Puzzling over protein cysteine phosphorylation – assessment of proteomic tools for S-phosphorylation profiling

A. K. Buchowiecka*

Cysteine phosphorylation has recently been discovered in both prokaryotic and eukaryotic systems, and is thought to play crucial roles in signaling and regulation of cellular responses. This article explores the topics of chemical stability of this type of structural modification and the resulting issues regarding affinity enrichment of S-phosphopeptides and their mass spectrometry-based detection in the course of general proteomics studies. Together, this work suggests that the current advances in phosphoproteomic methodologies provide adequate tools for investigating protein cysteine phosphorylation and appear to be immediately available for practical implementation. The article provides useful information necessary for designing experiments in the emerging cysteine phosphoproteomics. The examples of methodological proposals for S-linked phosphorylation detection are included herein in order to stimulate development of new approaches by the phosphoproteomic community.

Introduction

The last decade has witnessed dynamic progress in proteomic studies. A global view of the advanced technologies that drive development of new research methodologies and that enable impressive achievements in the field is given in the recent review articles.1–3 A comprehensive evaluation of the current challenges in large-scale phosphoproteomics is presented by Engholm-Keller and Larsen.4 To date, the majority of phosphoproteomic studies are focused on various aspects of serine, threonine and tyrosine O-phosphorylation that dominate phosphoproteomic studies are focused on various aspects of serine, threonine and tyrosine O-phosphorylation that dominate signaling and biological regulation in eukaryotic organisms. In vivo phosphorylation of side chains of six other amino acids (His, Lys, Arg, Asp, Glu and Cys) is much less explored due to the chemical lability of the respective peptidyl phosphoramidates,5–7 peptidyl acyl phosphates and peptidyl phosphorothiolates.

Cysteine is one of the least abundant amino acids incorporated into proteins. Studies done on single-residue compositions of proteomes revealed an average cysteine frequency of around 2% for eukaryotic organisms, while for Eubacteria and Archea frequency of ca. 1%. It was also shown that on average, approximately 70% of the reduced thiol equivalents within cells are those originating from proteins. It is thought that the abundance of cysteine is governed by its function in proteins and that the high reactivity of thiol groups shapes distribution and topology of cysteine residues in protein structures. Isolated cysteine accounts for 55–60% of all protein cysteine residues, but due to their anomalous hydrophobic behavior, cysteine residues are infrequently found on protein surfaces.10 Exposed protein thiols are potentially accessible for redox reactions10,11,12 and, presumably, only a certain fraction of them can experience a reversible, enzymatic S-phosphorylation.

Cysteine phosphorylation is important in the context of cysteine-dependent protein phosphatases (CDPs) which belong...
to a subfamily of protein tyrosine phosphatases (PTPs) and catalyze the hydrolysis of phosphoester bonds via the formation of a phospho-cysteine intermediate. Unlike the so-called dual specificity phosphatases, which catalyze dephosphorylation of both Tyr and Ser/Thr residues in proteins, CDPs exhibit specificity exclusively for phosphotyrosine-containing substrates. The sulfur atom of the reactive cysteine present in the CDP catalytic center plays the role of the phosphate group acceptor, thus a cysteiny1-phosphate intermediate is formed. In the second step of catalysis, the cysteiny1-phosphate intermediate undergoes nucleophilic attack by a water molecule. Due to the enzyme active site architecture, the phosphate is prevented from being transferred to acceptors other than water. More information on a fundamental role of Cys-dependent PTPs in regulation of cell signaling is accessible in the extensive review by Nicholas K. Tonks. The author presents substantial body of data emphasizing the contribution of phosphatases to signal transduction networks. The work explains that kinases and phosphatases are specific and competing regulators of signaling that are coordinated to determine the response to a physiological stimulus.

In 2012 Sun et al. detected phosphorylation that occurs on a reactive cysteine residue conserved in several global transcriptional regulators that control a broad spectrum of genes in the human pathogen Staphylococcus aureus. The authors have shown that this reversible posttranslational modification (PTM) is critical in regulating expression of virulence determinants and bacterial resistance to antibiotics. Moreover, these studies also suggest that cysteine phosphorylation, previously considered a rare modification, could be more prevalent in nature and may play significant previously unappreciated roles in biological regulation in various organisms. Another group of researchers, reexamining PTMs of recombinantly expressed human cystic fibrosis transmembrane conductance regulator, detected in its sequence S-phosphocysteine. These new findings extend a short list of known S-phosphorylated proteins (RESID ID AA0034) and provoke questions regarding the possibility of concurrent detection of S- and O-phosphorylation in large-scale phosphoproteomic studies. What factors interfere with the identification of S-linked phosphorylation, and which ones are favorable? This article suggests that through current advances in phosphoproteomic methodologies, adequate tools for detecting protein cysteine phosphorylation appear to be accessible and ready for implementation.

**Stability**

O-phosphorylated peptides are quite durable under a broad range of pH values, and their stability can be attributed to the resistance of the phosphate group to hydrolysis or beta-elimination reactions. However, under specific conditions, O-phosphopeptides are susceptible to both of these transformations. In turn, some chemical properties of S-phosphorylated peptides can be inferred from early studies involving several low molecular S-substituted phosphorothioic acids. In general, this class of compounds is very stable at pH >7 up to 12, but is prone to phosphothiolester bond hydrolysis with its maximum rate at a pH of 2–4. At pH <2 hydrolysis slows down according to a mechanism discussed by Åkerfeldt. In another study, it was observed that S-phosphorylated cystamine and S-phosphorylated mercapto propionic acids undergo hydrolysis at a maximal rate at a pH of 3–4, reaching a half-life of about 15 min. at 37 °C. Moreover, thiophilic metal ions such as Ag I and Cu II were efficient promoters of phosphothiolester bond hydrolysis. Until now, only two examples of kinetic studies on hydrolysis of S-phosphorylated peptides have been described. In 1988 Pas and Robillard demonstrated that at pH 4 the KIIAPCDAGMGSSAM peptide loses a phosphate group with a maximum rate comparable to the kinetics of phosphoenolpyruvate hydrolysis (half-life 30 min. at 25 °C). Later Meins et al. applied the same procedure to study chemical dephosphorylation of the peptide ENITNLDApCITR. Both peptides originated from Escherichia coli enzymes involved in carbohydrate transport by the phosphoenolpyruvate-dependent phosphotransferase system.

**Affinity enrichment**

Successfully navigating the experimental steps of phosphopeptide analysis beginning with protein lysate and ending with the identification of phosphopeptides is a technically demanding task. Generally, in a typical large-scale phosphoproteomics complex protein samples are first digested with selected proteinase(s) and are then subjected to an obligatory affinity enrichment. The sample preparation process usually involves lengthy operations that are performed at both low and high pH and may run as long as two days. Most frequently, metal oxide affinity chromatography is used for the specific catch and release of phosphopeptides. The enrichment process occurs by electrostatic interactions between charged peptides and the charged solid support surface. However, the electrostatic-mediated deposition does not exclude phosphate–metal cation covalent interactions. For example, detailed studies on stability of the dodecyl phosphate films on titanium dioxide demonstrated enhanced hydrolysis of phosphoester bonds at an alkaline pH, possibly due to phosphate covalent binding. In general, charge transfer from phosphate oxygen atoms to coordinately unsaturated titanium cations on the metal oxide surface increases the electrophilicity of the phosphorus atom, making it more prone to a nucleophilic attack by a water molecule or by a hydroxide ion.
observed on lanthanide-based affinity materials as exemplified in several studies. The hydrolysis taking place even in a slightly alkaline or slightly acidic milieu can be connected to catalytic activity of lanthanide ions towards phosphate esters. However, in the following MALDI experiments attention should be paid on distinguishing between the MS-signals from hydrolytically dephosphorylated peptides and the identical signals that may originate from a metastable loss of the phosphate group (−80 Da). Moreover, there is also possibility of partial dephosphorylation of phosphopeptides due to their acid hydrolysis with common MALDI matrices.

Together, these findings suggest that ammonia elution performed under conditions that hydrolytically affect O-phosphopeptides, may cause more extensive, if not complete hydrolysis of S-linked phosphopeptides. Thus, one can expect the co-occurrence of phosphopeptides and cysteine-containing peptides in the eluate from materials that do not display affinity for thiol groups, unlike notably Fe III-IMAC resins. Nonetheless, to specify conditions for S-phosphopeptide survival in the recovery step, enrichment experiments with the use of model S-phosphorylated peptides are recommended. Such compounds can be obtained by Michael addition of thiophosphoric acid to dehydroalanine-containing proteins and peptides.

Several articles regarding phosphopeptide enrichment on innovative supports offer fast techniques, which should not be harmful to a phosphothiolester link. Some of these novel affinity materials can work at almost neutral milieu, using nonvolatile buffers for loading and elution processes. This necessitates subsequent removal of salts. Recently, effective methods for desalting phosphopeptides under both neutral and alkaline conditions as well as very fast desalting techniques under acidic conditions have been considerably improved. It ought to be mentioned that unavoidable steps involving low-pH operations, if limited to very short time duty, could also be incorporated into enrichment technology.

**Detection**

This brief summary outlines procedural conditions required to make co-enrichment for O- and S-phosphopeptides possible. Presented below is the rudimentary concept of an MS-based approach for detection of presumably infrequent S-linked phosphopeptide among abundant O-phosphopeptides, which is strongly supported by information garnered from real-world studies (Fig. 1).

It demonstrates selection of currently available research tools to obtain anticipated results, which seem reliable based on certain experimental data that are cited herein.

The mixture of O- and S-phosphopeptides co-enriched from a digest of fully reduced and thoroughly S-protected protein sample should be split into two equal parts and then subjected to separate dephosphorylation procedures (path 1 and path 2 in Fig. 1). One half of the sample is acid-hydrolyzed in order to dephosphorylate cysteine residues and to generate cysteine-containing peptides, which are further isolated on thiol specific resins (path 1). The recovered peptides are subjected to LC-MS/MS analysis and the obtained identifiable spectra are then used for building a reference spectral library (RefLib). Optional but recommended S-tagging of the cysteine-containing peptides can be performed to improve their fragmentation pattern and further sequencing by database searching or de novo sequencing methods can be performed. Alternatively, real MS/MS-spectra of S-tagged and cysteine-containing peptides can be replaced by computationally predicted semi-empirical spectra of respective S-phosphorylated peptides, by analogy to the recently proposed approach of O-phosphopeptide spectra prediction. The reference library being “extracted” from the analyzed sample covers sequences of peptides potentially bearing S-linked phosphorylation. At the final stage of the analysis, this reference library will serve in the detection and identification of intact S-phosphopeptides from the second half of the analyte, by performing spectrum to spectrum matching.

The second part of co-enriched O- and S-phosphopeptides (path 2 in Fig. 1) is treated with an appropriate, non-specific acid phosphatase at a pH of about 7, which causes a loss of phosphate groups by O-phosphopeptides. Enzymatic dephosphorylation of phosphoproteins by cocktails containing acid phosphatases or with acid phosphatases alone has been described before. Subsequently, most likely intact S-phosphorylated peptides are recovered from the mixture by repeating the harshless enrichment procedure. A conceptually similar approach has been implemented in the study of Hu and Tao aiming at isolating low abundant phosphotyrosine-containing peptides from a complex mixture of phosphopeptides. Under highly alkaline conditions the beta-elimination reaction removes the phosphate from phosphoserine and phosphothreonine residues while phosphotyrosine residues remain unaffected by hydrolysis.

After enzymatic depletion of interfering O-phosphopeptides, targeted S-phosphopeptides need to be fractionated prior to tandem mass spectrometry. In typical phosphoproteomic experiments, a positive ion ESI-MS/MS is preceded by reversed phase chromatography with low pH elution, which would be destructive for S-phosphorylated structures. Yet, in regard to their predicted half-life, rapid acidic separation might be acceptable for highly enriched samples. The actual rate of hydrolysis of S-phosphopeptides during their separation on the RP-LC column may actually be much lower than hydrolysis in a homogenous water solution, analogous to the variable context-dependent stability of N-phosphohistidine containing peptides. The latter undergo rapid dephosphorylation at pH <5 (half-life ca. 20 min) but are stable when stored on an acidic C18 RP column can thus be detected by ESI-MS. An alternative 2D-LC system involving reversed phase fractionation at alkaline pH followed by acidic reversed phase HPLC with a short run time appears as a more suitable approach. S-phosphopeptides that survive the LC-ESI step have a chance to be selected for fragmentation in data dependent acquisition (DDA) mode. However data independent acquisition (DIA), while delivering mixture spectra of co-fragmented peptides, seems to be a method of choice providing higher dynamic range of signal detection. Regardless of the MS/MS acquisition method, further identification of intact S-phosphorylated peptide spectra...
Minireview

**Final remarks and conclusions**

The conceptual discussion outlined above is based heavily on previous empirical accomplishments. A methodology similar to the one presented herein was applied for the MS-based detection and localization of the S-glycosylation sites in a model mixture of chemically modified hen egg white lysozyme.

Briefly, one part of the lysozyme derivatives that contained on average three S-glucosyl residues indeterminately distributed over eight S-protected cysteine residues per molecule was subjected to S-glycosidic bond hydrolysis in a mild alkaline milieu. This reaction transformed S-glycosylated proteins into respective proteins containing freed thiol groups. After S-tagging and trypsinolysis the resultant peptides were analyzed by complementary MS/MS techniques. The spectra of S-tagged sequences incorporating up to three modified cysteine residues per single peptide were selected for constructing the reference spectral library assigned as “a vicarious library”. The second part of S-glycosylated lysozyme derivatives after trypsinolysis followed by the same tandem MS analyses provided spectra of respective S-glycosylated peptides. The complexity of each reference spectrum was reduced by careful selection of appropriate productions/tag loss ions. This approach allowed detection of the respective spectral pairs of S-tagged/S-glycosylated peptide (of the same sequence) as well as further determination of a glycan size. In the light of the unambiguously confirmed occurrence of S-glycosylated bacteriocins the designed and experimentally verified methodology can be eventually tested on native proteomes. It is worthwhile to mention that thoughtful attention to possible co-detection of mono-ADP-S-ribosylation and S-phosphorylation was also given in those studies.

In conclusion, detection of S-linked phosphopeptides by MS-based phosphoproteomics can be successful only if appropriate experimental conditions are used. In the course of routine workflows certain pitfalls jeopardizing the survival of peptidyl phosphorothiolates are to be expected. Prolonged acidic treatments, ammonia elution from commonly used metal oxide materials, and sample contamination with thiophilic metal ions may cause dephosphorylation of S-phosphopeptides. Thus, the possible appearance of cysteinyl peptides in a sample of enriched phosphopeptides may indicate the presence of their S-phosphorylated precursors in the initial protein lysate.

Recognition of mechanisms that lie behind hydrolytic split of the P-S bonding calls for an analytical approach that is relatively harmless for the labile analyte. Detrimental acidic steps should be limited to a few minutes, or rather replaced by operations performed under pH conditions close to neutral. Standard affinity enrichment materials can be exchanged for new ones that insignificantly affect the density of negative charge on the complex-forming phosphate group, therefore minimizing phosphothiolester bond susceptibility to nucleophilic split in an alkaline environment. The final stage of workflow, MS-spectra acquisition in synergy with bioinformatic
analysis, can be successfully performed by the use of appropriately selected methods currently applied in phosphoproteomics.

Some concerns in relation to reactivity of S-phosphorylated peptides during isolation or detection processes can be experimentally resolved by the use of model peptides that are easily synthesizable.

Acknowledgements

I gratefully thank Katarzyna M. Skrzypczynska from the University of California, San Francisco, for her critical reading of this manuscript and constructive suggestions.

Notes and references

42 B. Yoo and M. D. Pagel, Molecules, 2013, 18, 3894.
72. X. S. Yue and A. B. Hummon, *J. Proteome Res.*, 2013, 12, 4176.