Proteomic tools to study phosphorylation of intrinsically disordered proteins

Barbara Spolaore[a], Luca Secco[b], Giulia Rocca[c], Guidalberto Manfioletti[b], Giorgio Arrigoni[c] and Riccardo Sgarra[b]

[a]Department of Pharmaceutical and Pharmacological Sciences, University of Padua, Padova, Italia; [b]Life Sciences Department, University of Trieste, Trieste, Italy; [c]Department of Biomedical Sciences, University of Padova, Padova, Italy

ABSTRACT

Introduction: Intrinsically disordered proteins (IDPs) represent a family of proteins that lack secondary or tertiary structure. IDPs are hubs in interaction networks, participate in liquid–liquid phase separation processes, and drive the formation of proteinaceous membrane-less organelles. Their unfolded structure makes them particularly prone to post-translational modifications (PTMs) that play key functional modulatory roles.

Areas covered: We discuss different analytical approaches to study phosphorylation of IDPs starting from methods for IDP enrichment (strong acid extractions and heat-based pre-fractionation), strategies to enrich and map phosphopeptides/proteins, and mass spectrometry-based tools to study the phosphorylation-dependent conformational alterations of IDPs (limited proteolysis, HDX, chemical cross-linking, covalent labeling, and ion mobility).

Expert opinion: There is a growing interest in IDPs and their PTMs since they are involved in several diseases. The intrinsic disorder could be exploited to facilitate purification and synthetic production of IDPs taking full advantage of those structural mass-spectrometry-based methods that can be used to investigate IDPs and their phospho-dependent conformational alterations. The diffusion and implementation of mass spectrometers with ion mobility devices and electron transfer dissociation capabilities could be key-elements for increasing information on IDP biology.

1. Introduction

The perception that disordered or unstructured regions in proteins could have functional roles expanded in the 1990s when different studies showed that these amino acid sequences where more than just linker regions, providing of the necessary flexibility to allow different domains to juxtapose. Recently, Uversky and Kulkarni published an interesting chronology of research that led to the concept of a function being associated with intrinsic disorder [1]. Some relevant examples of the first functional evidence of the roles of unstructured protein regions regard (i) the yeast alpha 2 protein that relies on the presence of an unstructured region to form a complex with MCM1 and to bind to an operator [2] and (ii) the p65 subunit of the NF-kappa B transcription factor and its disordered C-terminal tail. This was demonstrated to be essential for its transactivating ability, working with an 'induced fit' mechanism implying a disorder/order transition upon binding to specific targets/partners [3]. Additional evidences provided by limited proteolysis experiments performed on NF-κB pointed in the same direction. These indicated how the AB loop of p50, not involved in the direct contact with DNA, was readily subjected to proteolytic cleavages in the DNA-unbound state, but was highly protected upon contact of NF-κB with DNA, thus suggesting a disorder/order transition guided by DNA binding [4]. Evidences of a disorder/order transition were also provided by NMR analyses of the flagellum-specific sigma factor, sigma 28, and its inhibitor, FlgM. The latter undergoes ordering that involves up to 50% of

its amino acid sequence on binding to sigma 28 [5]. However, it was only in 2000 that the term 'intrinsically disordered proteins (regions) – IDP(R)s' became popular, recognizing the peculiarities, unicity, and diffusion of this particular 'family' of proteins.

Dunker at al. in 2001 [6] showed that IDP(R)s can be recognized based on the amino acid composition and indeed they showed that some amino acids are underrepresented in IDP(R)s (W, C, F, I, Y, V, L, and N), while others are strongly enriched (A, R, G, Q, S, P, E, and K), and proposed to call them order-promoting and disorder-promoting amino acids, respectively. In addition, other attributes can be exploited to identify the propensity of a protein region to be disordered, such as for instance the hydropathy and the net charge. All these indications prompted the development of tools for the prediction of disordered regions within proteins. To date, about 100 such predictors have been developed exploiting different methods [7], and given the importance of these tools, there has been a drive toward the evaluation and comparison of their performances [7,8]. The Critical Assessment of protein Intrinsic Disorder prediction (CAID) [8] evaluated the performance of 43 methods for their effectiveness in predicting intrinsically disordered regions (IDRs), fully disordered proteins (IDPs), and disorders binding regions. fIDPnn (http://biomine.cs.vcu.edu/servers/fIDPnn/), SPOT-Disorder2 (https://sparks-lab.org/server/spot-disorder2/), RawMSA (https://bitbucket.org/clami66/rawmsa/src/master/), and AUCpreD (http://raptorx2.uchicago.edu/StructurePropertyPred/predict/) were among the top performing tools/algorithms [8].

Article highlights

- Intrinsically disordered proteins (IDPs) represent an emerging heterogeneous family of proteins involved in different diseases;
- IDPs are highly post-translationally modified and phosphorylation is key in modulating their functions;
- IDPs can be extracted/pre-fractionated by unconventional methods such as strong acid- or heat-based extractions;
- The characterization of IDP phosphorylation can take advantage of consolidated phosphopeptide enrichment strategies;
- The conformational role of phosphorylation must be studied with non-conventional structural methods such as limited proteolysis, hydrogen-deuterium exchange (HDX) mass spectrometry, chemical cross-linking mass spectrometry (XL-MS), covalent labeling mass spectrometry, ion mobility mass spectrometry (IM-MS), and electron transfer dissociation (ETD);
- The spread of mass spectrometers with both IM and ETD capabilities would pave the way in boosting the characterization of this family of proteins from both a conformational and functional point of view.

The importance of IDPs/IDRs is underlined also by the fact that several annotated databases for IDPs/IDRs are available, each providing different and relevant information [9].

A peculiar property of IDPs is their ability to establish a myriad of interactions with nucleic acids (both DNA and RNA), membranes, and other proteins, and for this reason, they have been defined as networks' hubs [10]. IDP interaction modalities differ on the basis of the specific modules employed: these being molecular recognition features (MoRFs), short linear motifs (SLiMs), and low-complexity regions (LCRs). MoRFs usually undergo a disorder-to-order transition upon binding, and SLiMs can retain disorder on binding, while LCRs are highly repetitive modules with an elevated binding promiscuity [11].

IDP's interactions are explained by different models: (i) two state binding, in which a single module of the IDP interacts with a single site on the partners; (ii) avidity, in which multiple specific sites on both the IDP and the partner are involved with a cooperative effect; (iii) allovalency, in which multiple modules on the IDP compete for the same site on the binding partner; (iv) fuzzy binding, in which multiple sites on both the IDP and the binding partner are present and can establish short-lived interactions also at the same moment [12].

It is important to highlight that, since in the fuzzy binding modality several interactions occur in parallel, they all contribute to a 'cooperative' effect that leads to high-affinity interactions. However, the fact that disorder is maintained and that each single interaction is short lived allows for a higher accessibility to the surfaces involved in the interactions, thus making their modulation by PTMs possible; taken together, these features lead to stable complexes that can be easily disassembled [8].

This high interactivity has been demonstrated to be key for the participation of IDPs in liquid–liquid phase separation processes that are at the basis of the formation of proteinaceous membrane-less organelles (PMLOs). PMLOs are peculiar intracellular subdomains (both nuclear and cytoplasmic) formed following the increase of certain critical factors, temperature fluctuations, PTMs, and other alterations. Within these domains, proteins and nucleic acids can be stored or different activities can be strongly enhanced by providing high local concentrations of the factors responsible for these processes/transformations [13,14]. Key features of IDPs have been summarized in Figure 1. For more details about IDPs, their involvement in LLPS, and their PTMs, we suggest a series of recent reviews, particularly focused on these specific topics [15–20].

It seems clear that PTMs are strategic for the modulation of IDP's functions and, as a matter of fact, given their high accessibility, they undergo a plethora of PTMs [21]. Among these, phosphorylation is the most studied and perhaps the most relevant modification affecting IDPs' functions [22].

In this review, we will focus on several technical aspects regarding the methods used to study phosphorylation of IDPs, ranging from sample preparation and enrichment strategies for IDPs, to mass spectrometry-based approaches to map and identify phosphorylation sites, and then to the analytical strategies that can be adopted to extrapolate conformational information regarding structural alterations induced by this PTM.

2. Article body

2.1. IDP enrichment methods

In general, one of the main problems that must be overcome when studying specific protein subfamilies is related to the wide dynamic range of expression levels. Indeed, very abundant proteins quite often mask the detection of those that are present at lower levels. Moreover, specifically concerning the study of protein phosphorylation, an additional obstacle derives from the fact that phosphorylations usually occur at sub-stoichiometric levels, therefore making phosphopeptide enrichment methods essential (see below). In light of these considerations, any strategy that could assist in focusing on a specific subset of proteins can help in providing a better information coverage. As highlighted in the introduction, IDPs have a peculiar amino acid composition that causes them to behave differentially from conventionally 3D-structured proteins. When a protein with a canonical set of secondary structural elements forming a tertiary structure is exposed to harsh conditions (i.e. extreme pH values or high temperature), it usually undergoes a denaturation/misfolding process, which results in the exposure of the inner hydrophobic residues that are normally in the core. This phenomenon triggers an aggregation process that ends up with protein precipitation. IDPs have a very low content of hydrophobic residues and lack stable secondary and tertiary structures and, therefore, when exposed to these harsh conditions, do not aggregate or precipitate. This peculiar feature has led to the development of two IDP enrichment strategies based on the application of high temperature or very low pH values (Figure 2).

Regarding the first strategy, the various proposed protocols are comparable. After the extraction of proteins from the biological sample of interest, temperatures ranging from 95°C to 100°C can be used for either a short time (5 or 10 min) or for a longer period (1 h), yielding similar results. Despite the time element, the key factor is temperature: heating samples at 60°C for 10 min results in lower IDP enrichment [23]. The same method has also been
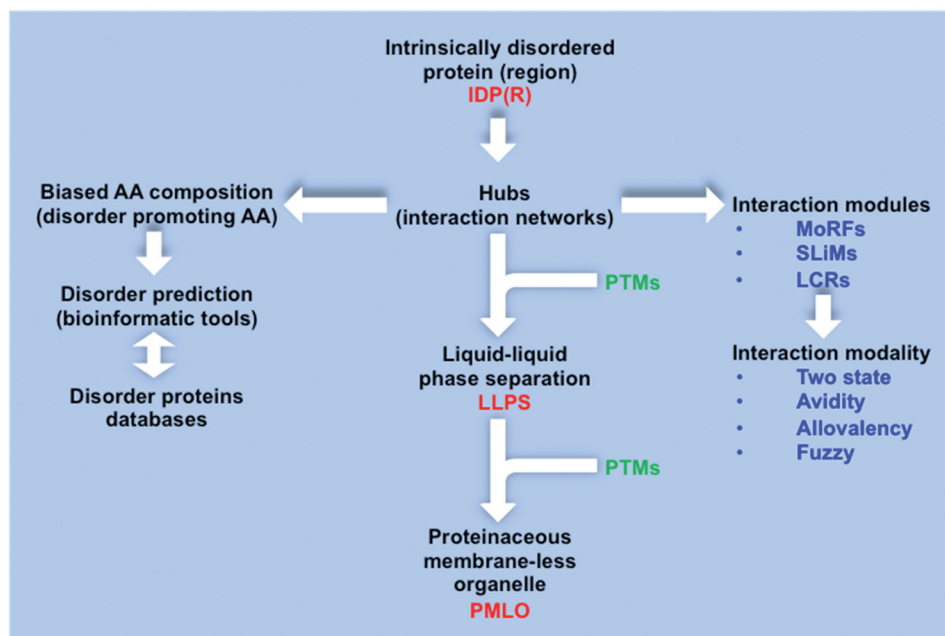
Figure 1. Main features of intrinsically disordered proteins (IDPs). Intrinsically disordered proteins (IDPs) and intrinsically disordered protein regions (IDPRs) are hubs in protein/protein interaction networks. They are characterized by different modules, i.e. molecular recognition features (MoRfs), short linear motifs (SLiMs) and low-complexity regions (LCRs) that enable them to interact with other factors using different strategies: two state, avidity, allovalency, and fuzzy. IDPs have a biased amino acid composition and this allowed the development of disorder prediction tools that in turn aided the generation of databases of IDPs. The main relevant biological aspect of IDPs is that they participate in liquid-liquid phase separation (LLPS) processes thus significantly contributing to the generation of proteinaceous membrane-less organelles (PMLOs).

proposed to enrich IDPs starting from plant samples [24]. After the heating step, samples can be centrifuged at 4°C for 30 min at high speed (from 20.400 g up to 100.000 g) [25,26], or centrifuged after cooling on ice for 15 min [23,24,27,28]. This allows heat-unstable proteins to precipitate, while IDPs are enriched in the supernatant. If required, IDPs can then be precipitated with cold (−20°C) trichloroacetic acid (TCA, 20% solution), followed by cold (−20°C) acetone, and finally pelleted through centrifugation [23,24,27,28]. In mammalian biological samples, this strategy permits to enrich up to 69–77% of IDPs [26,27] and up to 80% for enrichment performed on nuclear extracts [28]. To underline the reliability of this enrichment strategy, it is worthwhile noting that the total number of the heat-stable proteins extracted with these procedures and identified using high-sensitivity methods are comparable among the different published works [26,27].

Regarding the second strategy, i.e. the application of very low pH values, slightly different protocols have been proposed, but the common characteristic is the use of diluted (5%) perchloric acid (PCA) solution. PCA can be directly present in the extraction solution [29] or added to the supernatant after cell lysis [30,31]. In the first case, the solution is then frozen (−20°C)/thawed/vortexed (10 s) thrice and centrifuged at 15,000 g for 10 min. To precipitate acid-soluble proteins, the clear supernatant is collected, and cold trifluoroacetic acid (TFA) is added to a final concentration of 33%. Precipitation occurs on ice for 1 h and proteins are collected through centrifugation as described previously, washed once with 0.2% HCl in acetone, and twice with pure acetone. In the second case, the sample is kept on ice for 15 min after addition of PCA and then centrifuged at 4500 g for

20 min and the supernatant collected. To precipitate acid-soluble proteins, 8 volumes of acetone are added and left at −20°C overnight. The use of different acids, like trichloroacetic acid (TCA) at different concentrations has been evaluated [30]: 3% PCA is the minimum quantity to allow for a strong enrichment. 5% and 9% PCA do not result in different spot patterns on 2D-gels, suggesting that 5% can be considered a suitable concentration to perform the enrichment. TCA is less selective compared with PCA at the same percentage, but results in a different spot pattern on the gels, suggesting that the combination of these two extraction methods can increase the coverage of the acid-soluble proteins. Cortese and colleagues, using a small set of already characterized proteins, estimated that about 70% of the acid-soluble proteins are IDPs [30]. Despite the small sample of proteins considered, this percentage is concordant with our analysis performed on the 154 acid-soluble proteins isolated by Zougman and Wiśniewski [29]. Using PONDR software, we estimated that 74% of these proteins have a percentage of disorder higher than 30% (for 55.8% of the proteins, the percentage is higher than 50%). Using the same classification of Galea et colleagues [23], it turns out that 61.7% of proteins are disordered proteins (DPs), 35.1% are folded proteins (FPs) and 3.3% are classified as mixed proteins (MXPs), confirming the specificity of the enrichment strategy for IDPs.

Comparing heat-based and acid-based enrichment strategies, it is clear that there is a relevant difference in terms of identified IDPs when using these two methods (about 1300/1400 obtained using the heat-based method versus only 154 adopting the acid-based one). To the best of our knowledge,

Figure 2. Prefractionation strategies for enriching intrinsically disordered proteins. Heat- and strong acid-based extraction methods exploit the precipitation of non-IDPs (folded proteins) in these harsh conditions (high temperature/low pH) to enrich for IDPs that, due to their strong hydrophilic characteristic, remain in solution and are then recovered by precipitation using hydrophobic solvents.

the work of Zougman and Wiśniewski, published in 2005, is the last that attempted to define the so-called 'unfoldome' by acid-based enrichment; on the other hand, heat-based strategies are more recent (2009 and 2022). This difference might be explained by the lower sensitivity of the mass spectrometer that was used in the first case so that a new investigation of the unfoldome derived from acid treatment, using more advanced, high-resolution MS systems, should be considered. In addition, a comparison between the heat- and acid-stable proteins has never been performed. We are convinced that this analysis should be conducted to better understand the common and different features of the two subsets, which could lead to a better characterization of IDPs.

2.2. MS-based structural methods

From a structural point of view, IDPs or IDRs exist as an ensemble of interconverting conformers. PTMs, and in particular phosphorylation, can affect this ensemble and in the case of IDPs can modulate structural elements and enable conversion into a conformation able to mediate a specific biological function. IDPs are indeed highly modified *in vivo*, and in particular, phosphorylation often occurs in flexible protein regions [32,33]. The acquisition of a distinct conformation by an IDP or IDR can be important to determine the function of the protein, such as the involvement in a particular cell signaling pathway, its localization, or its degradation. For this reason, the structural characterization of conformational changes induced by phosphorylation becomes important. Several studies on the structural effect of phosphorylation have been reported in the literature [34,35]. Adding a phosphate group to Ser, Thr, Tyr, and His introduces a negatively charged moiety in the sequence in place of neutral or partially cationic (His) residues. In IDPs, this modification can induce intramolecular contacts due to specific hydrogen bonds established by the phosphate group or to nonspecific electrostatic effects, can stabilize helix dipoles or simply provide a novel binding site to phosphoprotein-binding domains.

The structural biology of conformational effects induced by PTMs, and in particular by phosphorylation, is quite challenging. IDPs are not amenable to crystallization and even if the disordered region is part of a larger globular arrangement, in X-ray structures they result in missing electron density. NMR spectroscopy has instead been successfully applied to the in-depth characterization of the conformational changes induced by phosphorylation on IDPs [35,36].

In this review, we will focus on the possibility to apply structural mass spectrometry (MS) methods to the characterization of the conformational changes that phosphorylation induces in IDPs or IDRs. In particular, we provide an overview of limited proteolysis (LiP), hydrogen-deuterium exchange (HDX), chemical cross-linking (XL), covalent labeling, and ion mobility (IM) (Figure 3). All these techniques, which exploit mass spectrometry, do not allow the residue-resolution of NMR, but have the advantage to analyze proteins under physiological conditions, without a molecular mass limit or the need of isotopic labeling [37]. Moreover, MS-based techniques require only minute amounts of protein, which is an advantage in the context of phosphorylated proteins, given the difficulty to quantitatively reproduce this modification *in vitro* and to purify the modified protein in sufficient amounts [34]. In particular, with *in vitro* phosphorylation it is challenging to obtain a site-specific phosphorylation of IDPs due to the accessibility of nonspecific sites that determine an excess of phosphorylation [36]. The production of a large protein containing intrinsically disordered regions can also be limited by its high tendency to degrade and fragment during the purification procedures [38]. Finally, MS-based techniques allow to identify and study the presence of discrete conformers of a protein in a conformational ensemble, as is often the case in IDPs [39]. In general, structural MS techniques provide complementary data that can be combined to unravel the
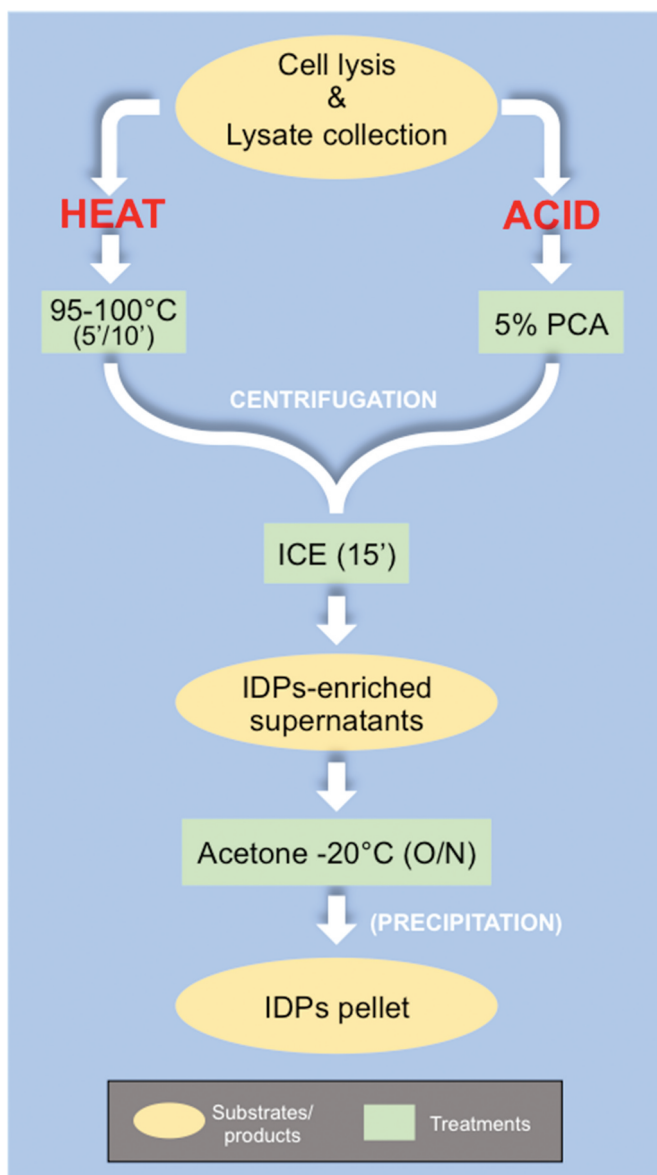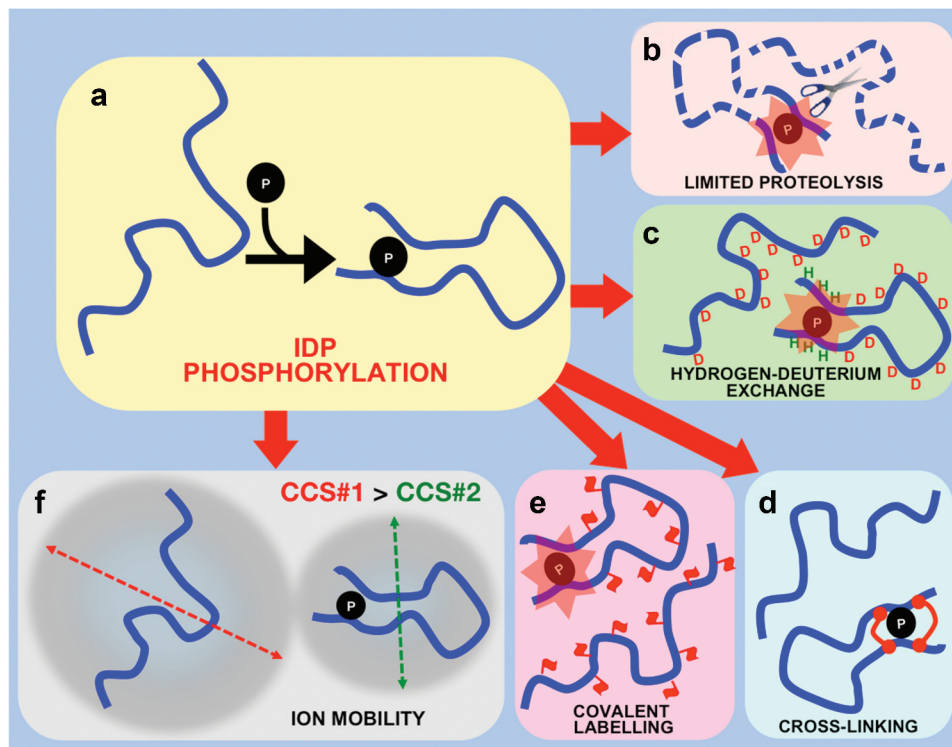
Figure 3. Intramolecular interactions induced by IDP phosphorylation and mass spectrometry–based methods that can be used to characterize conformational changes. (A) In this example, phosphorylation of the IDP favors a conformation in which the *N*- and C-termini establish long range contacts. (B) Limited proteolysis. In the phosphorylated IDP, the *N*- and C-termini of the protein display a higher resistance to protease hydrolysis. (C) Hydrogen – deuterium exchange. After phosphorylation, protein regions involved in intramolecular interactions show a slower exchange rate of amide hydrogens with deuterium. (D) Cross-linking mass spectrometry. Cross-linkers allow to link two functional groups that after phosphorylation have been placed in close spatial proximity in the structure of the protein. (E) Covalent labeling mass spectrometry. After phosphorylation, protein regions involved in intramolecular interactions are not accessible to chemical labeling. (F) Ion mobility. The conformational change induced by phosphorylation can be investigated by measuring the collision cross-sectional (CCS) area of the protein.

conformational preferences of proteins in an integrative structural approach [40]. In the following paragraphs, we provide a concise background of the structural mass spectrometry techniques considered for the conformational characterization of phosphorylated IDPs and a description of studies in which these methodologies were applied.

2.2.1. Limited proteolysis

Limited proteolysis (LiP) exploits proteases to map regions in the protein structure that are exposed and flexible or that undergo a conformational change in different conditions [41,42]. In a LiP experiment, the protein substrate is incubated with the protease at very low enzyme to substrate ratios and for short times (generally only minutes). Aliquots of the reaction mixture are quenched for enzymatic activity at different times and analyzed by LC-MS [43]. Mass spectrometry allows to measure the accurate mass of the large fragments of the protein that are produced after the initial nicking by the protease and thus to determine the first sites of hydrolysis in the polypeptide sequence. It has been previously demonstrated on several proteins, whose structure was determined by NMR or X-ray crystallography, that the first sites of hydrolysis occur at the level of flexible regions of the protein substrates such as those characterized by a high crystallographic temperature factor (B value) [44]. The localization in the

sequence of the sites of hydrolysis is determined, besides by the accessibility provided by the conformation of the protein, also by the specificity of the enzyme. Generally, a broad-specificity protease such as proteinase K is preferred since this allows to hydrolyze the protein substrate almost independently of the sequence. Interestingly, when proteases of different specificities are used, they all hydrolyze the protein in the same narrow region [45]. In the case of IDPs, LiP can be used to identify disordered regions in proteins, but also conformational changes induced upon binding or due to the introduction of PTMs. Since IDPs-IDRs are hydrolyzed very rapidly due to their flexible conformation, it is necessary to adjust the digestion conditions to map the sites of initial proteolysis, for example, by lowering the amount of enzyme used. A recent interesting development of LiP is its application on a proteomic scale in cell lysates to study structural changes induced upon molecular interactions [46]. Drug affinity responsive target stability (DARTS) is a method linked to LiP in which the target protein of a small molecule is identified based on a change in susceptibility to proteolysis upon binding. This possibility is certainly of interest with respect to IDPs since it could aid to decipher the conformational changes induced in IDPs by the binding of small molecules as well as the influence of PTM-induced conformational alterations toward the binding of specific drugs [47].

2.2.2. Hydrogen deuterium exchange mass spectrometry (HDX-MS)

In a typical HDX experiment, the protein sample under physiological solution conditions is diluted at least 10-fold in the same buffer prepared in D_2O, leading to the exchange of backbone amide hydrogens with deuterium [48,49]. In HDX-MS, the mass difference of 1 Da between hydrogen and deuterium allows the analysis of the extent of deuterium incorporation upon accurate mass measurement of the biomolecule. In a typical HDX workflow, hydrogens are allowed to exchange with deuterium for different incubation times ranging from seconds to hours and the reaction is quenched by dilution in ice-cold acid, where the labeling rate is slower. The exchange can be monitored by MS on the intact protein (global exchange) or at the level of peptides produced by digestion of the protein with a protease active at acidic pH (local exchange). The rate of exchange of backbone amide hydrogens is determined by the presence of hydrogen bonding, of local secondary structure, as well as by their solvent accessibility. In the case of IDPs, due to the lack of secondary structure, the exchange is fast and short deuteration times (milliseconds) are used to unravel the ensembles of conformational states that are present at the equilibrium under physiological conditions [50]. Quench-flow systems [51] or rapid mixing devices [52] allow the analysis of HDX exchange on millisecond labeling times and thus to study the conformational features of disordered regions of proteins.

PTEN (phosphatase and tensin homologue deleted on chromosome 10) is a tumor suppressor that can be inactivated by mutations in many types of solid tumors. This 47 kDa protein is a challenging structural target due to its intrinsically disordered C-terminal segment and the conformational heterogeneity caused by phosphorylation of its C-terminus. Phosphorylation of the intrinsically disordered C-terminal region of the protein appears to inhibit the phosphatase activity of the protein and to influence its membrane binding. HDX-MS was used to study the effect of phosphorylation of the C-terminal tail on the conformational properties of the protein [53]. The results demonstrated that phosphorylation at the level of six residues of the disordered tail causes a decrease in HDX at the phosphatase active site and at the membrane-binding interface of PTEN. These effects suggest that upon phosphorylation, the tail establishes a contact with these regions of the protein, causing auto-inhibitory interactions. Analysis by HDX-MS was also performed on the phosphorylated form of the protein modified at only four sites (instead of six) on the C-terminal tail since experimental data indicated that this modified form shows phosphatase activity but still lower membrane binding. HDX-MS data confirmed that phosphorylation at four sites induces the formation of an intermediate state with a reduced protection to HDX at the active site of the phosphatase domain. Conformational changes induced in PTEN by phosphorylation at four sites of the C-terminal tail were also monitored by LiP with trypsin [54], revealing a higher resistance to hydrolysis for the phosphorylated form, which is indicative of a more structured conformation.

In neuronal cells, the IDP α-synuclein (αSyn) is involved in Parkinson's disease (PD) in which it forms fibrils in Lewy bodies (LB) and Lewy neurites (LN). In LB and LN, 96% of αSyn is phosphorylated, but the contribution of phosphorylation *in vivo* to the aggregation process is still not understood [55]. Monomeric αSyn has three characteristic main regions: an N-terminal region that is positively charged, a central non-amyloid β component (NAC) region that is hydrophobic and forms the core of fibrils during aggregation, and a negatively charged C-terminus that can bind metal ions. As an IDP, αSyn populates an ensemble of conformations in solution. In order to study the effect of phosphorylation of Ser129 on αSyn, HDX-MS was used to unravel the protection to HD exchange of the protein after 30 s of incubation at 20°C in D_2O buffer at pD 7.2 [56]. Upon comparison with wild type (WT) αSyn, there was no significant difference in exchange due to phosphorylation. However, the same experiments were conducted in the presence of calcium since cytosolic calcium concentration increases in PD, and it was shown that it enhances the aggregation rate of WT αSyn. Calcium binds at the level of the acidic C-terminus of the protein. In the presence of calcium, in both WT and phosphorylated αSyn, a difference in HD exchange is observed in the central region (NAC region), which becomes deprotected and, at the C-terminus that, in turn, shows an increase in protection. However, while in phosphorylated αSyn the HD exchange of the N-terminus is unaffected by calcium binding, under the same conditions in WT αSyn it shows a higher HDX rate indicating that it becomes more solvent exposed likely due to the disruption of long-range contacts. Since *in vitro* in the presence of calcium phosphorylation lowers the aggregation propensity of αSyn, these differences in the level of protection of the N-terminus between phosphorylated and non-phosphorylated αSyn suggest a role of this region of the protein in the aggregation process.

HDX-MS was also applied to the study of phosphorylated tau, an IDP that has the physiological role to stabilize microtubules in the neuronal axis but that in tauopathies such as Alzheimer's disease is involved in the formation of amyloid aggregates, named neurofibrillary tangles (NFT). Hyperphosphorylation of tau appears to have a role in the pathology, since this modification inhibits the physiological function of tau and favors the formation of NFT [57], making it important to understand the conformational changes induced by phosphorylation. To this aim, hyperphosphorylated and native tau have been compared by HDX-MS using the Time-Resolved ElectroSpray Ionization (TRESI) HDX methodology that allows to monitor HDX resulting from labeling times between 42 ms and 8000 ms [58,59]. TRESI-HDX kinetics indicated that hyperphosphorylation of tau causes the adoption of a more extended structure characterized by a general increase in deuterium uptake. Measured changes in HDX along different regions of the phosphorylated protein were rationalized with the observation that hyperphosphorylation increases the amyloidogenicity propensity of tau.

2.2.3. Chemical cross-linking mass spectrometry (XL-MS)

In the XL-MS technique, chemical cross-linking of proteins and MS analysis are combined to study protein structure and interactions. Cross-linking allows to link two functional groups

that are in close spatial proximity in the structure of the protein or in a protein complex using a bifunctional modification reagent (cross-linker) with a defined length [60,61]. The formation of the covalent linkage between residues can be determined by enzymatic digestion of the modified proteins and LC-MS analysis of the peptides, often performed after an enrichment step of the cross-linked species in order to facilitate their assignment. The length of the cross-linker works as a 'molecular ruler' that allows to estimate distances between cross-linked residues and it can be used in computational modeling to derive protein 3D-structures and map protein interfaces [62,63]. XL-MS has been applied to IDPs to characterize the ensemble of protein conformations [37]. An advantage is that most used cross-linking reagents target lysine residues, and these charged residues are often present in high number in IDPs and IDRs. IDPs involved in neurodegenerative diseases were studied by XL-MS with the aim to derive model structures of the monomeric protein as for tau protein [64] and α-synuclein [65] or to unravel molecular mechanisms of aggregation of disease-associated variants as for tau [66].

2.2.4. Covalent labelling mass spectrometry

Covalent labeling of proteins and protein complexes allows to obtain structural information by mapping the accessibility of certain residues to chemical modification with a modifying reagent [67,68]. Again, the location in the sequence of the modification can be detected by MS since it determines a specific mass addition, while an estimate of differences in derivatization can be useful to monitor conformational changes between two states of a protein. Several chemical probes for protein labeling are currently used. Among these, important labeling agents are hydroxyl radicals that react preferentially with aromatic, heterocyclic, and sulfur-containing amino acid side chains [69]. They are produced with different methods including laser-induced photolysis of hydrogen peroxide in fast photochemical oxidation of proteins (FPOP) [70] and synchrotron radiolysis of water [71]. In FPOP, the lifetime of the radicals can be controlled by introducing a scavenger, thus allowing to tune the timescale of the experiment up to ~1 μs and to monitor the modification only of reactive and solvent-exposed residues avoiding over-labeling [72]. As an example of application in the field of IDPs, FPOP has been used to monitor changes in solvent accessibility of amino acid side chain in peptide Aβ1 − 42 during the aggregation process [73].

2.2.5. Ion mobility

Ion mobility devices exploit the principle for which molecules with the same m/z, traveling in a buffer gas and subjected to a weak electric field, move with different velocity as a function of the friction they experience. This is strictly dependent upon their collision cross section (CCS), i.e. the extent of the surfaces colliding with gas molecules [74]. Several ion mobility platforms differing in their separation principles are nowadays commercially available: TWIMS (traveling wave ion mobility spectrometry), DTIMS (drift tube ion mobility spectrometry), and TIMS (trapped ion mobility spectrometry) [75]. Regardless of the differences between these technologies, all provide the

possibility to follow a specific m/z value and to visualize the different conformations of a protein, as peaks differ from each other for their drift time; the higher the drift time, the more extended the conformation and vice-versa.

As was previously anticipated, IDPs exist as an ensemble of interconverting conformers due to the fact that their folding energy funnel landscape is made of several small funnels separated by small energy barriers. If these conformers differ in their compactness (i.e. CCS) they can be spatially resolved and visualized separately by a mass analyzer placed downstream the ion mobility device.

Ion mobility has been widely used to probe the presence of different conformations of IDPs, especially as concerns proteins involved in neurodegenerative diseases such as alpha-synuclein and tau. Some recent examples regard the effect of singly (Na+, K+, and Cs+) versus multiply (Ca2+, Mg2+, Cu2+, Zn2+, and La3+) charged metal ions on the αSyn compaction [76], the binding and accelerating effect of ScSERF (the S. cerevisiae homolog of human SERF and Caenorhabitis elegans MOAG) with respect to the αSyn amyloid formation [77] and the propensity of different tau isoforms to form oligomeric complexes [78]. However, notwithstanding the fact that IDPs are among the most post-translationally modified proteins [32,79–82], very few works focus on the influence of phosphorylation on conformational transitions in IDPs.

High Mobility Group A proteins are chromatin architectural factors often cited as prototypes of IDPs and characterized by a negatively charged acidic tail and three highly positively charged DNA binding domains. It was demonstrated by IM-MS that the acidic tail was responsible for the presence of a compact conformer. Indeed, when a C-terminal truncated HMGA form missing the acidic tail was analyzed, IM-MS evidenced the presence of a slower protein conformation (higher drift time) not present in the wild-type form. Since multiple CK2 phosphorylation sites are embedded on the HMGAs' acidic tail, the effect of these PTMs was also evaluated and it turned out that HMGA phosphorylations were responsible for a shift toward a more compact protein conformation [83]. Limited proteolysis experiments using Lys-C digestion supported these data, showing that the central region of the protein is more prone to hydrolysis in the truncated form with respect to the full-length protein, confirming a shielding effect of the C-terminal tail on these regions. IDPs are strongly enriched in charged amino acid residues and in certain cases can be considered as polyampholytes. It has been first shown by atomistic Metropolis Monte Carlo simulations that conformations of IDPs can be influenced by the linear sequence distributions of oppositely charged residues [84] and later, taking p27Kip1 as a model, that the extent of linear mixing versus segregation of oppositely charged residues influences the ion mobility-measured collision cross section and thus its conformational distribution [85]. This was exactly the case for HMGA proteins; indeed, the three DNA-binding domains and the acidic tail represent oppositely charged protein regions that can undertake electrostatic interactions. These interactions are responsible for the formation of a sort of loop that was demonstrated to shield HMGA proteins to the accessibility of other HMGAs modifying enzymes [83]. In a recent NMR

study, the phosphorylation of HMGA1a was demonstrated to strengthen long-range contacts between the C-terminal tail and adjacent regions, in agreement with the IM-MS and limited proteolysis results obtained on HMGA2 and further demonstrating the impact of phosphorylation on the structural ensemble of IDPs [86]. HMGA proteins are among the most post-translational modified nuclear proteins [87], and it was demonstrated that phosphorylation at different sites by different kinases (CK2 and cdc2) lead to different conformational ensembles and that phosphorylation by CK2 impairs DNA binding as corroborated by isothermal titration calorimetry (ITC)-based analyses. These findings support the interest to study the effect of phosphorylation on IDPs to explain the consequences of this modification on the function of these proteins and to support the development of therapeutic strategies.

Studying the conformational effects of the molecular tweezer CLR01 on the tau protein, it was observed an effect of phosphorylation toward the compactness of tau/CLR01 complex. CLR01 is able to bind to Lys residues, acting as a nanochaperone preventing the formation of oligomers and aggregates [88], and it is of particular interest in those neurodegenerative pathologies characterized by aggregation of specific proteins such as the Alzheimer's and Parkinson's diseases. Since a pathological hallmark of Alzheimer's disease is the accumulation of hyperphosphorylated tau protein [89], the effect of phosphorylation on a tau fragment was evaluated with respect to the binding of CLR01. Mono- and bi-phosphorylated tau fragments were more shifted toward a compact conformation with respect to the unmodified fragment, suggesting a phosphorylation dependency on the conformational effect of the molecular tweezer CLR01.

2.3. Phospho-enrichment strategies

Despite the tremendous progresses that phosphoproteomics has witnessed in the past 20 years, the study of phosphorylated proteins remains very challenging. Since phosphorylation is a transient and sub-stoichiometric modification, a phosphoproteomic analysis can require up to 100 times more material compared to a classical whole-proteome study [90]. Protein phosphorylation is one of the most common PTM, and indeed, the number of phosphorylation sites that have been mapped since the advent of proteomics is extremely high. In the PhosphositePlus database (www.phosphosite.org) almost 300,000 phosphosites are currently listed, and it has been suggested that phosphorylation events might involve more than 90% of the expressed proteome [91]. However, due to the dynamic nature of this PTM, only part of these sites actually exists at a given time [92]. Furthermore, among this population, there may be phospho-isoforms that differ in the phosphorylation sites as well as in the occupancy of each of them [93].

To address this daunting task, mass spectrometry (MS)-based approaches have been proven to be the most appropriate. Standard methods are based on the complete digestion of the proteins into peptides with the use of specific proteases. The limited dynamic range of the current instrumentation, however, still represents a major obstacle in the identification of the sub-stoichiometric phosphopeptides in shotgun experiments [94]. Conventionally, in these experiments, the mass spectrometer operates in a data-dependent acquisition (DDA) mode. The consequence is that only the most abundant peptides are selected for the MS/MS scans, while the low abundant ones are generally lost. Data-independent acquisition (DIA), in which all peptides present in a predetermined m/z window are selected and co-fragmented, has emerged as a powerful alternative to DDA, but it has not yet been broadly applied to protein phosphorylation analysis, because of the difficulties in implementing the workflow and the necessity of a reference spectral library [95,96]. A notable example of the application of DIA-MS to a phosphoproteomic study is the work of Bekker-Jensen et al. [95]. These authors included a phosphorylation localization algorithm in the DIA computational workflow that allows the profiling of phosphoproteomes in a robust and reproducible manner, without the need for experiment-specific spectral libraries. Moreover, very recently, a comprehensive study comparing several software tools and workflows for DIA applications has been published [97]. Using specifically designed benchmark data, the authors compare the performance of several DIA software packages for proteomics and phosphoproteomics, highlighting the very good performance of some of these tools also with in-silico generated libraries, therefore without the need for building experiment-specific spectral libraries.

In any case, to correctly identify the low abundant phosphoproteins, improve the coverage, and determine the localization of the modifications, a prior enrichment step is often required to reduce the complexity of the samples [93,96]. Enrichment strategies, commonly performed at the peptide level, can be divided into different categories. Some of the most commonly used techniques are based on chromatography and involve the interaction with metal ions (IMAC), metal oxides (MOAC), or with hydrophilic resins (HILIC). Alternatively, immunoprecipitation with specific antibodies can be performed. Other, less established protocols are based on the chemical derivatization of phosphate groups.

In the following paragraphs, we provide a brief description of the most popular strategies that can be used in MS-based phosphoproteomics, as a standard and well-accepted procedure is still lacking.

2.3.1. Chromatographic techniques
2.3.1.1. Strong Cation Exchange (SCX). SCX, often performed prior to MS/MS analysis to pre-fractionate complex samples, was first used as an enrichment strategy in 2004 [98]. At low pH (~2.7), cation exchange columns retain peptides with positive charges, which are then eluted with increasing salt concentrations. It can be assumed that phosphorylated peptides are eluted earlier than their non-phosphorylated counterparts because their interaction with the anionic stationary phase is weaker since the negatively charged phosphate groups lower the affinity for the resin. Thus, the early-eluting fractions are enriched in phosphopeptides. The main drawback of the SCX enrichment is its low selectivity: peptides with other acidic modifications can be

enriched as well, while phosphopeptides with multiple basic amino acids can be lost [93]. For this reason, SCX has often been applied in combination with other enrichment strategies, such as IMAC or MOAC.

2.3.1.2. IMAC.

Immobilized metal affinity chromatography (IMAC) is one of the most widely used enrichment strategies. IMAC relies on the strong affinity between transition metal cations (such as Fe^{3+} or Ga^{3+}) non-covalently immobilized on a resin, and the negatively charged phosphate groups [96]. IMAC columns can also suffer from low specificity because peptides that contain multiple acidic amino acids tend to strongly bind to the resin. This problem, however, can be reduced by lowering the pH of the buffer below 3 to protonate the acidic residues (carboxylic acids have a pKa of ~5), while the phosphate groups remain negatively charged (pKa ~2) [99]. The elution is then conducted at a basic pH (>10).

To date, various IMAC supports for phosphoproteomic analysis are commercially available, and research on new supports and materials is actively ongoing. Recently, some very promising new chelating agents and matrices have been developed, such as core-shell composite microspheres based on nanodiamond that can bind titanium ions (Ti^{4+}) and thus be used for selective phosphopeptide enrichment [100].

2.3.1.3. MOAC.

MOAC (Metal oxide affinity chromatography) is another very popular metal-based chromatography method, which differs from the IMAC technology for the use of metal oxide matrices, such as titanium dioxide (TiO_2), zirconium dioxide (ZrO_2) or magnetite (Fe_3O_4) [96]. Interestingly, even though both IMAC and MOAC have been widely used for large-scale applications, different protocols show a modest overlap [90]. For example, a comparative analysis of four different enrichment protocols (including TiO_2 and IMAC enrichment) performed in melanoma A375 cells concluded that each method allows for the isolation of different fractions of the phosphoproteome [101], indicating that the implementation of more sensitive, selective, and reproducible enrichment methods is still required. To overcome the diverse enrichment preferences of the different methods, several protocols that combine more than one strategy have been developed, such as the Sequential Elution from IMAC (SIMAC method) [102].

2.3.1.4. Phos-tag.

A popular alternative to these techniques is represented by phos-tag molecules. Phos-tag is a binuclear metal complex developed by Kinoshita and coworkers that, unlike the other methods mentioned, can specifically capture phosphate groups at a neutral pH [103], a characteristic that makes it particularly suited for the analysis of acidic labile phosphoresidues such as histidine and aspartate. Various strategies based on phos-tag derivatives have been applied in phosphoproteomics studies, such as Phos-tag polymer beads that can be used to perform affinity chromatography [104].

2.3.2. Antibodies-based enrichment

Immunoprecipitation with specific antibodies is another method extensively used for the investigation of the phosphorylated residues (pTyr, pSer, and pThr). The identification of phosphorylated tyrosine residues is particularly challenging because of the very low abundance of this PTM in proteins. Indeed, p-Tyr residues account for <1% of the total phosphorylation events [91]. In this context, specific antibodies against p-Tyr residues turned out to be particularly useful because the methods discussed so far allow for the detection only of a small number of pTyr-containing proteins, in line with the lower level of pTyr compared to pSer and pThr [105]. To date, pTyr-specific antibodies have been applied in studies addressing the role of tyrosine phosphorylation and allowed the recovery of thousands of pTyr-peptides with high specificity [96]. The main drawback of their usage, apart from the cost, is the large amount of biological material required [106]. As discussed for the previous strategies, immunoprecipitation with pTyr antibodies can also be combined with other techniques to improve selectivity and efficiency of the enrichment.

It is worth mentioning that although phosphorylation mostly occurs on the canonical sites (Ser, Thr, and Tyr), six other residues, namely His, Lys, Arg, Asp, Glu, and Cys, can be phosphorylated as well [96]. Various attempts to isolate these phosphopeptides have been made, but since they are very unstable in the acidic conditions necessary for phospho-enrichment and LC-MS analyses they have not been deeply investigated; however, current efforts are predominantly focused on pHis because of its unique properties [107]. Indeed, pHis can act as an enzyme intermediate in the transfer of the phosphate moiety to other residues and can exist in the form of two isomers that have different reactivity and stability [108]. In 2015, the development of monoclonal antibodies against each of these isomers [109] helped to accelerate the characterization of the pHis phosphoproteome: interestingly, in eukaryotic cells, pHis sites tend to be located in unstructured regions exposed to the solvent [110].

2.3.3. General principles of quantitative phosphoproteomics

Quantitative phosphoproteomic experiments are generally performed using the same methods that are applied for the 'classical' proteome-wide studies, such as label-free quantification and techniques based on chemical labeling. The removal from the datasets of phosphopeptides with a high number of missing values has been shown to be important to reduce variations and obtain more robust quantifications [111]. Moreover, a normalization step is generally required to minimize systematic biases introduced in the sample preparation, which is often not very reproducible when phosphopeptide enrichment methods are applied. However, how to properly normalize the abundances of phosphopeptides is one of the main issues still associated with quantitative phosphoproteomics [112]. Indeed, most of the normalization methods used for proteomics data assume that the abundances of the majority of the peptides do not vary between the samples, but this may not be true in the context of phosphoproteomics because there could be global changes in the phosphorylation if kinases/phosphates are altered. The normalization could be performed using the abundance of the proteins from which phosphopeptides were generated, but a pairwise normalization method has also been proposed, which uses the

phosphopeptides identified both in the enriched and in the non-enriched samples to adjust the abundance values in label-free experiments [112]. Alternatively, PhosR, an R package for phosphoproteomic workflow, uses 'stably phosphorylated sites' (SPSs) to normalize the data and remove unwanted variations [113].

2.3.4. Tools for the assignment of phosphorylation sites

It is well known that determining the precise localization of the phosphosites within peptide sequences can be very challenging. In fact, the unambiguous localization of phosphorylation sites requires the identification of specific product ions in the MS/MS spectra, which can be used to discriminate between the possible candidate sites [111]. To address this issue several localization tools have been developed: some of them, as the Ascore algorithm [114], assign a probability of correct phosphosite localization based on the presence and intensity of site-determining fragment ions in MS/MS spectra, others, as Mascot delta score [115], use the difference between the scores of the highest ranked peptide spectrum matches (PSMs) to assess the confidence of the localization [116,117].

The popular software MaxQuant (freely available at http://www.maxquant.org/) integrates the PTM-score algorithm, while MS Amanda search engine (freely available at https://ms.imp.ac.at/?goto=msamanda) employs as localization score the phosphoRS algorithm [118]. The commercial software Proteome Discoverer (Thermo Fisher Scientific) can exploit both the delta score and the phosphoRS tool to assess the confidence in the localization of phosphorylation sites.

In 2020, Locard-Paulet et al. conducted an extensive evaluation study on the performances of three software suites (MaxQuant, Proteome Discoverer, and PeptideShaker), investigating different combinations of search engines, PSMs validation strategies, and localization score algorithms [119]. This investigation showed that results obtained with different pipelines are not easily comparable and highlighted the importance of setting a proper score threshold for each algorithm to limit the rate of false localization.

2.4. Electron transfer dissociation for the study of IDPs

Electron transfer dissociation (ETD) is a protein/peptide fragmentation process that is substantially different from collision-induced dissociation (CID) and higher-energy collisional dissociation (HCD) strategies: while the later accelerate ions so that they collide with a neutral gas causing the fragmentation, the former uses the transfer of a low energy electron on the polypeptide chain in order to initiate a series of molecular rearrangements finally leading to fragment generation [120]. CID/HCD fragmentation generates b/y fragments (rupture of the amidic C-N bond), while ETD causes cleavage of the Cα-N bond of the different amino acids, generating c/z fragments. CID/HCD works preferentially with 2+/3+ charged peptides that have been generated by trypsin, Lys-C, or Arg-C digestion in order to obtain peptides with a length of 4–20 residues carrying a positive charge at the C-terminus; longer peptides or peptides with a higher charge state are not fragmented efficiently. On the contrary, ETD fragmentation requires only that the charge state is higher or equal to 2+ without limitation concerning the enzyme used for generating the peptides [121]. The main problem regarding the study of protein/peptide phosphorylation by CID/HCD is the preferential loss of phosphoric acid (loss of 98 Da) from the phosphorylated S/T/Y residues, implying that the fragmentation pattern is dominated by the m/z peak corresponding to the neutral loss of phosphoric acid and that other fragments are poorly detectable, thus impairing the localization/mapping of the phosphorylation site [122]. This limitation is generic and does not apply only to IDPs. However, there are some considerations that can be made with respect to the advantages of the application of ETD in the context of IDPs:

a. IDPs have a bias in their amino acid content, and indeed, they are rich in basic residues (K and R) that are recognized and cleaved by the proteolytic enzymes preferentially used in proteomics (trypsin, Lys-C, and Arg-C). This means that when analyzing IDPs with conventional protocols, there is a higher probability to obtain low mass peptides or highly charged ones that are not ideal for the CID-based PTM mapping process. ETD fragmentation has been shown to overcome this issue because of the possibility of utilizing 'middle-down' proteomic strategies to obtain sequencing information from longer peptides. Indeed, it has been adopted to map phosphorylation and methylation sites in arginine-rich RNA binding proteins [123,124];

b. ETD provides sequence information also on very long peptides/short proteins [125,126]. This aspect is relevant especially concerning the mapping of multiple PTMs on the same amino acid chain. In the introduction, we have underlined that IDPs are heavily post-translationally modified (due to their higher accessibility), thus implying that different combinations of PTMs can be present on the same protein. Being able to gain information on the mapping of different co-existing PTMs is a necessary step to comprehend the PTM code, its role in influencing the conformational ensemble of IDPs and thus their functional implications. Obviously, these considerations apply also to non-IDPs, but we believe they are particularly relevant for this protein family. An example of the utility of ETD for the characterization of coexisting PTMs comes from a study on histone H3 [127]. Histones are considered to be IDPs [128], and their protruding tails are subjected to a myriad of PTMs whose different combinations constitute the histone epigenetic code [129]. In particular, ETD allowed the identification of 114 different combinations of PTMs on histone H3 extracted from mouse embryonic stem cells [127].

c. ETD also opens the possibility of obtaining information at the level of single amino acids upon fragmentation of the peptides that have undergone HDX. Indeed, the electron-driven fragmentation process in ETD is suitable for HDX experiments since it prevents intramolecular hydrogen/deuterium migration (hydrogen scrambling) that, conversely, occurs using CID [130]. However, there are constraints in the

applicability of this technique given the requirement for multiply charged ions.

d. It has been shown that, in the context of top-down approaches for the characterization of protein structures by electron-driven fragmentation, flexible protein regions turn out to be preferentially fragmented with respect to those having a well-defined and rigid 3D structure [131]. This peculiarity could be of significant interest in studying those IDRs that are embedded in structured proteins, either to confirm and identify the IDRs themselves, or to map phosphorylation specifically affecting these regions.

It is worthwhile to underline that instrumentation with ETD capabilities was initially confined to costly FT-ICR MS systems using the electron capture dissociation (ECD) method and then extended to low-resolution tridimensional or linear traps. For logical reasons, this limited the applicability of this fragmentation technique. Nowadays, however, this limit has been overcome, and high-performance orbitrap- and Q/TOF-based mass spectrometers are available for ETD experiments.

3. Conclusion

We have presented a brief overview of different strategies that can be used to study phosphorylation and its effect on the conformational ensemble of intrinsically disordered proteins. We believe that the key to approaching to the field of IDPs is 'combination.' Indeed, it is evident that there is no single method or analytical tool that can overcome all the problems intrinsic to IDP analysis. However, the combination of different approaches, starting from IDP prefractionation strategies (strong acid- or heat-prefractionations), phospho-enrichment, and mapping tools to the conformational analysis by limited proteolysis, HDX, chemical cross-linking, covalent labeling, ion mobility, and ETD can aid in solving functional questions regarding IDPs' phosphorylation.

4. Expert opinion

The field of IDPs is an emerging one, and a deeper understanding of IDP mechanisms of action would be of considerable impact on the understanding of the pathological conditions they are connected with, also beyond their well-established involvement in neurodegenerative disorders, such as αSyn and tau for Parkinson's and Alzheimer's diseases, respectively. It is indeed evident nowadays that IDPs and their PTMs are widely involved in a plethora of diseases.

IDPs possess a defining feature, that is the lack of secondary or tertiary structure. Despite the fact that this characteristic has often been considered as a limitation in the understanding of their mechanism of action, it is important to underline that it allows to adopt less stringent conditions for their purification and even to chemically synthesize them (at least for those proteins with a relatively low number of amino acids). This last possibility is an important strategy to produce IDPs with site-specific PTMs. For the same reasons, the production and purification of recombinant IDPs in bacteria could also be less problematic with respect to non-IDPs. We believe

that these features will bring further improvements to the pipeline required to obtain 'pure' IDPs. These can then be subjected to the conformational studies essential to understand their conformational landscapes but can also be used to study the role of PTMs in contributing to or driving structural/conformational shifts. In our opinion, one of the main problems is to unravel the function(s) of IDRs found within proteins that are otherwise structured to a large percentage, because of the well-known limitations in the process of purification of folded proteins.

We are surprised by the fact that many IDPs with well-established impact on cell biology have not yet been explored by conformational mass spectrometry approaches. For example, stathmin, cyclin-dependent kinase inhibitors 1, 1C, and 1B, eukaryotic translation initiation factor 4E-binding protein 1, nuclear protein 1, and PCNA-associated factor are all 100% disordered proteins listed in the Disprot database (https://www.disprot.org), whose activity is modulated by phosphorylation but that have not been investigated (to our knowledge, at least).

We believe that the IDP field will greatly benefit when specific technologies, and in particular ion mobility mass spectrometry, will gain popularity. As a matter of fact, an estimate of its limited diffusion can be obtained by a simple PubMed search using 'protein ion mobility mass spectrometry' as query, which returns 1761 entries, while searching for 'protein mass spectrometry' returns 168,177 results. This means that currently, the impact of ion mobility mass spectrometry is roughly only about 1% with respect to all the applications of mass spectrometry in the protein field.

One of the areas that could reasonably progress in the future is the application of ECD/ETD in combination with middle-down or top-down approaches to elucidate the combinatorial PTM code and the cross-talk between different PTMs and to obtain relevant information on IDRs and their PTMs within structured proteins. Indeed, while the availability of MS systems with ECD/ETD capabilities is already a well-established reality, there is still an urgent need for the development of more robust and efficient tools for middle-down and top-down proteomic approaches. In this context, the use of new and non-canonical proteases for proteomic studies will probably allow for a deeper and more comprehensive mapping of PTMs, also in the field of IDPs.

Furthermore, in the case of the HDX-MS technique, the application of ETD/ECD will move it toward a more quantitative description of the change in HDX rate at the level of single amide hydrogens. This single-residue-resolution HDX combined with the possibility to analyze millisecond exchange times will open to the possibility of obtaining a more detailed description of the conformational changes induced by phosphorylation of IDPs. An important development will also be an improvement of computational tools that allow to integrate HDX data with results from other structural MS techniques to obtain 3D structural models of the conformational ensemble within IDPs and of the conformational changes induced by their phosphorylation. Indeed, these models can advance our understanding at a molecular level of the differences in biological function determined by the introduction of PTMs.

The growing interest in the IDP field, in our opinion, will be a stimulus for the optimization of pipelines aimed at the generation of synthetic/recombinant IDPs carrying specific PTMs. These 'pure' proteins will represent powerful tools to study conformational changes induced by specific modifications and to elucidate how different combinations of PTMs influence the balance of IDP conformational ensembles.

The game changer in the IDP field could be the recognition of the essential role played by IDPs in the liquid–liquid phase separation process and thus in the organization of proteinaceous membraneless organelles. This could provide a relevant push toward investing in the comprehension of IDP biology.

Declaration of interest

The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Reviewer disclosures

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

Author contributions

Writing original draft: all authors; Visualization: B Spolaore, L Secco, and R Sgarra; Reviewing: all authors; Editing: all authors; Conceptualization: B Spolaore, G Arrigoni, and R Sgarra. All authors read and approved the final version of the manuscript for publication.

References

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Uversky VN, Kulkarni P. Intrinsically disordered proteins: chronology of a discovery. Biophys Chem. 2021;279:106694.
2. Vershon AK, Johnson AD. A short, disordered protein region mediates interactions between the homeodomain of the yeast alpha 2 protein and the MCM1 protein. Cell. 1993;72:105–112.
3. Schmitz ML, dos Santos Silva MA, Altmann H, et al. Structural and functional analysis of the NF-kappa B p65 C terminus. An acidic and modular transactivation domain with the potential to adopt an alpha-helical conformation. J Biol Chem. 1994;269:25613–25620.
4. Matthews JR, Nicholson J, Jaffray E, et al. Conformational changes induced by DNA binding of NF-kappa B. Nucleic Acids Res. 1995;23:3393–3402.
5. Daughdrill GW, Chadsey MS, Karlinsey JE, et al. The C-terminal half of the anti-sigma factor, FlgM, becomes structured when bound to its target, sigma 28. Nat Struct Biol. 1997;4:285–291.
6. Dunker AK, Lawson JD, Brown CJ, et al. Intrinsically disordered protein. J Mol Graph Model. 2001;19:26–59.
7. Zhao B, Kurgan L. Surveying over 100 predictors of intrinsic disorder in proteins. Expert Rev Proteomics. 2021;18:1019–1029.
   • This paper provides a relevant overview of predictors of intrinsic disorder.
8. Necci M, Piovesan D, Predictors CAID, et al. Critical assessment of protein intrinsic disorder prediction. Nat Methods. 2021;18:472–481.
9. Piovesan D, Monzon AM, Quaglia F, et al. Databases for intrinsically disordered proteins. Acta Crystallogr D Struct Biol. 2022;78:144–151.
10. Uversky VN. Analyzing IDPs in Interactomes. Methods Mol Biol. 2020;2141:895–945.
11. Morris OM, Torpey JH, Isaacson RL. Intrinsically disordered proteins: modes of binding with emphasis on disordered domains. Open Biol. 2021;11:210222.
12. Olsen JG, Teilum K, Kragelund BB. Behaviour of intrinsically disordered proteins in protein-protein complexes with an emphasis on fuzziness. Cell Mol Life Sci. 2017;74:3175–3183.
    •• This paper provides a comprehensive review on IDPs' binding modality.
13. Uversky VN. Supramolecular fuzziness of intracellular liquid droplets: liquid-liquid phase transitions, membrane-less organelles, and intrinsic disorder. Molecules. 2019;24:3265.
    •• This paper provides a short review on IDPs and their involvement in liquid–liquid phase transition and in the assembly of proteinaceous membrane-less organelles.
14. Uversky VN. Recent developments in the field of intrinsically disordered proteins: intrinsic disorder-based emergence in cellular biology in light of the physiological and pathological liquid-liquid phase transitions. Annu Rev Biophys. 2021;50:135–156.
15. Fonin AV, Antifeeva IA, Kuznetsova IM, et al. Biological soft matter: intrinsically disordered proteins in liquid-liquid phase separation and biomolecular condensates. Essays Biochem. 2022;66:831–847.
16. Borcherds W, Bremer A, Borgia MB, et al. How do intrinsically disordered protein regions encode a driving force for liquid-liquid phase separation? Curr Opin Struct Biol. 2021;67:41–50.
17. Darling AL, Zaslavsky BY, Uversky VN. Intrinsic disorder-based emergence in cellular biology: physiological and pathological liquid-liquid phase transitions in cells. Polymers. 2019;11:990.
18. Abyzov A, Blackledge M, Zweckstetter M. Conformational dynamics of intrinsically disordered proteins regulate biomolecular condensate chemistry. Chem Rev. 2022;122:6719–6748.
19. Wu S, Wen J, Perrett S. Unravelling the microscopic characteristics of intrinsically disordered proteins upon liquid-liquid phase separation. Essays Biochem. 2022;66:891–900.
20. Darling AL, Uversky VN. Intrinsic disorder and posttranslational modifications: the darker side of the biological dark matter. Front Genet. 2018;9:158.
21. Pejaver V, Hsu WL, Xin F, et al. The structural and functional signatures of proteins that undergo multiple events of post-translational modification. Protein Sci. 2014;23:1077–93.
22. Newcombe EA, Delaforge E, Hartmann-Petersen R, et al. How phosphorylation impacts intrinsically disordered proteins and their function. Essays Biochem. 2022;66:901–913.
23. Galea CA, Pagala VR, Obenauer JC, et al. Proteomic studies of the intrinsically unstructured mammalian proteome. J Proteome Res. 2006;5:2839–2848.
    • This paper is the first example of a heat-based prefractionation method for the enrichment of IDPs.
24. Ginsawaeng O, Gorka M, Erban A, et al. Characterization of the heat-stable proteome during seed germination in Arabidopsis with special focus on LEA proteins. Int J Mol Sci. 2021;22:8172.
25. Csizmók V, Szollosi E, Friedrich P, et al. A novel two-dimensional electrophoresis technique for the identification of intrinsically unstructured proteins. Mol & Cell Proteomics. 2006;5:265–273.
26. Park H, Yamanaka T, Nukina N. Proteomic analysis of heat-stable proteins revealed an increased proportion of proteins with compositionally biased regions. Sci Rep. 2022;12:4347.
27. Galea CA, High AA, Obenauer JC, et al. Large-scale analysis of thermostable, mammalian proteins provides insights into the intrinsically disordered proteome. J Proteome Res. 2009;8:211–226.

28. Skupien-Rabian B, Jankowska U, Kedracka-Krok S. Analysis of a nuclear intrinsically disordered proteome. Methods Mol Biol. 2020;2175:181–196.

29. Zougman A, Wiśniewski JR. Beyond linker histones and high mobility group proteins: global profiling of perchloric acid soluble proteins. J Proteome Res. 2006;5:925–934.
   • **This paper is the first example of an acid-based prefractionation method for the enrichment of IDPs.**

30. Cortese MS, Baird JP, Uversky VN, et al. Uncovering the unfoldome: enriching cell extracts for unstructured proteins by acid treatment. J Proteome Res. 2005;4:1610–1618.

31. Tantos A, Tompa P. Proteomic methods for the identification of intrinsically disordered proteins. Methods Mol Biol. 2012;896:429–437.

32. Iakoucheva LM, Radivojac P, Brown CJ, et al. The importance of intrinsic disorder for protein phosphorylation. Nucleic Acids Res. 2004;32:1037–1049.

33. Csizmok V, Forman-Kay JD. Complex regulatory mechanisms mediated by the interplay of multiple post-translational modifications. Curr Opin Struct Biol. 2018;48:58–67.

34. Bah A, Forman-Kay JD. Modulation of intrinsically disordered protein function by post-translational modifications. J Biol Chem. 2016;291:6696–6705.

35. Kumar A, Narayanan V, Sekhar A. Characterizing post-translational modifications and their effects on protein conformation using NMR spectroscopy. Biochemistry. 2020;59:57–73.

36. Bah A, Vernon RM, Siddiqui Z, et al. Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. Nature. 2015;519:106–109.

37. Beveridge R, Calabrese AN. Structural proteomics methods to interrogate the conformations and dynamics of intrinsically disordered proteins. Front Chem. 2021;9:603639.
   • **An interesting recent review on structural mass spectrometry methods useful for the characterization of IDPs.**

38. Wright PE, Dyson HJ. Intrinsically disordered proteins in cellular signalling and regulation. Nat Rev Mol Cell Biol. 2015;16:18–29.

39. Beveridge R, Chappuis Q, Macphee C, et al. Mass spectrometry methods for intrinsically disordered proteins. Analyst. 2013;138:32–42.

40. Rout MP, Sali A. Principles for integrative structural biology studies. Cell. 2019;177:1384–1403.

41. Fontana A, Polverino de Laureto P, De Filippis V, et al. Probing the partly folded states of proteins by limited proteolysis. Fold Des. 1997;2:R17–26.

42. Fontana A, de Laureto PP, Spolaore B, et al. Probing protein structure by limited proteolysis. Acta Biochim Pol. 2004;51:299–321.

43. Fontana A, de Laureto PP, Spolaore B, et al. Identifying disordered regions in proteins by limited proteolysis. Methods Mol Biol. 2012;896:297–318.

44. Fontana A, Fassina G, Vita C, et al. Correlation between sites of limited proteolysis and segmental mobility in thermolysin. Biochemistry. 1986;25:1847–1851.

45. Fontana A, Zambonin M, Polverino de Laureto P, et al. Probing the conformational state of apomyoglobin by limited proteolysis. J Mol Biol. 1997;266:223–230.

46. Feng Y, De Franceschi G, Kahraman A, et al. Global analysis of protein structural changes in complex proteomes. Nat Biotechnol. 2014;32:1036–1044.

47. Lomenick B, Hao R, Jonai N, et al. Target identification using drug affinity responsive target stability (DARTS). Proc Natl Acad Sci U S A. 2009;106:21984–21989.

48. Konermann L, Pan J, Liu Y-H. Hydrogen exchange mass spectrometry for studying protein structure and dynamics. Chem Soc Rev. 2011;40:1224–1234.

49. Masson GR, Burke JE, Ahn NG, et al. Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. Nat Methods. 2019;16:595–602.

50. James EI, Murphree TA, Vorauer C, et al. Advances in hydrogen/deuterium exchange mass spectrometry and the pursuit of challenging biological systems. Chem Rev. 2022;122:7562–7623.

51. Rist W, Rodriguez F, Jørgensen TJD, et al. Analysis of subsecond protein dynamics by amide hydrogen exchange and mass spectrometry using a quenched-flow setup. Protein Sci. 2005;14:626–632.

52. Keppel TR, Weis DD. Analysis of disordered proteins using a simple apparatus for millisecond quench-flow H/D exchange. Anal Chem. 2013;85:5161–5168.

53. Masson GR, Perisic O, Burke JE, et al. The intrinsically disordered tails of PTEN and PTEN-L have distinct roles in regulating substrate specificity and membrane activity. Biochem J. 2016;473:135–144.
   •• **This paper demonstrates the potential of HDX-MS for the identification of IDRs in proteins and to characterise conformational changes induced by phosphorylation of IDRs.**

54. Bolduc D, Rahdar M, Tu-Sekine B, et al. Phosphorylation-mediated PTEN conformational closure and deactivation revealed with protein semisynthesis. Elife. 2013;2:e00691.

55. Anderson JP, Walker DE, Goldstein JM, et al. Phosphorylation of Ser-129 is the dominant pathological modification of alpha-synuclein in familial and sporadic Lewy body disease. J Biol Chem. 2006;281:29739–29752.

56. Stephens AD, Zacharopoulou M, Moons R, et al. Extent of N-terminus exposure of monomeric alpha-synuclein determines its aggregation propensity. Nat Commun. 2020;11:2820.

57. González A, Singh SK, Churruca M, et al. Alzheimer's disease and tau self-assembly: in the search of the missing link. Int J Mol Sci. 2022;23:4192.

58. Zhu S, Shala A, Bezginov A, et al. Hyperphosphorylation of intrinsically disordered tau protein induces an amyloidogenic shift in its conformational ensemble. PLoS ONE. 2015;10:e0120416.
   • **This paper provides an interesting example of HDX-MS analysis on a millisecond time scale of a native and phosphorylated IDP using the TRESI-HDX technique.**

59. Rob T, Liuni P, Gill PK, et al. Measuring dynamics in weakly structured regions of proteins using microfluidics-enabled subsecond H/D exchange mass spectrometry. Anal Chem. 2012;84:3771–3779.

60. Piersimoni L, Kastritis PL, Arlt C, et al. Cross-linking mass spectrometry for investigating protein conformations and protein-protein interactions – a method for all seasons. Chem Rev. 2022;122:7500–7531.

61. Leitner A, Faini M, Stengel F, et al. Crosslinking and mass spectrometry: an integrated technology to understand the structure and function of molecular machines. Trends Biochem Sci. 2016;41:20–32.

62. Green NS, Reisler E, Houk KN. Quantitative evaluation of the lengths of homobifunctional protein cross-linking reagents used as molecular rulers. Protein Sci. 2001;10:1293–1304.

63. Brodie NI, Popov KI, Petrotchenko EV, et al. Solving protein structures using short-distance cross-linking constraints as a guide for discrete molecular dynamics simulations. Sci Adv. 2017;3:e1700479.

64. Popov KI, Makepeace KAT, Petrotchenko EV, et al. Insight into the Structure of the "Unstructured" Tau Protein. Structure. 2019;27:1710–1715.e4.

65. Brodie NI, Popov KI, Petrotchenko EV, et al. Conformational ensemble of native α-synuclein in solution as determined by short-distance crosslinking constraint-guided discrete molecular dynamics simulations. PLoS comput Biol. 2019;15:e1006859.

66. Chen D, Drombosky KW, Hou Z, et al. Tau local structure shields an amyloid-forming motif and controls aggregation propensity. Nat Commun. 2019;10:2493.

67. Chance MR, Farquhar ER, Yang S, et al. Protein footprinting: auxiliary engine to power the structural biology revolution. J Mol Biol. 2020;432:2973–2984.

68. Liu XR, Zhang MM, Gross ML. Mass spectrometry-based protein footprinting for higher-order structure analysis: fundamentals and applications. Chem Rev. 2020;120:4355–4454.

69. Xu G, Chance MR. Hydroxyl radical-mediated modification of proteins as probes for structural proteomics. Chem Rev. 2007;107:3514–3543.

70. Liu XR, Rempel DL, Gross ML. Protein higher-order-structure determination by fast photochemical oxidation of proteins and mass spectrometry analysis. Nat Protoc. 2020;15:3942–3970.

71. Maleknia SD, Brenowitz M, Chance MR. Millisecond radiolytic modification of peptides by synchrotron X-rays identified by mass spectrometry. Anal Chem. 1999;71:3965–3973.

72. Li KS, Shi L, Gross ML. Mass spectrometry-based Fast Photochemical Oxidation of Proteins (FPOP) for higher order structure characterization. Acc Chem Res. 2018;51:736–744.

73. Li KS, Rempel DL, Gross ML. Conformational-sensitive fast photochemical oxidation of proteins and mass spectrometry characterize amyloid beta 1–42 aggregation. J Am Chem Soc. 2016;138:12090–12098.

74. Valentine SJ, Liu X, Plasencia MD, et al. Developing liquid chromatography ion mobility mass spectrometry techniques. Expert Rev Proteomics. 2005;2:553–565.

75. Webb IK. Recent technological developments for native mass spectrometry. Biochim Biophys Acta Proteins Proteom. 2022;1870:140732.

76. Moons R, Konijnenberg A, Mensch C, et al. Metal ions shape α-synuclein. Sci Rep. 2020;10:16293.

77. Meinen BA, Gadkari VV, Stull F, et al. SERF engages in a fuzzy complex that accelerates primary nucleation of amyloid proteins. Proc Natl Acad Sci U S A. 2019;116:23040–23049.

78. Shahpasand-Kroner H, Portillo J, Lantz C, et al. Three-repeat and four-repeat tau isoforms form different oligomers. Protein Sci. 2022;31:613–627.

79. Collins MO, Yu L, Campuzano I, et al. Phosphoproteomic analysis of the mouse brain cytosol reveals a predominance of protein phosphorylation in regions of intrinsic sequence disorder. Mol & Cell Proteomics. 2008;7:1331–1348.

80. Holt LJ, Tuch BB, Villén J, et al. Global analysis of Cdk1 substrate phosphorylation sites provides insights into evolution. Science. 2009;325:1682–1686.

81. Dyson HJ. Expanding the proteome: disordered and alternatively folded proteins. Q Rev Biophys. 2011;44:467–518.

82. Owen I, Shewmaker F. The role of post-translational modifications in the phase transitions of intrinsically disordered proteins. Int J Mol Sci. 2019;20:5501.

83. Maurizio E, Cravello L, Brady L, et al. Conformational role for the C-terminal tail of the intrinsically disordered High Mobility Group a (HMGA) chromatin factors. J Proteome Res. 2011;10:3283–3291.
   •• This paper is one of the first works describing the use of ion mobility mass spectrometry for the study of phosphorylation-induced conformational alterations in IDP.

84. Das RK, Pappu RV. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. Proc Natl Acad Sci U S A. 2013;110:13392–13397.

85. Beveridge R, Migas LG, Das RK, et al. Ion mobility mass spectrometry uncovers the impact of the patterning of oppositely charged residues on the conformational distributions of intrinsically disordered proteins. J Am Chem Soc. 2019;141:4908–4918.

86. Kohl B, Zhong X, Herrmann C, et al. Phosphorylation orchestrates the structural ensemble of the intrinsically disordered protein HMGA1a and modulates its DNA binding to the NFκB promoter. Nucleic Acids Res. 2019;47:11906–11920.

87. Zhang Q, Wang Y. High mobility group proteins and their post-translational modifications. Biochim Biophys Acta Proteins Proteom. 2008;1784:1159–1166.

88. Schrader T, Bitan G, Klärner F-G. Molecular tweezers for lysine and arginine - powerful inhibitors of pathologic protein aggregation. Chem Commun (Camb). 2016;52:11318–11334.

89. Goel P, Chakrabarti S, Goel K, et al. Neuronal cell death mechanisms in Alzheimer's disease: an insight. Front Mol Neurosci. 2022;15:937133.

90. Paulo JA, Schweppe DK. Advances in quantitative high-throughput phosphoproteomics with sample multiplexing. Proteomics. 2021;21:e2000140.

91. Sharma K, D'Souza RCJ, Tyanova S, et al. Ultradeep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling. Cell Rep. 2014;8:1583–1594.

92. Palumbo AM, Smith SA, Kalcic CL, et al. Tandem mass spectrometry strategies for phosphoproteome analysis. Mass Spectrom Rev. 2011;30:600–625.

93. Low TY, Mohtar MA, Lee PY, et al. Widening the bottleneck of phosphoproteomics: evolving strategies for phosphopeptide enrichment. Mass Spectrom Rev. 2021;40:309–333.

94. Needham EJ, Parker BL, Burykin T, et al. Illuminating the dark phosphoproteome. Sci Signal. 2019;12:eaau8645.

95. Bekker-Jensen DB, Bernhardt OM, Hogrebe A, et al. Rapid and site-specific deep phosphoproteome profiling by data-independent acquisition without the need for spectral libraries. Nat Commun. 2020;11:787.

96. Urban J. A review on recent trends in the phosphoproteomics workflow. From sample preparation to data analysis. Anal Chim Acta. 2022;1199:338857.

97. Lou R, Cao Y, Li S, et al. Benchmarking commonly used software suites and analysis workflows for DIA proteomics and phosphoproteomics. Nat Commun. 2023;14:94.

98. Beausoleil SA, Jedrychowski M, Schwartz D, et al. Large-scale characterization of HeLa cell nuclear phosphoproteins. Proc Natl Acad Sci U S A. 2004;101:12130–12135.

99. Tsai C-F, Wang Y-T, Chen Y-R, et al. Immobilized metal affinity chromatography revisited: pH/acid control toward high selectivity in phosphoproteomics. J Proteome Res. 2008;7:4058–4069.

100. Jia S, Tang R, Zhang S, et al. Design and fabrication of reusable core-shell composite microspheres based on nanodiamond for selective enrichment of phosphopeptides. Mikrochim Acta. 2022;189:124.

101. Finamore F, Ucciferri N, Signore G, et al. Proteomics pipeline for phosphoenrichment and its application on a human melanoma cell model. Talanta. 2020;220:121381.

102. Thingholm TE, Jensen ON, Robinson PJ, et al. SIMAC (sequential elution from IMAC), a phosphoproteomics strategy for the rapid separation of monophosphorylated from multiply phosphorylated peptides. Mol & Cell Proteomics. 2008;7:661–671.

103. Kinoshita E, Ishikawa K, Kinoshita-Kikuta E, et al. An alkoxide-bridged dinuclear Zinc(II) hexaazacryptate: a novel phosphate capture molecule in aqueous solution. BCSJ. 2005;78:125–131.

104. Kinoshita E, Kinoshita-Kikuta E, Koike T. History of Phos-tag technology for phosphoproteomics. J Proteomics. 2022;252:104432.

105. Hunter T. A journey from phosphotyrosine to phosphohistidine and beyond. Mol Cell. 2022;82:2190–2200.

106. Leitner A. Enrichment strategies in phosphoproteomics. Methods Mol Biol. 2016;1355:105–121.

107. Kalagiri R, Hunter T. The many ways that nature has exploited the unusual structural and chemical properties of phosphohistidine for use in proteins. Biochem J. 2021;478:3575–3596.

108. Makwana MV, Muimo R, Jackson RF. Advances in development of new tools for the study of phosphohistidine. Lab Invest. 2018;98:291–303.

109. Fuhs SR, Meisenhelder J, Aslanian A, et al. Monoclonal 1- and 3-phosphohistidine antibodies: new tools to study histidine phosphorylation. Cell. 2015;162:198–210.

110. Zhao J, Zou L, Li Y, et al. HisPhosSite: a comprehensive database of histidine phosphorylated proteins and sites. J Proteomics. 2021;243:104262.

111. Lee DCH, Jones AR, Hubbard SJ. Computational phosphoproteomics: from identification to localization. Proteomics. 2015;15:950–963.

112. Kauko O, Laajala TD, Jumppanen M, et al. Label-free quantitative phosphoproteomics with novel pairwise abundance normalization reveals synergistic RAS and CIP2A signaling. Sci Rep. 2015;5:13099.

113. Kim HJ, Kim T, Hoffman NJ, et al. PhosR enables processing and functional analysis of phosphoproteomic data. Cell Rep. 2021;34:108771.

114. Beausoleil SA, Villén J, Gerber SA, et al. A probability-based approach for high-throughput protein phosphorylation analysis and site localization. Nat Biotechnol. 2006;24:1285–1292.

115. Savitski MM, Lemeer S, Boesche M, et al. Confident phosphorylation site localization using the mascot delta score. Mol & Cell Proteomics. 2011;10:M110.003830.

116. Xiao D, Chen C, Yang P. Computational systems approach towards phosphoproteomics and their downstream regulation. Proteomics. 2023;23:e2200068.

117. Potel CM, Lemeer S, Heck AJR. Phosphopeptide fragmentation and site localization by mass spectrometry: an update. Anal Chem. 2019;91:126–141.

118. Taus T, Köcher T, Pichler P, et al. Universal and confident phosphorylation site localization using phosphoRS. J Proteome Res. 2011;10:5354–5362.

119. Locard-Paulet M, Bouyssié D, Froment C, et al. Comparing 22 popular phosphoproteomics pipelines for peptide identification and site localization. J Proteome Res. 2020;19:1338–1345.

120. Lermyte F, Valkenborg D, Loo JA, et al. Radical solutions: principles and application of electron-based dissociation in mass spectrometry-based analysis of protein structure. Mass Spectrom Rev. 2018;37:750–771.

121. Mikesh LM, Ueberheide B, Chi A, et al. The utility of ETD mass spectrometry in proteomic analysis. Biochim Biophys Acta. 2006;1764:1811–1822.
    •• **This paper provides a useful overview of ETD, in particular as concern its utility in the process of PTM mapping and in providing PTM data on large peptides.**

122. Boersema PJ, Mohammed S, Heck AJR. Phosphopeptide fragmentation and analysis by mass spectrometry. J Mass Spectrom. 2009;44:861–878.

123. Hart-Smith G, Low JKK, Erce MA, et al. Enhanced methylarginine characterization by post-translational modification-specific targeted data acquisition and electron-transfer dissociation mass spectrometry. J Am Soc Mass Spectrom. 2012;23:1376–1389.

124. Kundinger SR, Bishof I, Dammer EB, et al. Middle-down proteomics reveals dense sites of methylation and phosphorylation in arginine-rich RNA-binding proteins. J Proteome Res. 2020;19:1574–1591.
    • **This paper describes the use of ETD for the mapping of PTMs overcoming technical sequence coverage problems.**

125. Kline JT, Mullen C, Durbin KR, et al. Sequential ion-ion reactions for enhanced gas-phase sequencing of large intact proteins in a tribrid orbitrap mass spectrometer. J Am Soc Mass Spectrom. 2021;32:2334–2345.

126. Duselis EM, Panepinto MC, Syka JEP, et al. Improved sequence analysis of intact proteins by parallel ion parking during electron transfer dissociation. Anal Chem. 2021;93:15728–15735.

127. Jung HR, Sidoli S, Haldbo S, et al. Precision mapping of coexisting modifications in histone H3 tails from embryonic stem cells by ETD-MS/MS. Anal Chem. 2013;85:8232–8239.
    • **This paper is an example of how ETD can be a key technology to solve the problem of combinatorial PTMs.**

128. Peng Z, Mizianty MJ, Xue B, et al. More than just tails: intrinsic disorder in histone proteins. Mol Biosyst. 2012;8:1886–1901.

129. Jenuwein T, Allis CD. Translating the histone code. Science. 2001;293:1074–1080.

130. Abzalimov RR, Kaplan DA, Easterling ML, et al. Protein conformations can be probed in top-down HDX MS experiments utilizing electron transfer dissociation of protein ions without hydrogen scrambling. J Am Soc Mass Spectrom. 2009;20:1514–1517.

131. Zhang H, Cui W, Wen J, et al. Native electrospray and electron-capture dissociation in FTICR mass spectrometry provide top-down sequencing of a protein component in an intact protein assembly. J Am Soc Mass Spectrom. 2010;21:1966–1968.