Analysis of Protein Phosphorylation Using Mass Spectrometry

Hsin-Yi Wu; Pao-Chi Liao, PhD

Protein phosphorylation has been known to be a pivotal modification regulating many cellular activities and functions. Except for several conventional techniques, mass spectrometry-based strategies are increasingly considered as vital tools that can be utilized to characterize phosphorylated peptides or proteins. In this article, we summarized currently available mass spectrometry-based techniques for the analysis of phosphorylation. Due to the low abundance of phosphopeptides, enrichment steps such as specific antibodies, immobilized metal affinity chromatography, and specific tags are crucial for their use in detection. Since the non-specific binding of the enrichment techniques are constantly of major concerns, phosphatase treatment, neutral loss scan, or precursor ion scan enable the recognition of the phosphopeptide signals. In addition, quantitative methods including isotope labeling and mass tags are also discussed. Phosphoproteome analysis seems to provide elucidation of signaling networks and global decipherment of cell activities, which require powerful analytical methods for complete and routine identification of the phosphorylation event. Despite that numerous approaches have been exploited, comprehensive analysis of protein phosphorylation remains a challenging task. With the progressively more improvements of instruments and methodologies, we can foresee the implementation of a comprehensive approach for the analysis of phosphorylation states of proteins. (Chang Gung Med J 2008;31:217-27)

Key words: phosphorylation, mass spectrometry, IMAC, phosphoproteome, quantitative analysis

Protein phosphorylation, an essential post-translational modification, affects most cellular activities including signal transduction, gene expression, cell cycle progression and other biological functions. Conventional, radioactive $^{32}$P is introduced into cellular proteins via labeled ATP to trace phosphorylation. Fractionation techniques, such as high-performance liquid chromatography and one or two dimensional gel electrophoreses are utilized subsequently for the detection of the radioactive proteins. The localization of the phosphorylation sites are achieved using Edman degradation. Although this technique has demonstrated its capability of analyzing protein phosphorylation on some cases, it can be laborious, time consuming and Edman degradation may fail if the N-terminus of the peptide is blocked. Mass spectrometry has been regarded as a powerful tool not only for the identification of proteins but also the analysis of post-translational modifications. Its tremendous accuracy of mass measurement suggests
the presence of protein modifications which may result in specific mass differences between modified and unmodified proteins. A number of researchers have featured the use of mass spectrometry in protein phosphorylation analysis. Nevertheless, some intrinsic characters of phosphopeptides have hindered the analysis by mass spectrometry. First, the stoichiometric level of phosphoprotein may be very low. Second, phosphopeptides tend to have relatively low ion abundance and suppression effect especially in the presence of non-phosphorylated peptides. Third, phosphate groups on phosphoserine and phosphothreonine are labile, which could decompose during improper sample preparation or peptide fragmentation. Many efforts have been made to improve the analysis of protein phosphorylation, as shown in Fig. 1 which will be discussed in this report.

**Enrichment methods**

**Antibody**

Antibodies are frequently used to recognize specific proteins. For phosphoproteins, a more generally useful tool would be amino-specific antibodies being raised against phosphorylated proteins/peptides. Those site-specific antibodies can be utilized as prior enrichment (e.g. immunoprecipitation) of phosphorylated species from complex samples. There are several commercially available antibodies that are able to bind to phosphotyrosine, phosphoserine, and phosphothreonine residues. However, phosphotyrosine-specific antibodies are considered a relatively efficient way for analyzing tyrosine phosphorylation and have demonstrated several successes in characterizing tyrosine phosphorylation events. The lack of excellent antibodies have been known to be the limitation of analyzing serine- and threonine-phosphorylated proteins. Yet in 2002, Grønborg and his coworkers worked out a new set of antibodies directed against phosphoserine and phosphothreonine residues. These antibodies demonstrated the capabilities of enriching phosphoserine- and phosphothreonine-containing proteins using a global approach.

**Chromatographic and affinity tag enrichment**

As mentioned before, phosphopeptides are often of low abundance and their detection can be interfered with by the existence of non-phosphorylated peptides. Therefore, the complicated peptide mixture samples call for procedures for extraction and enrichment of phosphopeptides before further MS analysis. The most widely used method for selectivity is immobilized metal affinity chromatography (IMAC). This technique incorporates metal ions, usually Fe³⁺, Ga⁺⁺⁺ or others (e.g. Zr⁴⁺), and bind them to a chelating group. Phosphopeptides are bound due to the affinity between the metal ions and phosphate groups. Even though the IMAC technique can be used successfully during both on-line and off-line mass spectrometry analysis, there are several limitations. Phosphopeptides could be lost due to their inability to bind to the IMAC column. In addition, multiple phosphorylated peptides are more enriched and may have difficulty in elution. A major limitation may also be the non-specific binding of nonphosphorylated peptides such as acidic peptides (rich in glutamic and aspartic acid). The conversion of peptides to their corresponding methyl esters proposed by Ficarro et al. not only significantly increased IMAC selectivity for phosphopeptides but also illustrated the potential of IMAC-based enrichment for global phosphoproteome analysis.

Another highly specific phosphopeptide isolation was demonstrated by Pinkse et al. using on-line titanium dioxide (TiO₂) columns. The approach was based on the selective interaction of phosphates with porous titanium dioxide microsphere via biden-
tate binding at the TiO₂ surface. Larsen et al. subsequently modified the method by packing the TiO₂ beads in GELoader tips as TiO₂ microcolumns and peptide loading in 2,5-dihydroxybenzoic acid (DHB), which showed less non-specific binding than IMAC. Titanium dioxide has been successfully used for mapping phosphorylation sites at low femt mole levels in combination with multiple protein cleavage using different protease and sensitive nanoLC-MS/MS. In addition, the use of magnetic Fe₃O₄/TiO₂ core/shell nanoparticles as affinity probes to concentrate phosphopeptides has been reported. Nevertheless, few phosphoproteome analyses using TiO₂ enrichment have been reported so far, and their accessibility for large scale protein phosphorylation analysis still need to be investigated.

Aside from IMAC and TiO₂ enrichment, several materials have also been mentioned to be able to enrich phosphopeptides/phosphoproteins. The use of a metal hydroxide, Al(OH)₃, was proved effective and more selective than commercial phosphoprotein enrichment kits. The utility of zirconium dioxide (ZrO₂) microtips for phosphopeptide isolation prior to mass spectrometric analysis has been demonstrated displaying similar overall performance as TiO₂ microtips.

Chemical modification methods

Chemical replacements of the phosphate group by an affinity group or tags that can be recognized using mass spectrometry analysis are alternative strategies of phosphopeptide/phosphoprotein enrichment. Both phosphoserine- and phosphothreonine containing peptides can lose H₃PO₄ by β-elimination reaction under high pH, resulting in dehydroalanine and dehydroaminobutyric acid residues. Ethanedithiol (EDT), acts as a nucleophile, reacts to double bond, and provides a new reactive thiol group. Subsequent reaction of the free thiol group with different biotin derivatives enables the isolation of labeled peptides by an avidin chromatography (Fig. 2A). To avoid side reactions, the thiol groups on the cysteine must be blocked prior to the reaction. Goshe et al. proposed a modified strategy coupled with labeled EDT (H3 or D4-ethanedithiol), called phosphoprotein isotope-coded affinity tags (PhIAT), for not only the isolation but also the quantitation of phosphopeptides. Goshe et al. even expanded its application for enrichment and identification of low abundance phosphoproteins by integrating capillary reversed-phase liquid chromatography for separating the recovered peptides. These chemical modification strategies encounter a main constraint that they are not suitable for tyrosine phosphorylation since tyrosine phosphorylated residues seldom undergo β-elimination. However, this procedure is easy to perform and reduced the losses due to the complicated reactions.

Another chemical modification method, reported by Zhou et al., was comprehensively applicable to all phosphopeptides (Fig. 2B). Using a carbodiimide condensation reaction, the cystamine group was attached to the phosphate moiety. Purification of the modified peptides was accomplished by attachment to aodoacetyl resin and released by treatment with trifluoroacetic acid. To avoid unwanted reac-

![Fig. 2](image-url) Schemes for enrichment approach using chemical modification. (A) Oda et al. utilized ethanedithiol to provide a thiol group, allowing subsequent reaction with an affinity biotin tag. (B) Zhou et al. couple thiol tags to phosphate groups by a carbodiimide activation. Tagged peptides are bound to iodoacetamide-functionalized beads, released by TFA.
tion, the amino groups and the carboxyl groups of the peptides were blocked with tert-butyl oxycarbonyl (tBoc) chemistry and amidation, respectively. This approach required multiple chemical reaction and purification steps before mass spectrometry analysis, which could introduce great losses.

Phosphoproteins can be enriched by specific anti-phospho antibodies, which greatly reduce the complexity of proteome. Anti-phosphotyrosine antibodies are currently available that can efficiently immunoprecipitate tyrosine phosphorylated proteins. However, anti-phosphoserine or anti-phosphothreonine antibody does not give the similar efficiency, which confines the studies of the complete phosphoproteome using this method. In addition, the enriched proteins are still too complicated for phosphorylation analysis after enzymatic digestion. It can be ameliorated by coupling with a chromatographic enrichment. IMAC or TiO$_2$ has become a common tool for phosphopeptide enrichment. No experimental evidence has shown inconsistencies in binding ability between serine-, treonine-, and tyrosine-phosphorylated species to those chromatographic approaches, which can be used for large-scale and comprehensive phosphoproteome analysis. The greatest limitation of the technique has been the non-specific binding of acidic peptides. Even though several chemical modifications enrichment strategies have been proposed to enhance the selectivity, they require multiple steps for chemical reactions and the reaction conditions must be well-controlled to get better yields. The existing methods require further improvements or refinements and should be combined to augment purity of phosphopeptides.

**Recognition of phosphopeptides**

*Phosphatase treatment*

Phosphatase can remove phosphate group(s) from phosphopeptides, yielding a “mass shift” relative to the original mass of the phosphopeptide. Depending on the number of phosphate groups, the mass shift could be -80 Da (HPO$_3$ = 79.966) or its multiples (–80 $\times$ n Da, where n is the number of phosphate groups). Phosphopeptide identification or sequencing can be achieved during further experiments by focusing on the signals that exhibit mass shifts after phosphatase treatment. Liao et al. applied the combination of phosphatase treatment and MS-based identification. This approach was applied to analyze phosphopeptides using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) before and after being treated with a phosphatase to differentiate phosphopeptide signals from others. Several researchers also have demonstrated that phosphopeptides enriched by IMAC or fractionated using reverse-phase High performance liquid chromatography (HPLC) can be identified by observing the mass shift in the MALDI spectrum following phosphatase treatment.

In 2005, Torres et al. developed a strategy named Phosphatase-directed Phosphorylation-site Determination (PPD). Because dephosphorylated peptides are known to be more detectable in the MS, IMAC-enriched peptides were first treated with phosphatase to yield dephosphorylated peptide signals. Their sequences were consequently defined by MS/MS analysis and the total number of Ser, Thr, or Tyr residues could hypothetically be used to predict the location of phosphopeptide signals. On the basis of this information, a mass list was used to direct MALDI-MS/MS on the phosphorylated peptides bound to IMAC beads for phosphorylation site determination. Although this approach was powerful for phosphoprotein analysis, it is not applicable for comprehensive analysis.

During high-throughput phosphopeptide identification, database search tools may categorize large numbers of false-positive/false-negative phosphopeptide assignments due to their inefficient fragmentation. A scheme containing a dephosphorylation reaction was used to increase the reliability of phosphopeptide identification results in a comprehensive study by choosing the peptide sequencing results with mass shifts and close retention times before and after phosphatase treatment.

Recently, we proposed a strategy to mine phosphopeptide signals by observing mass shifts generated from dephosphorylation reaction in liquid chromatography-mass spectrometry data (Fig. 4). LC-MS analysis was performed on TiO$_2$-enriched peptides before and after phosphatase treatment. Real phosphopeptide signals were expected to emerge mass shifts between the two LC-MS data set due to loss of phosphate moiety. Since LC-MS analysis may generate tremendous amount of signals, manual interpretation of these spectra may be very tedious and inefficient. We programmed an in-house com-
puter program, DeltaFinder, which is used to process data and differentiate possible phosphopeptide signals. The potential signals were sorted out of the complicated LC-MS data set. As shown in Fig. 5A, peptide mixtures of α- and β-caseins revealed that 61 peptide signals consisting of phosphopeptides along with many nonphosphorylated peptides were observed after TiO₂ enrichment. Followed by an alkaline phosphatase treatment, 93 peptide masses were present in the LC-MS data (Fig. 5B). Among them, only 10 pairs were referred to as potential phosphopeptides and their counterparts after being processed by the DeltaFinder (Fig. 5C). The retention times and m/z values of these selected LC-MS signals were used to facilitate subsequent LC-MS/MS experiments for phosphorylation site determination. This scheme has showed its capability of identifying more phosphorylation sites, in comparison with conventional data-dependent LC-MS/MS experiments, in the mixture of α- and β-caseins. For analyzing phosphoproteome, our approach also dominated over conventionally-used mass spectrometric analysis sets in the data-dependent mode and defined much more phosphopeptides as well as sites. The results shown in this work also demonstrated the value of computational algorithms for not only processing the data but also assisting in phosphorylation analysis.

Mass-spectrometric detection

Except for the examination of mass shifts, the specific ions derived from the side chains of phosphorylated residues, which are called “reporter ions”, can also be used to confirm the existence of phosphorylation. Therefore, several mass spectrometric techniques have been employed to distinguish those ions.

Precursor ion scan

Phosphate groups carried by peptides have the tendency to lose ions during fragmentation under high pH conditions and give rise to phosphate-derived anions at m/z 79 (PO₃⁻) as a reporter ion. A mass spectrometer operating in the negative ion mode is set to detect this particular signal in the precursor ion scans. In contrast to the neutral loss scans,
Neutral loss scan

In positive ion tandem MS, phosphoserine- and phosphothreonine-containing peptides prefer to undergo β-elimination reaction and give rise to a neutral loss of 98 Da (H₃PO₄) or 80 Da (HPO₃). This phenomenon can be used for the selective detection of phosphopeptides. In the MS/MS spectra, the loss of phosphoric acid (H₃PO₄) converts phosphoserine and phosphothreonine residues into dehydroalanine (69 Da) and dehydroaminobutyric residues (83 Da) respectively, pointing out the exact location of phosphoserine and phosphothreonine residues. In contrast, relatively few phosphoric acid neutral fragments were observed for phosphotyrosine-containing peptides. The phosphotyrosine-containing peptides are typically stable under these con-
ditions, which hampers the applicability of this method for tyrosine phosphorylation analysis. Take the work of Gruhler et al. as an example, using MS/MS and neutral loss-dependent MS[3]. Only four of 729 phosphorylation sites (0.5%) were identified as tyrosine phosphorylation in a yeast phosphoproteome.[42] Therefore, neutral loss scanning appears to attenuate for the detection of tyrosine phosphorylation.

Fluorescent affinity tags

A novel fluorescent affinity tag (FAT) was synthesized and used to selectively modify phosphorylated serine and threonine residues by beta-elimination and Michael addition. Fluorescence imaging was performed after a solution- or gel-based separation method to visualize these tagged phosphopeptides. The strong fluorescence signal enhanced the detection of phosphoproteins, allowing the subsequent identification.[43] Precursor ion scanning for ions producing a 79 Da fragment is useful because of the sensitive MS detection of the PO₄⁻ anion. The pitfall is the requirement to work in the negative mode, making the use of the data in conjunction with on-line HPLC difficult. Although phosphotyrosine immonium ion is detected at 216, immonium ions for phosphoserines and phosphothreonines are rarely seen because they

Fig. 5 Three dimensional plots of LC-MS data obtained from (A) the tryptic digest of α- and β-casein followed by TiO₂ microcolumn enrichment, and (B) its additional treatment with alkaline phosphatase. In (C), only 10 probable phosphopeptide signals picked by the program are shown. Figure was modified from reference 34.
are prone to lose phosphate groups. On the contrary, scanning for neutral loss of $H_3PO_4$ can hardly detect phosphotyrosine since it is apt to lose $HPO_3$ (80 Da) rather than $H_3PO_4$. Recognizing phosphopeptides by observing mass shifts owing to treatment with phosphatase may introduce less or no selection bias since the dephosphorylation reaction is universal to serine-, threonine-, and tyrosine-phosphorylated peptides/proteins.

Quantification of phosphorylation

The conventional methods used for analyzing stoichiometry of protein phosphorylation are phosphoamino acid analysis or Edman degradation after $^{32}P$ incorporation. It is a tedious procedure and requires handling large amounts of radioactive substances.

Oda et al. used stable isotopes to quantify phosphorylation events by growing two different cell populations in the presence of $^{15}N$-labeled or $^{14}N$-labeled medium. In addition, stable isotope labeling using amino acids in a cell culture (SILAC) was performed by growing cells with different isotopically labeled amino acids such as $[^{13}C_6]$arginine and $[^{13}C_6]$lysine. Using SILAC in combination with LC-MS/MS, Gruhler et al. characterize phosphorylation sites in yeast that are regulated during the mating response.

A strategy using phosphoprotein-specific isotope-coded affinity tags (PhIAT) has been proposed. PhIAT are biotin-containing affinity tags that can be introduced to phosphopeptides by way of the $\beta$-elimination reaction. It enables both purification and quantitation of phosphoserine- and phosphothreonine-containing peptides. Another approach, proposed by Weckwerth et al., incorporates H5- or D5-ethanethiol as a nucleophile. However in their study, no further biotin addition and thus no enrichment was performed.

A stable isotope-tagged amine-reactive reagent, isobaric tag for relative and absolute quantitation (iTRAQ), were design to be isobaric during MS and fragment during MS/MS to reveal differential low mass ions. It consisted of a reporter group (mass = 114-117), a balance group (mass = 31-28), and a peptide reactive group. Four iTRAQ reagents allowed four samples to be compared in a single analysis. Combining labeled peptides from different treatments, relative quantification were accomplished by inspecting the relative responses of reporter groups generated from neutral loss during MS/MS experiments. Zhang et al used iTRAQ to successfully identify the dynamics of tyrosine phosphorylation in response to epidermal growth factor (EGF) in epithelial cells.

Phosphoproteome analysis

It has been estimated that the human genome contains 518 genes for kinases, thus, is one of the largest protein families. In addition, 30% of all cellular proteins may be phosphorylated at any time, indicating that the phosphoproteome of each organism is vast and plays a vital role in regulating cellular activities. There are approximately 100000 potential phosphorylation sites in the human proteome of which fewer than 2000 are currently known. Knowing the phosphoproteome would be a valuable asset in understanding phosphorylation-based signaling networks. The existing mass spectrometry-based methods make the investigation of the phosphoproteome on a global scale possible. We have found many cases in which phosphoproteome analyses have been described. According to those investigations, enrichment of phosphorylated species before MS plays an important role in introducing successes of the characterization of phosphorylation in intricate samples. Among them, IMAC demonstrated much more success than TiO$_2$ microcolumns. Furthermore, the application of TiO$_2$ enrichment for phosphoproteome-wide scale analysis requires more investigation.

Conclusions

Despite the fact a number of analytical strategies have been developed for the characterization of protein phosphorylation, there is no single method that is superior to others for the identification of protein phosphorylation sites. No matter what methods are used, enrichment of phosphorylated proteins and peptides increase the probability of success. With the increasing number of improvements of the mass spectrometry and sample preparation techniques, we can envisage the accessibility of a comprehensive approach for the analysis of protein phosphorylation.

REFERENCES

Protein phosphorylation analysis

2000;100:113-27.
利用質量分析方法研究蛋白質磷酸化

吳欣怡 廖寶琦

蛋白質磷酸化是一種轉譯後修飾，在調節生物體的許多功能方面扮演著重要的角色。對於蛋白質磷酸化的分析，過去有許多方法被提出，其中利用質量分析方法來研究蛋白質的磷酸化越來越受到重視。本文文章描述目前可以被應用的質量分析方法，包括對於少量的磷酸化肽的濃縮，如金屬親合層析管柱 (IMAC)。另外還有在質量分析中辨識磷酸化肽訊號增加鑑定成功率的策略。文中還討論了對於蛋白質磷酸化的定量分析方法，此外還有針對磷酸化蛋白質 (phosphoproteome) 的許多研究報告。目前並沒有一個單一的分析策略可以分析所有的磷酸化蛋白質，但隨著儀器的發展與方法的開發，可以預期往後對於研究蛋白質磷酸化會有更好的進展。(長庚醫誌 2008;31:217-27)

關鍵詞：蛋白質磷酸化，質量儀，金屬親合層析管柱，磷酸化蛋白質體，定量分析