S.; Person, M. D.; Harris, R. A.; Mayfield, R. D. Dynamin-1 Co-associates with Native

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

Mouse Brain BK_{Ca} Channels: Proteomics Analysis of Synaptic Protein Complexes. *FEBS Lett.* **2010**, *584* (5), 845−851.
(207) An, Y.; Zhou, L.; Huang, Z.; Nice, E. C.; Zhang, H.; Huang, C. Molecular Insights into Cancer Drug Resistance from a Proteomics Perspective. *Expert Rev. Proteomics* **2019**, *16* (5), 413−429.
(208) Agostini, M.; Traldi, P.; Hamdan, M. Mass Spectrometry-Based Proteomics: Analyses Related to Drug-Resistance and Disease Biomarkers. *Medicina (Kaunas)* **2023**, *59* (10), 1722.
(209) Tierney, C.; Bazou, D.; Majumder, M. M.; Anttila, P.; Silvennoinen, R.; Heckman, C. A.; Dowling, P.; O'Gorman, P. Next Generation Proteomics with Drug Sensitivity Screening Identifies Sub-Clones Informing Therapeutic and Drug Development Strategies for Multiple Myeloma Patients. *Sci. Rep* **2021**, *11* (1), 12866.
(210) Meng, F.; Zou, L.; Zhang, T.; Jiang, L.; Ding, Y.; Yu, P.; Peng, J. Using LC–MS/MS-Based Targeted Proteomics to Monitor the Pattern of ABC Transporters Expression in the Development of Drug Resistance. *Cancer Manage. Res.* **2018**, *10*, 2859−2870.
(211) Tapia, I. J.; Perico, D.; Wolos, V. J.; Villaverde, M. S.; Abrigo, M.; Di Silvestre, D.; Mauri, P.; De Palma, A.; Fiszman, G. L. Proteomic Characterization of a 3D HER2+ Breast Cancer Model Reveals the Role of Mitochondrial Complex I in Acquired Resistance to Trastuzumab. *Int. J. Mol. Sci.* **2024**, *25* (13), 7397.
(212) Hao, B.-B.; Ma, K.; Xu, J.-Y.; Fan, R.-F.; Zhao, W.-S.; Jia, X.-L.; Zhai, L.-H.; Lee, S.; Xie, D.; Tan, M.-J. Proteomics Analysis of Histone Deacetylase Inhibitor-Resistant Solid Tumors Reveals Resistant Signatures and Potential Drug Combinations. *Acta Pharmacol Sin* **2024**, *45* (6), 1305−1315.
(213) Abril, A. G.; Carrera, M.; Böhme, K.; Barros-Velázquez, J.; Calo-Mata, P.; Sánchez-Pérez, A.; Villa, T. G. Proteomic Characterization of Antibiotic Resistance in Listeria and Production of Antimicrobial and Virulence Factors. *Int. J. Mol. Sci.* **2021**, *22* (15), 8141.
(214) Saboia-Vahia, L.; Cuervo, P.; Wiśniewski, J. R.; Dias-Lopes, G.; Pinho, N.; Padrón, G.; de Pilla Varotti, F.; Murta, S. M. F. In-Depth Quantitative Proteomics Characterization of In Vitro Selected Miltefosine Resistance in Leishmania Infantum. *Proteomes* **2022**, *10* (2), 10.
(215) Tagliazucchi, L.; Pinetti, D.; Genovese, F.; Malpezzi, G.; Perea Martinez, A.; Manzano, J. I.; García-Hernández, R.; Cole, A. R.; Kwon, B. R.; Aiello, D.; Brooks, B. W.; Thoré, E. S. J.; Bertram, M. G.; Gamarro, F.; Costi, M. P. Deciphering Host-Parasite Interplay in Leishmania Infection through a One Health View of Proteomics Studies on Drug Resistance. *ACS Infect Dis* **2024**, *10* (9), 3202−3221.
(216) Ankley, G. T.; Edwards, S. W. The Adverse Outcome Pathway: A Multifaceted Framework Supporting 21st Century Toxicology. *Curr. Opin Toxicol* **2018**, *9*, 1−7.
(217) Gust, K. A.; Ji, Q.; Luo, X. Example of Adverse Outcome Pathway Concept Enabling Genome-to-Phenome Discovery in Toxicology. *Integr Comp Biol.* **2020**, *60* (2), 375−384.
(218) Korn, D.; Thieme, A. J.; Alves, V. M.; Yeakey, M.; Borba, J. V. V. B.; Capuzzi, S. J.; Fecho, K.; Bizon, C.; Edwards, S. W.; Chirkova, R.; Colvis, C. M.; Southall, N. T.; Austin, C. P.; Muratov, E. N.; Tropsha, A. Defining Clinical Outcome Pathways. *Drug Discov Today* **2022**, *27* (6), 1671−1678.
(219) von Coburg, E.; Dunst, S. The Adverse Outcome Pathway for Breast Cancer: A Knowledge Management Framework Bridging Biomedicine and Toxicology. *Discover Oncology* **2023**, *14* (1), 223.
(220) Saarimäki, L. A.; Fratello, M.; Pavel, A.; Korpilähde, S.; Leppänen, J.; Serra, A.; Greco, D. A Curated Gene and Biological System Annotation of Adverse Outcome Pathways Related to Human Health. *Sci. Data* **2023**, *10* (1), 409.
(221) Kleinstreuer, N. C.; Sullivan, K.; Allen, D.; Edwards, S.; Mendrick, D. L.; Embry, M.; Matheson, J.; Rowlands, J. C.; Munn, S.; Maull, E.; Casey, W. Adverse Outcome Pathways: From Research to Regulation Scientific Workshop Report. *Regul. Toxicol. Pharmacol.* **2016**, *76*, 39−50.
(222) Conolly, R. B.; Ankley, G. T.; Cheng, W.; Mayo, M. L.; Miller, D. H.; Perkins, E. J.; Villeneuve, D. L.; Watanabe, K. H. Quantitative Adverse Outcome Pathways and Their Application to Predictive Toxicology. *Environ. Sci. Technol.* **2017**, *51* (8), 4661−4672.
(223) Pognan, F.; Beilmann, M.; Boonen, H. C. M.; Czich, A.; Dear, G.; Hewitt, P.; Mow, T.; Oinonen, T.; Roth, A.; Steger-Hartmann, T.; Valentin, J.-P.; Van Goethem, F.; Weaver, R. J.; Newham, P. The Evolving Role of Investigative Toxicology in the Pharmaceutical Industry. *Nat. Rev. Drug Discov* **2023**, *22* (4), 317−335.
(224) Vinken, M. Adverse Outcome Pathways and Drug-Induced Liver Injury Testing. *Chem. Res. Toxicol.* **2015**, *28* (7), 1391−1397.
(225) Svingen, T.; Villeneuve, D. L.; Knapen, D.; Panagiotou, E. M.; Draskau, M. K.; Damdimopoulou, P.; O'Brien, J. M. A Pragmatic Approach to Adverse Outcome Pathway Development and Evaluation. *Toxicol. Sci.* **2021**, *184* (2), 183−190.
(226) Veltman, C. H. J.; Pennings, J. L. A.; van de Water, B.; Luijten, M. An Adverse Outcome Pathway Network for Chemically Induced Oxidative Stress Leading to (Non) Genotoxic Carcinogenesis. *Chem. Res. Toxicol.* **2023**, *36* (6), 805−817.
(227) Bakker, R.; Ellers, J.; Roelofs, D.; Vooijs, R.; Dijkstra, T.; van Gestel, C. A. M.; Hoedjes, K. M. Combining Time-Resolved Transcriptomics and Proteomics Data for Adverse Outcome Pathway Refinement in Ecotoxicology. *Science of The Total Environment* **2023**, *869*, No. 161740.
(228) Liu, Z.; Li, Y.; Sepúlveda, M. S.; Jiang, Q.; Jiao, Y.; Chen, Q.; Huang, Y.; Tian, J.; Zhao, Y. Development of an Adverse Outcome Pathway for Nanoplastic Toxicity in Daphnia Pulex Using Proteomics. *Science of The Total Environment* **2021**, *766*, No. 144249.
(229) Neagu, A.-N.; Jayathirtha, M.; Baxter, E.; Donnelly, M.; Petre, B. A.; Darie, C. C. Applications of Tandem Mass Spectrometry (MS/MS) in Protein Analysis for Biomedical Research. *Molecules* **2022**, *27* (8), 2411.
(230) Åkesson, J.; Hojjati, S.; Hellberg, S.; Raffetseder, J.; Khademi, M.; Rynkowski, R.; Kockum, I.; Altafini, C.; Lubovac-Pilav, Z.; Mellergård, J.; Jenmalm, M. C.; Piehl, F.; Olsson, T.; Ernerudh, J.; Gustafsson, M. Proteomics Reveal Biomarkers for Diagnosis, Disease Activity and Long-Term Disability Outcomes in Multiple Sclerosis. *Nat. Commun.* **2023**, *14* (1), 6903.
(231) Hällqvist, J.; Bartl, M.; Dakna, M.; Schade, S.; Garagnani, P.; Bacalini, M.-G.; Pirazzini, C.; Bhatia, K.; Schreglmann, S.; Xylaki, M.; Weber, S.; Ernst, M.; Muntean, M.-L.; Sixel-Döring, F.; Franceschi, C.; Doykov, I.; Śpiewak, J.; Vinette, H.; Trenkwalder, C.; Heywood, W. E.; Mills, K.; Mollenhauer, B. Plasma Proteomics Identify Biomarkers Predicting Parkinson's Disease up to 7 Years before Symptom Onset. *Nat. Commun.* **2024**, *15* (1), 4759.
(232) Zhang, D.; Li, Y.; Liang, M.; Liang, Y.; Tian, J.; He, Q.; Yang, B.; Jin, J.; Zhu, W. LC-MS/MS Based Metabolomics and Proteomics Reveal Candidate Biomarkers and Molecular Mechanism of Early IgA Nephropathy. *Clin Proteomics* **2022**, *19* (1), 51.
(233) Harlan, R.; Zhang, H. Targeted Proteomics: A Bridge between Discovery and Validation. *Expert Rev. Proteomics* **2014**, *11* (6), 657−661.
(234) Kwon, Y. W.; Jo, H.-S.; Bae, S.; Seo, Y.; Song, P.; Song, M.; Yoon, J. H. Application of Proteomics in Cancer: Recent Trends and Approaches for Biomarkers Discovery. *Front Med. (Lausanne)* **2021**, *8*, No. 747333.
(235) Parker, C. E.; Borchers, C. H. Mass Spectrometry Based Biomarker Discovery, Verification, and Validation–Quality Assurance and Control of Protein Biomarker Assays. *Mol. Oncol* **2014**, *8* (4), 840−858.
(236) Babaei, M.; Kashanian, S.; Lee, H.; Harding, F. Proteomics Techniques in Protein Biomarker Discovery. *Quantitative Biology* **2024**, *12* (1), 53−69.
(237) Paver, E. C.; Morey, A. L. Biomarkers and Biomarker Validation: A Pathologist's Guide to Getting It Right. *Pathology* **2024**, *56* (2), 147−157.
(238) Ansong, C.; Purvine, S. O.; Adkins, J. N.; Lipton, M. S.; Smith, R. D. Proteogenomics: Needs and Roles to Be Filled by Proteomics in Genome Annotation. *Brief Funct Genomic Proteomic* **2008**, *7* (1), 50−62.
(239) Rudnick, P. A.; Markey, S. P.; Roth, J.; Mirokhin, Y.; Yan, X.; Tchekhovskoi, D. V.; Edwards, N. J.; Thangudu, R. R.; Ketchum, K. A.; Kinsinger, C. R.; Mesri, M.; Rodriguez, H.; Stein, S. E. A Description of

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

the Clinical Proteomic Tumor Analysis Consortium (CPTAC) Common Data Analysis Pipeline. *J. Proteome Res.* **2016**, *15* (3), 1023−1032.
(240) Chiou, S.-H.; Lee, K.-T. Proteomic Analysis and Translational Perspective of Hepatocellular Carcinoma: Identification of Diagnostic Protein Biomarkers by an Onco-Proteogenomics Approach. *Kaohsiung J. Med. Sci.* **2016**, *32* (11), 535−544.
(241) Huang, C.-H.; Kuo, C.-J.; Liang, S.-S.; Chi, S.-W.; Hsi, E.; Chen, C.-C.; Lee, K.-T.; Chiou, S.-H. Onco-Proteogenomics Identifies Urinary S100A9 and GRN as Potential Combinatorial Biomarkers for Early Diagnosis of Hepatocellular Carcinoma. *BBA Clin* **2015**, *3*, 205−213.
(242) Gillette, M. A.; Satpathy, S.; Cao, S.; Dhanasekaran, S. M.; Vasaikar, S. V.; Krug, K.; Petralia, F.; Li, Y.; Liang, W.-W.; Reva, B.; Krek, A.; Ji, J.; Song, X.; Liu, W.; Hong, R.; Yao, L.; Blumenberg, L.; Savage, S. R.; Wendl, M. C.; Wen, B.; Li, K.; Tang, L. C.; MacMullan, M. A.; Avanessian, S. C.; Kane, M. H.; Newton, C. J.; Cornwell, M.; Kothadia, R. B.; Ma, W.; Yoo, S.; Mannan, R.; Vats, P.; Kumar-Sinha, C.; Kawaler, E. A.; Omelchenko, T.; Colaprico, A.; Geffen, Y.; Maruvka, Y. E.; da Veiga Leprevost, F.; Wiznerowicz, M.; Gümüş, Z. H.; Veluswamy, R. R.; Hostetter, G.; Heiman, D. I.; Wyczalkowski, M. A.; Hiltke, T.; Mesri, M.; Kinsinger, C. R.; Boja, E. S.; Omenn, G. S.; Chinnaiyan, A. M.; Rodriguez, H.; Li, Q. K.; Jewell, S. D.; Thiagarajan, M.; Getz, G.; Zhang, B.; Fenyö, D.; Ruggles, K. V.; Cieslik, M. P.; Robles, A. I.; Clauser, K. R.; Govindan, R.; Wang, P.; Nesvizhskii, A. I.; Ding, L.; Mani, D. R.; Carr, S. A.; Clinical Proteomic Tumor Analysis Consortium. Proteogenomic Characterization Reveals Therapeutic Vulnerabilities in Lung Adenocarcinoma. *Cell* **2020**, *182* (1), 200−225.
(243) Ling, J. P.; Pletnikova, O.; Troncoso, J. C.; Wong, P. C. TDP-43 Repression of Nonconserved Cryptic Exons Is Compromised in ALS-FTD. *Science* **2015**, *349* (6248), 650−655.
(244) Viodé, A.; Fournier, C.; Camuzat, A.; Fenaille, F.; Latouche, M.; Elahi, F.; Ber, I. L.; Junot, C.; Lamari, F.; Anquetil, V.; Becher, F. New Antibody-Free Mass Spectrometry-Based Quantification Reveals That C9ORF72 Long Protein Isoform Is Reduced in the Frontal Cortex of Hexanucleotide-Repeat Expansion Carriers. *Front. Neurosci.* **2018**, *12*, 589.
(245) Álvez, M. B.; Edfors, F.; von Feilitzen, K.; Zwahlen, M.; Mardinoglu, A.; Edqvist, P.-H.; Sjöblom, T.; Lundin, E.; Rameika, N.; Enblad, G.; Lindman, H.; Höglund, M.; Hesselager, G.; Stålberg, K.; Enblad, M.; Simonson, O. E.; Häggman, M.; Axelsson, T.; Åberg, M.; Nordlund, J.; Zhong, W.; Karlsson, M.; Gyllensten, U.; Ponten, F.; Fagerberg, L.; Uhlén, M. Next Generation Pan-Cancer Blood Proteome Profiling Using Proximity Extension Assay. *Nat. Commun.* **2023**, *14* (1), 4308.
(246) Zertuche-Martínez, C.; Velázquez-Enríquez, J. M.; González-García, K.; Santos-Álvarez, J. C.; Romero-Tlalolini, M. de L. Á; Pina-Canseco, S.; Pérez-Campos Mayoral, L.; Muriel, P.; Villa-Treviño, S.; Baltiérrez-Hoyos, R.; Arellanes-Robledo, J.; Vásquez-Garzón, V. R. Discovery of Candidate Biomarkers from Plasma-Derived Extracellular Vesicles of Patients with Cirrhosis and Hepatocellular Carcinoma: An Exploratory Proteomic Study. *Mol. Omics* **2024**, *20* (7), 483−495.
(247) Lan, T.; Zeng, Q.; Fan, Y.; Liu, T.; Yao, P.; Liang, Z.; Dang, X.; Zhu, H.; Li, Y.; Jiang, W.; Lu, W. Proteomics Analysis of Serum Reveals Potential Biomarkers for Heart Failure Patients with Phlegm-Blood Stasis Syndrome. *J. Proteome Res.* **2024**, *23* (1), 226−237.
(248) Sanni, A.; Goli, M.; Zhao, J.; Wang, J.; Barsa, C.; El Hayek, S.; Talih, F.; Lanuzza, B.; Kobeissy, F.; Plazzi, G.; Moresco, M.; Mondello, S.; Ferri, R.; Mechref, Y. LC-MS/MS-Based Proteomics Approach for the Identification of Candidate Serum Biomarkers in Patients with Narcolepsy Type 1. *Biomolecules* **2023**, *13* (3), 420.
(249) Hällqvist, J.; Pinto, R. C.; Heywood, W. E.; Cordey, J.; Foulkes, A. J. M.; Slattery, C. F.; Leckey, C. A.; Murphy, E. C.; Zetterberg, H.; Schott, J. M.; Mills, K.; Paterson, R. W. A Multiplexed Urinary Biomarker Panel Has Potential for Alzheimer's Disease Diagnosis Using Targeted Proteomics and Machine Learning. *Int. J. Mol. Sci.* **2023**, *24* (18), No. 13758.
(250) Song, P.; Kwon, Y.; Joo, J.-Y.; Kim, D.-G.; Yoon, J. H. Secretomics to Discover Regulators in Diseases. *Int. J. Mol. Sci.* **2019**, *20* (16), 3893.
(251) Knecht, S.; Eberl, H. C.; Kreisz, N.; Ugwu, U. J.; Starikova, T.; Kuster, B.; Wilhelm, S. An Introduction to Analytical Challenges, Approaches, and Applications in Mass Spectrometry−Based Secretomics. *Molecular & Cellular Proteomics* **2023**, *22* (9), No. 100636.
(252) Wang, Y.-T.; Shi, T.; Srivastava, S.; Kagan, J.; Liu, T.; Rodland, K. D. Proteomic Analysis of Exosomes for Discovery of Protein Biomarkers for Prostate and Bladder Cancer. *Cancers (Basel)* **2020**, *12* (9), 2335.
(253) Dilsiz, N. A Comprehensive Review on Recent Advances in Exosome Isolation and Characterization: Toward Clinical Applications. *Transl Oncol* **2024**, *50*, No. 102121.
(254) Kugeratski, F. G.; Hodge, K.; Lilla, S.; McAndrews, K. M.; Zhou, X.; Hwang, R. F.; Zanivan, S.; Kalluri, R. Quantitative Proteomics Identifies the Core Proteome of Exosomes with Syntenin-1 as the Highest Abundant Protein and a Putative Universal Biomarker. *Nat. Cell Biol.* **2021**, *23* (6), 631−641.
(255) Ung, T. H.; Madsen, H. J.; Hellwinkel, J. E.; Lencioni, A. M.; Graner, M. W. Exosome Proteomics Reveals Transcriptional Regulator Proteins with Potential to Mediate Downstream Pathways. *Cancer Sci.* **2014**, *105* (11), 1384−1392.
(256) Li, S.; Qu, Y.; Liu, L.; Zhang, X.; He, Y.; Wang, C.; Guo, Y.; Yuan, L.; Ma, Z.; Bai, H.; Wang, J. Comparative Proteomic Profiling of Plasma Exosomes in Lung Cancer Cases of Liver and Brain Metastasis. *Cell Biosci* **2023**, *13* (1), 180.
(257) Gao, Z.; Jiang, C.; Zhang, J.; Jiang, X.; Li, L.; Zhao, P.; Yang, H.; Huang, Y.; Li, J. Hierarchical Graph Learning for Protein−Protein Interaction. *Nat. Commun.* **2023**, *14* (1), 1093.
(258) Silverbush, D.; Sharan, R. A Systematic Approach to Orient the Human Protein−Protein Interaction Network. *Nat. Commun.* **2019**, *10* (1), 3015.
(259) Wang, C.; Xu, S.; Sun, D.; Liu, Z.-P. ActivePPI: Quantifying Protein−Protein Interaction Network Activity with Markov Random Fields. *Bioinformatics* **2023**, *39* (9), No. btad567.
(260) Farooq, Q.ul A.; Shaukat, Z.; Aiman, S.; Li, C.-H. Protein-Protein Interactions: Methods, Databases, and Applications in Virus-Host Study. *World J. Virol* **2021**, *10* (6), 288−300.
(261) Szklarczyk, D.; Gable, A. L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N. T.; Morris, J. H.; Bork, P.; Jensen, L. J.; Mering, C. v. STRING V11: Protein−Protein Association Networks with Increased Coverage, Supporting Functional Discovery in Genome-Wide Experimental Datasets. *Nucleic Acids Res.* **2019**, *47* (D1), D607−D613.
(262) Szklarczyk, D.; Gable, A. L.; Nastou, K. C.; Lyon, D.; Kirsch, R.; Pyysalo, S.; Doncheva, N. T.; Legeay, M.; Fang, T.; Bork, P.; Jensen, L. J.; von Mering, C. Correction to 'The STRING Database in 2021: Customizable Protein−Protein Networks, and Functional Characterization of User-Uploaded Gene/Measurement Sets. *Nucleic Acids Res.* **2021**, *49* (18), 10800−10800.
(263) Oughtred, R.; Rust, J.; Chang, C.; Breitkreutz, B.; Stark, C.; Willems, A.; Boucher, L.; Leung, G.; Kolas, N.; Zhang, F.; Dolma, S.; Coulombe-Huntington, J.; Chatr-aryamontri, A.; Dolinski, K.; Tyers, M. The BioGRID Database: A Comprehensive Biomedical Resource of Curated Protein, Genetic, and Chemical Interactions. *Protein Sci.* **2021**, *30* (1), 187−200.
(264) del Toro, N.; Shrivastava, A.; Ragueneau, E.; Meldal, B.; Combe, C.; Barrera, E.; Perfetto, L.; How, K.; Ratan, P.; Shirodkar, G.; Lu, O.; Mészáros, B.; Watkins, X.; Pundir, S.; Licata, L.; Iannuccelli, M.; Pellegrini, M.; Martin, M. J.; Panni, S.; Duesbury, M.; Vallet, S. D.; Rappsilber, J.; Ricard-Blum, S.; Cesareni, G.; Salwinski, L.; Orchard, S.; Porras, P.; Panneerselvam, K.; Hermjakob, H. The IntAct Database: Efficient Access to Fine-Grained Molecular Interaction Data. *Nucleic Acids Res.* **2022**, *50* (D1), D648−D653.
(265) Porras, P.; Barrera, E.; Bridge, A.; del-Toro, N.; Cesareni, G.; Duesbury, M.; Hermjakob, H.; Iannuccelli, M.; Jurisica, I.; Kotlyar, M.; Licata, L.; Lovering, R. C.; Lynn, D. J.; Meldal, B.; Nanduri, B.; Paneerselvam, K.; Panni, S.; Pastrello, C.; Pellegrini, M.; Perfetto, L.;

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

Rahimzadeh, N.; Ratan, P.; Ricard-Blum, S.; Salwinski, L.; Shirodkar, G.; Shrivastava, A.; Orchard, S. Towards a Unified Open Access Dataset of Molecular Interactions. *Nat. Commun.* **2020**, *11* (1), 6144.
(266) Perez, J. J.; Perez, R. A.; Perez, A. Computational Modeling as a Tool to Investigate PPI: From Drug Design to Tissue Engineering. *Front Mol. Biosci* **2021**, *8*, No. 681617.
(267) Ding, H.; Li, X.; Han, P.; Tian, X.; Jing, F.; Wang, S.; Song, T.; Fu, H.; Kang, N. MEG-PPIS: A Fast Protein−Protein Interaction Site Prediction Method Based on Multi-Scale Graph Information and Equivariant Graph Neural Network. *Bioinformatics* **2024**, *40* (5), No. btae269.
(268) Gallagher, S. R.; Goldberg, D. S. *Clustering Coefficients in Protein Interaction Hypernetworks*. In Proceedings of the International Conference on Bioinformatics, Computational Biology and Biomedical Informatics; ACM: New York, NY, USA, 2013; pp 552−560. .
(269) A Neural Network for Large-Scale Clustering of Peptide Mass Spectra. *Nat. Methods* **2022**, *19* (6), 658−659. .
(270) Wang, X.-W.; Madeddu, L.; Spirohn, K.; Martini, L.; Fazzone, A.; Becchetti, L.; Wytock, T. P.; Kovács, I. A.; Balogh, O. M.; Benczik, B.; Pétervári, M.; Agg, B.; Ferdinandy, P.; Vulliard, L.; Menche, J.; Colonnese, S.; Petti, M.; Scarano, G.; Cuomo, F.; Hao, T.; Laval, F.; Willems, L.; Twizere, J.-C.; Vidal, M.; Calderwood, M. A.; Petrillo, E.; Barabási, A.-L.; Silverman, E. K.; Loscalzo, J.; Velardi, P.; Liu, Y.-Y. Assessment of Community Efforts to Advance Network-Based Prediction of Protein−Protein Interactions. *Nat. Commun.* **2023**, *14* (1), 1582.
(271) Gilbert, M.; Li, Z.; Wu, X. N.; Rohr, L.; Gombos, S.; Harter, K.; Schulze, W. X. Comparison of Path-Based Centrality Measures in Protein-Protein Interaction Networks Revealed Proteins with Phenotypic Relevance during Adaptation to Changing Nitrogen Environments. *J. Proteomics* **2021**, *235*, No. 104114.
(272) Kim, M.; Hwang, D. Network-Based Protein Biomarker Discovery Platforms. *Genomics Inform* **2016**, *14* (1), 2.
(273) Chisanga, D.; Keerthikumar, S.; Mathivanan, S.; Chilamkurti, N. Network Tools for the Analysis of Proteomic Data. *Methods Mol. Biol.* **2017**, *1549*, 177−197.
(274) Koh, G. C. K. W.; Porras, P.; Aranda, B.; Hermjakob, H.; Orchard, S. E. Analyzing Protein−Protein Interaction Networks. *J. Proteome Res.* **2012**, *11* (4), 2014−2031.
(275) Zhang, J.; Humphreys, I. R.; Pei, J.; Kim, J.; Choi, C.; Yuan, R.; Durham, J.; Liu, S.; Choi, H.-J.; Baek, M.; Baker, D.; Cong, Q. Predicting Protein-Protein Interactions in the Human Proteome. *Science* **2025**.
(276) Carvalho, L. B.; Teigas-Campos, P. A. D.; Jorge, S.; Protti, M.; Mercolini, L.; Dhir, R.; Wiśniewski, J. R.; Lodeiro, C.; Santos, H. M.; Capelo, J. L. Normalization Methods in Mass Spectrometry-Based Analytical Proteomics: A Case Study Based on Renal Cell Carcinoma Datasets. *Talanta* **2024**, *266*, No. 124953.
(277) Gardner, M. L.; Freitas, M. A. Multiple Imputation Approaches Applied to the Missing Value Problem in Bottom-Up Proteomics. *Int. J. Mol. Sci.* **2021**, *22* (17), 9650.
(278) Jin, L.; Bi, Y.; Hu, C.; Qu, J.; Shen, S.; Wang, X.; Tian, Y. A Comparative Study of Evaluating Missing Value Imputation Methods in Label-Free Proteomics. *Sci. Rep* **2021**, *11* (1), 1760.
(279) Jiang, Y.; Rex, D. A. B.; Schuster, D.; Neely, B. A.; Rosano, G. L.; Volkmar, N.; Momenzadeh, A.; Peters-Clarke, T. M.; Egbert, S. B.; Kreimer, S.; Doud, E. H.; Crook, O. M.; Yadav, A. K.; Vanuopadath, M.; Hegeman, A. D.; Mayta, M. L.; Duboff, A. G.; Riley, N. M.; Moritz, R. L.; Meyer, J. G. Comprehensive Overview of Bottom-Up Proteomics Using Mass Spectrometry. *ACS Measurement Science Au* **2024**, *4* (4), 338−417.
(280) Chen, C.; Hou, J.; Tanner, J. J.; Cheng, J. Bioinformatics Methods for Mass Spectrometry-Based Proteomics Data Analysis. *Int. J. Mol. Sci.* **2020**, *21* (8), 2873.
(281) Feng, Z.; Fang, P.; Zheng, H.; Zhang, X. DEP2: An Upgraded Comprehensive Analysis Toolkit for Quantitative Proteomics Data. *Bioinformatics* **2023**, *39* (8), No. btad526.
(282) Peng, H.; Wang, H.; Kong, W.; Li, J.; Goh, W. W. B. Optimizing Differential Expression Analysis for Proteomics Data via High-Performing Rules and Ensemble Inference. *Nat. Commun.* **2024**, *15* (1), 3922.
(283) Aguilan, J. T.; Kulej, K.; Sidoli, S. Guide for Protein Fold Change and *p* -Value Calculation for Non-Experts in Proteomics. *Mol. Omics* **2020**, *16* (6), 573−582.
(284) Locard-Paulet, M.; Doncheva, N. T.; Morris, J. H.; Jensen, L. J. Functional Analysis of MS-Based Proteomics Data: From Protein Groups to Networks. *Molecular & Cellular Proteomics* **2024**, *23* (12), No. 100871.
(285) Hu, J.; He, J.; Li, J.; Gao, Y.; Zheng, Y.; Shang, X. A Novel Algorithm for Alignment of Multiple PPI Networks Based on Simulated Annealing. *BMC Genomics* **2019**, *20* (Suppl. 13), 932.
(286) Menor-Flores, M.; Vega-Rodríguez, M. A. Decomposition-Based Multi-Objective Optimization Approach for PPI Network Alignment. *Knowl Based Syst* **2022**, *243*, No. 108527.
(287) Zhu, L.; Zhang, J.; Zhang, Y.; Lang, J.; Xiang, J.; Bai, X.; Yan, N.; Tian, G.; Zhang, H.; Yang, J. An Improved Method to Align PPI Networks Based on Gene Ontology and Graphlets. *Front. Bioeng. Biotechnol.* **2020**, *8*, 547.
(288) Ashburner, M.; Ball, C. A.; Blake, J. A.; Botstein, D.; Butler, H.; Cherry, J. M.; Davis, A. P.; Dolinski, K.; Dwight, S. S.; Eppig, J. T.; Harris, M. A.; Hill, D. P.; Issel-Tarver, L.; Kasarskis, A.; Lewis, S.; Matese, J. C.; Richardson, J. E.; Ringwald, M.; Rubin, G. M.; Sherlock, G. Gene Ontology: Tool for the Unification of Biology. *Gene Ontology Consortium. Nat. Genet* **2000**, *25* (1), 25−29.
(289) Kanehisa, M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **2000**, *28* (1), 27−30.
(290) Dennis, G.; Sherman, B. T.; Hosack, D. A.; Yang, J.; Gao, W.; Lane, H. C.; Lempicki, R. A. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* **2003**, *4* (5), P3.
(291) Eden, E.; Navon, R.; Steinfeld, I.; Lipson, D.; Yakhini, Z. GOrilla: A Tool for Discovery and Visualization of Enriched GO Terms in Ranked Gene Lists. *BMC Bioinformatics* **2009**, *10* (1), 48.
(292) Thomas, P. D.; Ebert, D.; Muruganujan, A.; Mushayahama, T.; Albou, L.; Mi, H. PANTHER: Making Genome-scale Phylogenetics Accessible to All. *Protein Sci.* **2022**, *31* (1), 8−22.
(293) Haw, R.; Stein, L. Using the Reactome Database. *Curr. Protoc. Bioinformatics* **2012**, 8.7.1−8.7.23.
(294) Sidiropoulos, K.; Viteri, G.; Sevilla, C.; Jupe, S.; Webber, M.; Orlic-Milacic, M.; Jassal, B.; May, B.; Shamovsky, V.; Duenas, C.; Rothfels, K.; Matthews, L.; Song, H.; Stein, L.; Haw, R.; D'Eustachio, P.; Ping, P.; Hermjakob, H.; Fabregat, A. Reactome Enhanced Pathway Visualization. *Bioinformatics* **2017**, *33* (21), 3461−3467.
(295) Gillespie, M.; Jassal, B.; Stephan, R.; Milacic, M.; Rothfels, K.; Senff-Ribeiro, A.; Griss, J.; Sevilla, C.; Matthews, L.; Gong, C.; Deng, C.; Varusai, T.; Ragueneau, E.; Haider, Y.; May, B.; Shamovsky, V.; Weiser, J.; Brunson, T.; Sanati, N.; Beckman, L.; Shao, X.; Fabregat, A.; Sidiropoulos, K.; Murillo, J.; Viteri, G.; Cook, J.; Shorser, S.; Bader, G.; Demir, E.; Sander, C.; Haw, R.; Wu, G.; Stein, L.; Hermjakob, H.; D'Eustachio, P. The Reactome Pathway Knowledgebase 2022. *Nucleic Acids Res.* **2022**, *50* (D1), D687−D692.
(296) Krämer, A.; Green, J.; Pollard, J.; Tugendreich, S. Causal Analysis Approaches in Ingenuity Pathway Analysis. *Bioinformatics* **2014**, *30* (4), 523−530.
(297) Ahmed, Z.; Wan, S.; Zhang, F.; Zhong, W. Artificial Intelligence for Omics Data Analysis. *BMC Methods* **2024**, *1* (1), 4.
(298) Gabriel, W.; The, M.; Zolg, D. P.; Bayer, F. P.; Shouman, O.; Lautenbacher, L.; Schnatbaum, K.; Zerweck, J.; Knaute, T.; Delanghe, B.; Huhmer, A.; Wenschuh, H.; Reimer, U.; Médard, G.; Kuster, B.; Wilhelm, M. Prosit-TMT: Deep Learning Boosts Identification of TMT-Labeled Peptides. *Anal. Chem.* **2022**, *94* (20), 7181−7190.
(299) Gessulat, S.; Schmidt, T.; Zolg, D. P.; Samaras, P.; Schnatbaum, K.; Zerweck, J.; Knaute, T.; Rechenberger, J.; Delanghe, B.; Huhmer, A.; Reimer, U.; Ehrlich, H.-C.; Aiche, S.; Kuster, B.; Wilhelm, M. Prosit: Proteome-Wide Prediction of Peptide Tandem Mass Spectra by Deep Learning. *Nat. Methods* **2019**, *16* (6), 509−518.
(300) Frejno, M.;et al.CHIMERYS: An AI-Driven Leap Forward in Peptide Identification. *ThermoFischer Scientific*,https://assets.

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX

thermofisher.cn/TFS-Assets/CMD/posters/po-66098-chimerys-peptide-identification-po66098-en.pdf **2023**.
(301) Matzinger, M.; Schmücker, A.; Yelagandula, R.; Stejskal, K.; Kršššáková, G.; Berger, F.; Mechtler, K.; Mayer, R. L. Micropillar Arrays, Wide Window Acquisition and AI-Based Data Analysis Improve Comprehensiveness in Multiple Proteomic Applications. *Nat. Commun.* **2024**, *15* (1), 1019.
(302) Agarwal, N.; Nupur; Paul, P. K.; Mishra, S. K. *Artificial Intelligence and Machine Learning for Analysis of Multi-Omics*. In Multi-Omics Analysis of the Human Microbiome; Springer Nature Singapore: Singapore, 2024; pp 339−354. .
(303) He, X.; Liu, X.; Zuo, F.; Shi, H.; Jing, J. Artificial Intelligence-Based Multi-Omics Analysis Fuels Cancer Precision Medicine. *Semin Cancer Biol.* **2023**, *88*, 187−200.
(304) Hartman, E.; Scott, A. M.; Karlsson, C.; Mohanty, T.; Vaara, S. T.; Linder, A.; Malmström, L.; Malmström, J. Interpreting Biologically Informed Neural Networks for Enhanced Proteomic Biomarker Discovery and Pathway Analysis. *Nat. Commun.* **2023**, *14* (1), 5359.
(305) Valous, N. A.; Popp, F.; Zörnig, I.; Jäger, D.; Charoentong, P. Graph Machine Learning for Integrated Multi-Omics Analysis. *Br. J. Cancer* **2024**, *131* (2), 205−211.
(306) Gao, F.; Huang, K.; Xing, Y. Artificial Intelligence in Omics. *Genomics Proteomics Bioinformatics* **2022**, *20* (5), 811−813.
(307) Patel, S. K.; George, B.; Rai, V. Artificial Intelligence to Decode Cancer Mechanism: Beyond Patient Stratification for Precision Oncology. *Front Pharmacol* **2020**, *11*, 1177.
(308) Stransky, S.; Sun, Y.; Shi, X.; Sidoli, S. Ten Questions to AI Regarding the Present and Future of Proteomics. *Front Mol. Biosci* **2023**, *10*, No. 1295721.
(309) Baysoy, A.; Bai, Z.; Satija, R.; Fan, R. The Technological Landscape and Applications of Single-Cell Multi-Omics. *Nat. Rev. Mol. Cell Biol.* **2023**, *24* (10), 695−713.
(310) Raufaste-Cazavieille, V.; Santiago, R.; Droit, A. Multi-Omics Analysis: Paving the Path toward Achieving Precision Medicine in Cancer Treatment and Immuno-Oncology. *Front Mol. Biosci* **2022**, *9*, No. 962743.
(311) Tarazona, S.; Arzalluz-Luque, A.; Conesa, A. Undisclosed, Unmet and Neglected Challenges in Multi-Omics Studies. *Nat. Comput. Sci.* **2021**, *1* (6), 395−402.
(312) Mani, D. R.; Krug, K.; Zhang, B.; Satpathy, S.; Clauser, K. R.; Ding, L.; Ellis, M.; Gillette, M. A.; Carr, S. A. Cancer Proteogenomics: Current Impact and Future Prospects. *Nat. Rev. Cancer* **2022**, *22* (5), 298−313.
(313) Zhang, D.; Zhao, M.; Wu, J.; Wang, R.; Xue, G.; Xue, Y.; Shao, J.; Zhang, Y.; Dong, E.; Li, Z.; Xiao, H. Dual-Omics Reveals Temporal Differences in Acute Sympathetic Stress-Induced Cardiac Inflammation Following A1 and β-Adrenergic Receptors Activation. *Acta Pharmacol Sin* **2023**, *44* (7), 1350−1365.
(314) Teibo, J. O.; Silvestrini, V. C.; Vargas, A. P.; Lanfredi, G. P.; Faça, V. M. *The Interplay Between the Transcriptomics and Proteomics Profiles*. In Transcriptomics in Health and Disease; Springer International Publishing: Cham, 2022; pp 187−208. .
(315) Nalbantoglu, S.; Karadag, A. Metabolomics Bridging Proteomics along Metabolites/Oncometabolites and Protein Modifications: Paving the Way toward Integrative Multiomics. *J. Pharm. Biomed Anal* **2021**, *199*, No. 114031.
(316) Chen, Y.; Ni, J.; Gao, Y.; Zhang, J.; Liu, X.; Chen, Y.; Chen, Z.; Wu, Y. Integrated Proteomics and Metabolomics Reveals the Comprehensive Characterization of Antitumor Mechanism Underlying Shikonin on Colon Cancer Patient-Derived Xenograft Model. *Sci. Rep* **2020**, *10* (1), 14092.
(317) Wörheide, M. A.; Krumsiek, J.; Kastenmüller, G.; Arnold, M. Multi-Omics Integration in Biomedical Research − A Metabolomics-Centric Review. *Anal. Chim. Acta* **2021**, *1141*, 144−162.
(318) Zaghlool, S. B.; Kühnel, B.; Elhadad, M. A.; Kader, S.; Halama, A.; Thareja, G.; Engelke, R.; Sarwath, H.; Al-Dous, E. K.; Mohamoud, Y. A.; Meitinger, T.; Wilson, R.; Strauch, K.; Peters, A.; Mook-Kanamori, D. O.; Graumann, J.; Malek, J. A.; Gieger, C.; Waldenberger, M.; Suhre, K. Epigenetics Meets Proteomics in an Epigenome-Wide Association Study with Circulating Blood Plasma Protein Traits. *Nat. Commun.* **2020**, *11* (1), 15.
(319) Bi, H.; Weng, X. Single-Cell Epigenomics and Proteomics Methods Integrated in Multiomics. *Fundamental Research* **2025**, *5*, 1988.
(320) Wang, B.; Mezlini, A. M.; Demir, F.; Fiume, M.; Tu, Z.; Brudno, M.; Haibe-Kains, B.; Goldenberg, A. Similarity Network Fusion for Aggregating Data Types on a Genomic Scale. *Nat. Methods* **2014**, *11* (3), 333−337.
(321) Mo, Q.; Shen, R.; Guo, C.; Vannucci, M.; Chan, K. S.; Hilsenbeck, S. G. A Fully Bayesian Latent Variable Model for Integrative Clustering Analysis of Multi-Type Omics Data. *Biostatistics* **2018**, *19* (1), 71−86.
(322) Rohart, F.; Gautier, B.; Singh, A.; Lê Cao, K.-A. MixOmics: An R Package for 'omics Feature Selection and Multiple Data Integration. *PLoS Comput. Biol.* **2017**, *13* (11), No. e1005752.
(323) Argelaguet, R.; Velten, B.; Arnol, D.; Dietrich, S.; Zenz, T.; Marioni, J. C.; Buettner, F.; Huber, W.; Stegle, O. Multi-Omics Factor Analysis—a Framework for Unsupervised Integration of Multi-omics Data Sets. *Mol. Syst. Biol.* **2018**, *14* (6), No. e8124.
(324) Liu, T.; Salguero, P.; Petek, M.; Martinez-Mira, C.; Balzano-Nogueira, L.; Ramšak, Ž.; McIntyre, L.; Gruden, K.; Tarazona, S.; Conesa, A. PaintOmics 4: New Tools for the Integrative Analysis of Multi-Omics Datasets Supported by Multiple Pathway Databases. *Nucleic Acids Res.* **2022**, *50* (W1), W551−W559.
(325) Hamzeiy, H.; Ferretti, D.; Robles, M. S.; Cox, J. Perseus Plugin "Metis" for Metabolic-Pathway-Centered Quantitative Multi-Omics Data Analysis for Static and Time-Series Experimental Designs. *Cell Reports Methods* **2022**, *2* (4), No. 100198.
(326) Calinawan, A. P.; Song, X.; Ji, J.; Dhanasekaran, S. M.; Petralia, F.; Wang, P.; Reva, B. ProTrack: An Interactive Multi-Omics Data Browser for Proteogenomic Studies. *Proteomics* **2020**, *20*, 21−22.
(327) Ballard, J. L.; Wang, Z.; Li, W.; Shen, L.; Long, Q. Deep Learning-Based Approaches for Multi-Omics Data Integration and Analysis. *BioData Min* **2024**, *17* (1), 38.
(328) Adossa, N.; Khan, S.; Rytkönen, K. T.; Elo, L. L. Computational Strategies for Single-Cell Multi-Omics Integration. *Comput. Struct Biotechnol J.* **2021**, *19*, 2588−2596.
(329) Cao, Z.-J.; Gao, G. Multi-Omics Single-Cell Data Integration and Regulatory Inference with Graph-Linked Embedding. *Nat. Biotechnol.* **2022**, *40* (10), 1458−1466.
(330) Zhang, W.; Mou, M.; Hu, W.; Lu, M.; Zhang, H.; Zhang, H.; Luo, Y.; Xu, H.; Tao, L.; Dai, H.; Gao, J.; Zhu, F. MOINER: A Novel Multiomics Early Integration Framework for Biomedical Classification and Biomarker Discovery. *J. Chem. Inf Model* **2024**, *64* (7), 2720−2732.
(331) Gao, Y.; Cao, D.; Li, M.; Zhao, F.; Wang, P.; Mei, S.; Song, Q.; Wang, P.; Nie, Y.; Zhao, W.; Wang, S.; Yan, H.; Wang, X.; Jiao, Y.; Liu, Q. Integration of Multiomics Features for Blood-Based Early Detection of Colorectal Cancer. *Mol. Cancer* **2024**, *23* (1), 173.
(332) Flores, J. E.; Claborne, D. M.; Weller, Z. D.; Webb-Robertson, B.-J. M.; Waters, K. M.; Bramer, L. M. Missing Data in Multi-Omics Integration: Recent Advances through Artificial Intelligence. *Front Artif Intell* **2023**, *6*, No. 1098308.
(333) Yetgin, A. Revolutionizing Multi-omics Analysis with Artificial Intelligence and Data Processing. *Quantitative Biology* **2025**, *13* (3), No. e70002.
(334) Picard, M.; Scott-Boyer, M.-P.; Bodein, A.; Périn, O.; Droit, A. Integration Strategies of Multi-Omics Data for Machine Learning Analysis. *Comput. Struct Biotechnol J.* **2021**, *19*, 3735−3746.
(335) Liu, W.; Vu, T.; Konigsberg, I. R.; Pratte, K. A.; Zhuang, Y.; Kechris, K. J. Smccnet 2.0: A Comprehensive Tool for Multi-Omics Network Inference with Shiny Visualization. *BMC Bioinf.* **2024**, *25* (1), 276.
(336) Singh, A.; Shannon, C. P.; Gautier, B.; Rohart, F.; Vacher, M.; Tebbutt, S. J.; Lê Cao, K.-A. DIABLO: An Integrative Approach for Identifying Key Molecular Drivers from Multi-Omics Assays. *Bioinformatics* **2019**, *35* (17), 3055−3062.
(337) Li, Z.; Katz, S.; Saccenti, E.; Fardo, D. W.; Claes, P.; Martins dos Santos, V. A. P.; Van Steen, K.; Roshchupkin, G. V. Novel Multi-Omics

Deconfounding Variational Autoencoders Can. Obtain Meaningful Disease Subtyping. *Brief Bioinform* **2024**, *25* (6), No. bbae512.
(338) Zheng, Y.; Liu, Y.; Yang, J.; Dong, L.; Zhang, R.; Tian, S.; Yu, Y.; Ren, L.; Hou, W.; Zhu, F.; Mai, Y.; Han, J.; Zhang, L.; Jiang, H.; Lin, L.; Lou, J.; Li, R.; Lin, J.; Liu, H.; Kong, Z.; Wang, D.; Dai, F.; Bao, D.; Cao, Z.; Chen, Q.; Chen, Q.; Chen, X.; Gao, Y.; Jiang, H.; Li, B.; Li, B.; Li, J.; Liu, R.; Qing, T.; Shang, E.; Shang, J.; Sun, S.; Wang, H.; Wang, X.; Zhang, N.; Zhang, P.; Zhang, R.; Zhu, S.; Scherer, A.; Wang, J.; Wang, J.; Huo, Y.; Liu, G.; Cao, C.; Shao, L.; Xu, J.; Hong, H.; Xiao, W.; Liang, X.; Lu, D.; Jin, L.; Tong, W.; Ding, C.; Li, J.; Fang, X.; Shi, L. Multi-Omics Data Integration Using Ratio-Based Quantitative Profiling with Quartet Reference Materials. *Nat. Biotechnol.* **2024**, *42* (7), 1133−1149.
(339) Fouodo, C. J. K.; Bleskina, M.; Szymczak, S. FuseMLR: An R Package for Integrative Prediction Modeling of Multi-Omics Data. *BMC Bioinformatics* **2025**, *26* (1), 221.
(340) Sharifi-Noghabi, H.; Zolotareva, O.; Collins, C. C.; Ester, M. MOLI: Multi-Omics Late Integration with Deep Neural Networks for Drug Response Prediction. *Bioinformatics* **2019**, *35* (14), i501−i509.
(341) Mohr, A. E.; Ortega-Santos, C. P.; Whisner, C. M.; Klein-Seetharaman, J.; Jasbi, P. Navigating Challenges and Opportunities in Multi-Omics Integration for Personalized Healthcare. *Biomedicines* **2024**, *12* (7), 1496.
(342) Sanches, P. H. G.; de Melo, N. C.; Porcari, A. M.; de Carvalho, L. M. Integrating Molecular Perspectives: Strategies for Comprehensive Multi-Omics Integrative Data Analysis and Machine Learning Applications in Transcriptomics, Proteomics, and Metabolomics. *Biology (Basel)* **2024**, *13* (11), 848.

https://doi.org/10.1021/acs.jmedchem.5c01986
J. Med. Chem. XXXX, XXX, XXX−XXX