Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Chromatography A, 1053 (2004) 7–14

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

Review

Strategies for shotgun identification of post-translational modifications by mass spectrometry

Greg T. Cantin, John R. Yates III*

Department of Cell Biology, The Scripps Research Institute, 10550 North Torrey Pines Road SR11, La Jolla, CA 92037, USA

Available online 28 July 2004

Abstract

The global identification of post-translationally modified proteins is a difficult challenge that is currently being addressed by many researchers in the field of mass spectrometry (MS)-based proteomics. The ability to identify thousands of proteins by shotgun-based strategies has made the mere idea of a global analysis of a particular protein modification seem reasonable. There has been much progress in the development of methods that make use of shotgun-based protein identification in the analysis of a wide variety of protein modifications, some of which will be discussed here.
© 2004 Elsevier B.V. All rights reserved.

Keywords: Reviews; Proteomics; Proteins

Contents

1. Introduction 7
2. Standard strategies for post-translational modification identification 9
3. Global affinity-based enrichments of post-translational modifications 9
4. Chemical derivatization strategies 11
5. Conclusion 13
Acknowledgements 13
References 13

1. Introduction

Post-translational modifications (PTMs) are of extreme biological importance. Although the mere presence or absence of a particular protein often determines if its designated function is carried out, this is not always the case. In fact, an enormous number of proteins have been shown to be post-translationally regulated by a variety of different modifications. Many of the documented effects of PTMs include a change in enzymatic activity, the ability to interact with other proteins, subcellular localization, targeted degradation, etc. Due to the wide spectrum of effects, identifying and associating PTMs with particular functional consequences of a protein, cell, or organism is a task carried out and pursued by researchers in all areas of the biological sciences. Prior to studying proteins via mass spectrometry (MS), the analysis of modifications occurring on individual proteins was usually a slow and laborious process. Since the advent of mass spectrometry-based protein identification, identifying certain modifications on proteins of interest has become quite routine due to the reduced amount of time and effort required; however, this ease in identification is applicable only to proteins that can be purified in sufficient quantity (hundreds of

* Corresponding author. Tel.: +1 858 784 8862; fax: +1 858 784 8883.
E-mail address: jyates@scripps.edu (J.R. Yates III).

0021-9673/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2004.06.046

nanograms to microgram amounts), thus limiting the scope and throughput of a typical analysis. The next stage in mass spectrometry and biological research is the development of methods to detect PTMs of interest at a global level (e.g. for an entire organism or cell type). To be successful, such methods need to reduce the amount of sample, time, and effort required compared to what is needed in standard mass spectrometry-based analyses of PTMs. This review will discuss recent advances in this very active area of research.

The recent explosion in genomic data has resulted in an increased ability to study hundreds to thousands of proteins simultaneously via mass spectrometry (termed mass spectrometry-based proteomics) in a variety of organisms. Generally, two different mass spectrometry-based proteomic strategies are used for the identification of proteins in complex mixtures: (1) two-dimensional (2D) gel-based procedures; and (2) shotgun-based approaches. 2D gel-based experiments separate intact proteins into individual spots, based on size in one-dimension and isoelectric point in the second dimension. The identification of the proteins corresponding to each gel spot is usually carried out by excising a stained protein spot from a gel, and subjecting the residing protein to an in-gel sequence specific proteolytic digestion (e.g. using trypsin) resulting in characteristic peptides that are subsequently analyzed by mass spectrometry. Either matrix-assisted laser desorption ionization (MALDI) time-of-flight (TOF) is used to identify the protein by its peptide mass fingerprint or nano-electrospray ionization (nanoESI) tandem mass spectrometry (MS–MS) is carried out, where collision-induced dissociation (CID) is used to generate sequence information for multiple peptides from the digest, leading to protein identification. Both approaches depend on computer-based algorithms that can associate mass spectra to a particular protein in a database. Spectra from MALDI-TOF are matched to in silico digests (peptide mass fingerprinting); whereas tandem mass spectra produced from CID are generally associated with a peptide from a particular protein using a program (e.g. SEQUEST) that correlates experimentally collected spectra with theoretical spectra from a protein database [1].

The alternative to standard 2D gel proteomic analysis is shotgun proteomics. Shotgun proteomic methods allow for the identification of hundreds to thousands of proteins within complex protein mixtures without the need for fractionation into individual proteins prior to analysis by mass spectrometry. One such method developed in this laboratory is multidimensional protein identification technology (MudPIT). Starting with a complex protein mixture, MudPIT consists of digesting the mixture and performing online two-dimensional chromatography coupled to a tandem mass spectrometer (LC–LC–MS–MS) [2,3]. This type of analysis allows tandem mass spectra to be collected on up to thousands of peptides from a 12–24 h run using an ion trap mass spectrometer; for an in depth review on ion traps see the following review [4]. Peptides are then identified by using a database searching algorithm such as SEQUEST.

Both 2D gel-based and shotgun proteomic methodologies are capable of identifying many different PTMs. This is done by simply altering the search parameters of either type of algorithm used to include the unmodified and modified masses of a particular amino acid. The two biggest limitations in identifying PTMs in typical analyses are: (1) whether a particular PTM is stable during mass spectrometry and thus identifiable; and (2) whether there is sufficient sequence coverage of a protein that will allow a certain modified region to be accessible to mass spectrometry (discussed below). Another critical point in PTM analyses is whether or not the exact site of the modification is identified. MS–MS-based experiments allows for the site of modification to be localized to a specific amino acid, whereas PTM identification by peptide mass fingerprinting can only localize a modification to a particular peptide. Altogether, mass spectrometry-based proteomics is particularly well suited to study many biologically interesting PTMs at a global level [5].

Early proteomic analyses of PTMs generally utilized 2D-gels, mostly focusing on phosphorylation and glycosylation. These include profiling experiments involving labeling phosphoproteins via the use of ^{32}P-labeled ATP [6], use of enzymatic cleavage of a specific modification and subsequent detection of affected proteins or peptides [6], or, more recently, the use of modification specific dyes [7,8] or antibodies [9]. Identified spots of interest are excised from the original or duplicate gel and subjected to MS-based analysis (either MALDI peptide mass fingerprinting, ESI-MS, or MS–MS peptide sequencing) for identification of the modified protein. Although these 2D gel-based methodologies have proven to be successful in identifying modified proteins, the critical drawbacks include the following: (1) the limits of detection of standard protein staining techniques [10], though improvements are being made by the introduction of fluorescently-based protein staining dyes [11]; (2) exclusion of certain classes of proteins with particular biochemical properties (e.g. integral membrane proteins due to high hydrophobicity); (3) difficulties in localizing the site of modification due to incomplete sequence coverage because of low protein amount in the gel spot; and (4) the time required to carry out the entire analysis at a proteomic level.

As an alternative to the 2D gel approach, there have been a number of recent developments that have incorporated the use of shotgun-based mass spectrometry into a variety of PTM profiling strategies. Shotgun-based approaches can overcome some of the limitations of 2D gel-based analysis. These include reduced biochemical bias [3], faster analysis time, and lower detection limits [12]. However, in some instances, sequence coverage may be higher in 2D gel approaches for some identified proteins. In PTM profiling experiments using a shotgun proteomic approach, the obtained sequence coverage for all of the identified proteins becomes critically important. As the complexity of a protein mixture increases, the sequence coverage obtained from a shotgun proteomic analysis will usually decrease due to: (1) incomplete separation of all peptides during chromatography and/or; (2) sup-

pression of ionization of low abundant or difficult-to-ionize peptides leading to a decreased ability to identify certain peptides by MS–MS analysis. Thus, as the complexity of a protein mixture is increased, the ability to identify all modifications within each protein will generally decrease. This problem is confounded even more in the analysis of some PTMs, phosphorylation for example, where modifications are often substoichiometric and occurring on low abundant proteins. Thus, shotgun-based proteomic methods that are designed to profile PTMs must address and overcome these obstacles in order to be proven successful. The remainder of this review will focus on recent progress made in this area of research.

2. Standard strategies for post-translational modification identification

MudPIT has been successfully used to identify PTMs from purified and moderately complex protein samples [13]. In this study by MacCoss et al. [13], methylation, acetylation, oxidation, and phosphorylation were probed for. Both known and previously undescribed PTMs were identified in commercially available purified proteins, in an immunoprecipitated protein complex from *Sacromyces cerevisciae*, and a moderately complex fraction of proteins from human cataract lens tissue. In this analysis, high sequence coverage was achieved by the separate use of multiple proteases, producing multiple peptide profiles for each protein; thus, increasing the probability of identifying PTMs. Ultimately, this study shows that when adequate amounts of a protein or protein complex can be isolated, a multi-protease analysis via MudPIT is a robust strategy for PTM identification.

The chemical stabilitites of different PTMs vary from extremely labile (e.g. phosphorylated histidine, occurring mostly in prokaryotes) to very stable (e.g. methylated lysine) under conditions typically used in liquid chromatography and mass spectrometry. Thus, not all PTMs are amenable to analysis by mass spectrometry. For instance, phosphotyrosine (pTyr) is quite stable, whereas phosphoserine (pSer) and phosphothreonine (pThr) are more labile during standard analysis by mass spectrometry. It has been shown that pSer and pThr lose phosphoric acid (a process called β-elimination) quite readily in the process of MS–MS analysis, whereas pTyr generally does not [14]. β-Elimination during MS–MS analysis results in an MS–MS spectra that is complicated by the reduction in intensity of expected fragment ions (those generated by peptide bond cleavage) and an introduction of fragment ions that are derived from the β-elimination of the parent and expected fragment ions [15]. This makes the identification of pSer and pThr-containing peptides difficult to identify by standard database searching algorithms. However, the unique fragmentation behavior of pSer and pThr containing peptides during MS–MS has actually been used to aid in their identification. Precursor ion scanning and neutral loss scanning strategies that focus on β-elimination are frequently used to increase the sensitivity of detecting these pSer/Thr-containing phosphopeptides [14]. Though successful in smaller scale studies, most of these strategies have limited sensitivity and are not robust enough for use in global analyses of phosphorylation, and thus will not be covered here.

3. Global affinity-based enrichments of post-translational modifications

If a PTM profiling experiment of a more complex sample (e.g. a whole cell lysate) is desired, some form of enrichment or selective pull down strategy (illustrated in Fig. 1) will most likely be required in order to identify specific peptides and sites that are modified. Many different techniques are being utilized, and a summary of some modifications that are being targeted and the types of strategies used are listed in Table 1. One such strategy has been successfully carried out by Peng et al. [16] in a global analysis of ubiquitination in yeast, a modification mainly involved in regulated protein degradation in eukaryotic cells. An 6× histidine-tagged version of ubiquitin was used to replace the endogenous gene in *S. cerevisciae*. This tag was used to pull down ubiquitin-conjugated proteins from a whole cell lysate via the use of Ni-NTA resin. This fraction was then digested and subjected to offline ion-exchange chromatog-

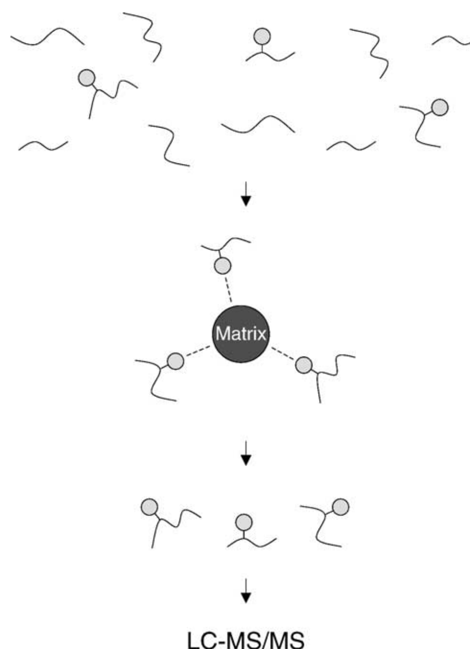

Fig. 1. General scheme for enrichment and identification of PTMs in a complex mixture of proteins. A protein sample (untreated or chemically derivatized at specific modification sites) is digested with one or multiple proteases. Peptides containing the PTM or derivatized-PTM of interest are immobilized on a matrix covalently or noncovalently, and unmodified peptides are washed away. Modified peptides are released from the matrix and subsequently analyzed by LC–MS–MS.

Table 1
Enrichment techniques used in global PTM profiling methods

Targeted PTM	Non-covalent methods	Comments	Chemical derivatization methods	Comments
Phosphorylation	(1) IMAC	(1) Targets pSer/Thr/Tyr	(1) β-Elimination/Michael addition of affinity tag that is or can later be immobilized	(1) Targets only pSer/Thr
	(2) Phosphotyrosine-specific antibodies		(2) Carbodiimide condensation, with subsequent immobilization	(2) Targets pSer/Thr/Tyr
Glycosylation	Lectin affinity	Has been used with specific glycosidase for identification of N-linked glycoproteins	Solid-phase hydrazide chemistry	Has been used with specific glycosidase for identification of N-linked glycoproteins
Ubiquitination	Immunoprecipitation via epitope tagged ubiquitin	Genetically engineered organism/cell line may be required		
Nitrosylation			Nitrosothiols modified w/biotin, subsequently immobilized	

raphy followed by reversed phase LC–MS–MS analysis. From this analysis, 1075 proteins were identified as candidate ubiquitin conjugates. Even though the strategy has been designed to selectively pull down ubiquitinated proteins, the possibility of nonspecific proteins cofractionating does exist and thus these proteins cannot be unequivocally assigned as being ubiquitinated unless the site of ubiquitination is identified, a difficult process since the modification is an M_r 8000 Da protein. To address this issue, Peng et al. [16] showed that modified sites could be determined based on the identification of peptides containing a dipeptide remnant of ubiquitin attached at a lysine (the amino acid target of ubiquitination) after proteolysis with trypsin. Peptides with this unique dipeptide attached are identified and the site of modification localized from MS–MS spectra using SEQUEST. Using this unique modification byproduct of proteolysis, 110 sites from 72 proteins were localized by MS–MS of the original 1075 identified ubiquitin-conjugate candidates. The large discrepancy between the number of proteins with sites localized and the number of candidates is probably due to insufficient sequence coverage of the pulled down target proteins. Localization of the modification for the remaining candidates from this study and of all candidates in future studies is desirable for not only the confirmation that the pulled down protein is in fact modified, but also for the biologically relevant information that could be extracted from the data for use in future studies (see Section 5).

A few monoclonal antibodies have been produced that selectively target a small number of PTMs irrespective of the protein that is modified. The only well-characterized antibodies that have been used in proteomic experiments target phosphorylated tyrosine and nitrosylated tyrosine. Antibodies against nitrosylated tyrosine have mostly been used in Western blotting experiments of 2D gels, where corresponding protein spots are excised and identified from duplicate gels [17]. Antibodies directed against pTyr have also been utilized in 2D gel/western blotting-based proteomic studies [9]; but more interestingly, these antibodies have been used as reagents to enrich for pTyr-containing proteins in solution [18,19]. Enrichment is achievable due to the affinity and specificity of these antibodies for the pTyr epitope in physiological conditions, allowing for immunoprecipitations to be carried out in complicated cell lysates. In one study of two mammalian signal transduction cascades, these antibodies were used to immunoprecipitate pTyr-containing proteins from a protein lysate [18]. Immunoprecipitated proteins were run on a one-dimensional gel to determine differences between stimulated and untreated protein samples. From excised protein bands, nine proteins were identified, and thus determined to undergo tyrosine-phosphorylation in response to a particular stimulus [18]. Though this study is not necessarily high throughput, the use of these antibodies show great promise for application to other signal transduction cascades and also for their use in more streamlined proteomic methods such as an approach used by Salomon et al. [19] that will be discussed later. Antibodies directed against pSer and pThr have not been as successfully produced as those against pTyr. There has been at least one report that has attempted to characterize and utilize antibodies generated to be specific for pSer and pThr-containing proteins [20]. However, in this study, only an antibody that recognizes R-X-pSer/Thr (common in protein kinase A-phosphorylated proteins) was shown to be useful for immunoprecipitation from a cell lysate. In an immunoprecipitation experiment, using this antibody, followed by one-dimensional gel electrophoresis, only seven proteins were identified as candidate pSer/Thr-containing proteins. Due to the nature of the modification, generation of antibodies that universally recognize pSer/Thr-containing proteins may not be achievable. Thus, generation of antibodies that recognize pSer/Thr in particular contexts of adjacent amino acids (as described above for protein kinase A substrates) may be the only recourse. If this is true, a significant amount of work will need to be invested in the generation and characterization of antibodies specific to pSer and pThr in particularly known motifs or contexts in order for antibodies to be successfully used in enrichment strategies. This may not be an attractive

venture, and thus time may be more wisely spent in the further development of other strategies designed for the analysis of serine and threonine phosphorylation events such as those discussed below.

Another type of modification-specific enrichment strategy is the use of immobilized metal affinity chromatography (IMAC). It has been determined that phosphoproteins and phosphopeptides can selectively bind and be eluted from chelated metals such as Fe and Ga quite readily [21,22]. A number of different strategies utilizing IMAC in both on-line and off-line methods have been developed [14,22]. One very successful method was reported by Ficarro et al. [23], where phosphorylation in *S. cerevisiae* was analyzed at a global level. In this strategy, yeast whole cell lysate was digested and phophopeptides were selectively purified by the use of IMAC. IMAC-bound phosphopeptides were eluted onto a reverse phase capillary column that was then hooked up on-line for LC–MS–MS analysis. From the analysis of soluble yeast extract, more than 1000 candidate phosphopeptides were identified by the aid of a search algorithm that uses neutral loss in the identification process. However, only 216 of these were manually validated, localizing the site of phosphorylation. Nevertheless, this is the most extensive list produced thus far on identified sites of phosphorylation in any organism. This procedure is ideal for localizing the phosphorylation site, because the IMAC enrichment step is performed on digested protein. Thus, only phosphopeptides should be enriched, effectively eliminating the sequence coverage problem confronted with when enriching for a particular modification on intact proteins. An apparent advance that was also introduced in this report was the incorporation of a methyl esterification step on the protein digest prior to IMAC. This derivatization step blocks acidic groups, thereby reducing the amount of acidic peptides retained during IMAC separations, a well-documented problem [22], thus increasing the selectivity for phosphopeptides.

It is also possible to utilize multiple enrichment strategies in order to enhance the detection of specific phosphorylation events. In one such instance, IMAC was combined with the use of pTyr specific antibodies to identify pTyr modification events in mammalian cells [19]. In this report, Salomon et al. first enriched for pTyr-containing proteins by immunoprecipitation, then carried out IMAC on digested protein from the immunoprecipitation, followed by LC–MS–MS. Analyzing two different signaling events, activation of T-cells and BCR-ABL-dependent phosphorylation in leukemia cells, Salomon et al. were able to identify 64 unique sites of tyrosine phosphorylation occurring on 32 different proteins. Identifying both known and unknown tyrosine phosphorylations events, this fusion of methods appears to be robust and thus well suited for the analysis of tyrosine phosphorylation occurring in other signal transduction pathways.

Similar to phosphorylation, glycosylation has also been targeted by researches by the use of chromatography selective for the modification. A number of carbohydrate-binding proteins (termed lectins) have been shown to selectively bind glycosylated proteins. Taking advantage of this, a number of researchers have utilized immobilized lectins as a means to profile protein glycosylation in complex protein mixtures [24–26]. One recently published method, termed IGOT, incorporates a lectin-pull down scheme in the analysis of N-linked glycosylation in *Caenorhabditis elegans* [24]. In this report by Kaji et al., immobilized lectin is used to pull down glycosylated proteins, followed by digestion and a second lectin pull down, this time of glycosylated peptides. Bound N-linked glycosylated peptides are then released by the use of an N-linked specific glycosidase (PNGaseF), and the peptide mixture is then analyzed by LC–LC–MS–MS. This specific deglycosylation event results in the conversion of the glycosylated asparagine residues into aspartic acid. This results in a shift in 1 mass unit, which can easily be detected by some mass spectrometers. Additionally, it was shown that this shift could be enhanced by 2 mass units by carrying out the deglycosylation in ^{18}O water, allowing for more reliable identifications of N-linked glycosylation sites. Ultimately, this method was used to identify 250 N-linked glycoproteins containing 400 sites of glycosylation within soluble protein lysate from *C. elegans*. Additionally, with the ability to carry out the deglycosylation in ^{18}O or ^{16}O water, this method appears to be adaptable for use in a quantitative proteomic analysis.

4. Chemical derivatization strategies

A number of PTM-specific proteomic strategies have been recently developed that incorporate reaction steps that allow a specific PTM to be chemically modified. The goal of these chemical derivatizations is usually to enhance the ability to detect the site of modification by mass spectrometry, incorporate a molecular tag that can then be subsequently targeted in some form of enrichment step, or both. By far, the majority of published methods that fall into this category are designed for profiling phosphorylation. A few of these phosphoproteomic methods will be discussed, as well as methods that target glycosylation and nitrosylation.

In an effort to study the labile modification of proteins containing nitrosylated-cysteine (*S*-nitrosylated) at a proteomic level, Jaffrey et al. [27] developed a method that incorporates a derivatization step designed to both stabilize and enrich for these modified proteins. In this method, labile nitrosylcysteine is modified with a biotin moiety that is then used as a tag in a subsequent enrichment step using avidin-conjugated beads, after which eluted proteins are digested and subjected to MALDI-TOF–MS. From a study of mammalian brain lysates, Jaffrey et al. used this method to successfully identify *S*-nitrosylated proteins. Even though no sites were localized in this report and only a few numbers of proteins were identified, the potential of this derivatization/enrichment strategy is quite enormous. As discussed below, this general strategy is increasingly being used to identify many different PTM events at the global level.

Published at the same time as Kaji et al. [24], Zhang et al. [28] reported on an alternative proteomic method to identify N-linked glycosylation that utilized a chemical derivatization procedure. In this method, carbohydrates are chemically modified and immobilized to a resin. Resin-bound glycoproteins are digested, unbound peptides washed away, and N-linked glycopeptides are then specifically released by the use of the glycosidase PNGaseF. Similar to the method of Kaji et al., Zhang et al. identify and localize the site of glycosylation of the enzymatically eluted peptides by searching for a 1 mass unit shift occurring at asparagine residues. Zhang et al. used this method to analyze serum proteins for the presence of N-linked glycosylations. From this analysis, 57 glycoproteins were identified. Additionally, a conclusive enrichment of glycoproteins and consequent de-enrichment of at least one highly abundant non-glycosylated protein was demonstrated. In a second application, 64 N-linked glycoproteins were found in a crude membrane fraction from a cultured mammalian cell line, where each identifying peptide was found to contain the consensus N-linked glycosylation motif (N-X-S/T). Finally, Zhang et al. also included a second derivatization step into the method where peptides are N-terminally labeled with one of two isotopic moieties. This will allow for protein quantification to be incorporated with this method for use in future comparative studies of protein glycosylation.

As previously mentioned, protein phosphorylation is an important modification in cell signaling and is intensely studied in many different areas biology. Due to the high degree of interest, phosphorylation has been the target of numerous proteomic strategies that have been developed within the last few years [14,18]. Amongst these strategies, of which some have already been discussed, include a group of methods that incorporate a chemical derivatization of phosphate that allows for an enhancement in the identification of this PTM.

The majority of derivatization strategies directed towards phosphorylation involve the β-elimination of phosphate from pSer and pThr-containing proteins. Early work led to the finding that high pH treatment of phosphoproteins causes β-elimination of phosphate from pSer and pThr, resulting in dehydroalanine and β-methyldehydroalanaine, respectively [29], similar to the β-elimination reaction occurring in the gas phase during CID [14]. By mass spectrometry, β-eliminated peptides can be identified directly [30] or with the Michael addition of thiol-containing compounds at sites of the β-elimination [31]. Additionally, O-linked (attached through serine or threonine) glycoproteins undergo base-induced β-elimination, and similar methods have been used to analyze these glycosylations [32]. Regarding phosphopeptides, the apparent benefit of this approach is in the enhancement in quality of MS–MS spectra due to an intentional pre-β-elimination of phosphate from pSer and pThr-containing peptides rather than being incomplete and unintentional in a typical analysis; however, this method by itself is not adequate for a global analysis of phophorylation. Though this chemistry has been utilized for a number of years, it has just recently been exploited as a means of enriching for pSer/Thr-containing phosphoproteins in the development of phosphoproteomic methods [6,33–36]. In these methods, β-elimination of pSer/Thr is carried out followed by Michael addition of various sulfhydral molecules which have been used as a tag to allow for enrichment, as a stable replacement of the phosphate moiety for MS–MS analysis, and as a means of quantitation.

Oda et al. [34] and Goshe et al. [35] both published similar methods that derivatize pSer/Thr to a biontinylated-molecular tag that allows for enrichment via the use of immobilized avidin. Both groups clearly achieved an enrichment of derivatized phosphopeptides that are amenable to MS–MS identification. However, the biotin tag used by Oda et al. was apparently more labile than that used by Goshe et al., and gives poorer quality MS–MS fragmentation due to internal fragmentation of the biotin-moiety. Though these strategies were not shown to profile complex protein mixtures, the reports were a promising first step in that direction. In a follow up of the Oda et al. paper, by McLachlin and Chait [6], and in a follow up of the Goshe et al. report, by Qian et al. [33], both groups report on slightly altered derivatization and enrichment strategies. Generally, both groups find alternatives to the use of biotin as a means to derivatize and enrich for phosphopeptides. These reports cite benefits over previous biotin-based methods due to the replacement of labile biotin with a more stable derivative, producing better quality tandem mass spectra [6], and the replacement of biotin–avidin-based enrichment with more efficient purification schemes [6,33]. Two potential drawbacks of these procedures include: (1) the fact that O-linked glycoproteins undergo β-elimination as well, and one may need to eliminate these proteins by affinity chromatography, or remove glycosylations by use of a glycosidase [36]; and (2) cysteines must be protected by oxidation in order to prevent their reactivity during the β-elimination process [37]; which also results in a heterogeneous population of oxidized tryptophan residues, complicating detection of certain peptides [34].

Related to the previous strategies, two methods for pSer/Thr site detection have recently been reorted by Knight et al. [36] and Rusnak et al. [38]. In the method used by Knight et al., β-elimination of pSer/Thr-containing proteins is followed by the covalent addition of cysteamine at the sites of β-elimination. This derivatization results in the conversion of pSer/Thr to a pseudo-lysine which is recognized by lysine-directed proteases. After derivatization and proteolysis, this method allows for the identification of C-terminally labeled peptides by MALDI-MS and ESI-MS–MS. This method has a unique feature over the other methods, and that is the ability to induce protein cleavage at a formerly phosphorylated site. This can be beneficial when the generation of particular phosphopeptides is not readily achievable by commonly used proteases. Also, this procedure has the potential to identify multiple sites of phosphorylation that are clustered in a small stretch of amino acids. This is possible because the derivatized phosphopeptides form stereoisomers where only

one of the two can be digested, thus allowing for a ladder of peptide fragments spanning a region containing multiple sites of phosphorylation to be created. Additionally, Knight et al. demonstrated that cysteamine could be coupled to a solid support through the amine group. In an alternative method, they demonstrated that β-eliminated phosphopeptides can be captured from a peptide mixture, enriched for by washing away unphosphorylated peptides, and subsequently released as cysteamine modified peptides. Though not tested, this method could possibly be used to enrich for phosphopeptides within complex mixtures, as is accomplished in other methods described above.

In a different type of phosporylation derivatization strategy reported by Zhou et al. [39], phosphotyrosine, in addition to pSer/Thr, are targeted. In this method, the phosphate of pSer/Thr/Tyr are modified to allow for their immobilization. After the digestion and enrichment of formerly phosphorylated peptides, phosphate groups are rederived and phosphopeptides are released. At the end of a six-step procedure the collected phosphopeptides are analyzed by LC–MS–MS to identify the peptide and localize the site of phosphorylation. Zhou et al. successfully demonstrated this method on a purified phosphoprotein. The method was also applied to whole cell lysate from yeast resulting in the identification of 24 phosphopeptides from 14 proteins. However, some sites could not be localized to a specific position. This may suggest one of the biggest limitations of the method, which lies in the inherent difficulties in generating quality MS–MS from pSer/Thr-containing phosphopeptides. Nevertheless, this is the longest list of identified phosphorylation sites generated from a proteomic analysis that includes a chemical derivatization step of a phosphate moiety.

5. Conclusion

Identifying post-translational modifications at a global level is undoubtedly one of the next big steps for the field of proteomics. For several of the methods described here, one exciting application is in probing particular cell signaling events. Identifying components and sites of modification that have previously gone unidentified in both well-known and poorly defined signaling cascades is an achievable goal with these methods. These identifications will provide a sound basis for future research; where determining not only which proteins, but what sites are modified will greatly contribute to the enormity of data that is required, for example, in determining kinase–substrate relationships and protein–protein interaction maps, both of which will help in the elucidation of whole signaling networks. Besides basic research, clinically-related research will also benefit from the application of these methods. The comparison of modification events occurring in diseased and healthy tissue may help identify diagnostic or prognostic markers for a particular disease state. Critical to this achievement is the further development and application of the methods discussed above. What also appears necessary for these methods is to not only have the ability to identify PTMs but also quantify and detect statistically significant differences between samples. Though most methods discussed here either included a means of quantification or at least mentioned the possibility in the original reports, the sensitivity and robustness of each remains to be determined when further reports come out, especially those from other laboratories attempting to reproduce and apply these methods to their own research. It will be interesting to see which strategies prevail as a method of choice for future researchers.

Acknowledgements

We thank Cristian Ruse for helpful comments on the manuscript. This work was supported by National Institutes of Health (grants 5 RO1 MH067880 and RR11823-08).

References

[1] J.K. Eng, A.L. McCormack, I. Yates, R. John, J. Am. Soc. Mass Spectrom. 5 (1994) 976.
[2] A.J. Link, J. Eng, D.M. Schieltz, E. Carmack, G.J. Mize, D.R. Morris, B.M. Garvik, J.R. Yates III, Nat. Biotechnol. 17 (1999) 676.
[3] M.P. Washburn, D. Wolters, J.R. Yates III, Nat. Biotechnol. 19 (2001) 242.
[4] K.R. Jonscher, I. Yates, R. John, Anal. Biochem. 244 (1997) 1.
[5] M. Mann, O.N. Jensen, Nat. Biotechnol. 21 (2003) 255.
[6] D.T. McLachlin, B.T. Chait, Anal. Chem. 75 (2003) 6826.
[7] T.H. Steinberg, K. Pretty On Top, K.N. Berggren, C. Kemper, L. Jones, Z. Diwu, R.P. Haugland, W.F. Patton, Proteomics 1 (2001) 841.
[8] T.H. Steinberg, B.J. Agnew, K.R. Gee, W.Y. Leung, T. Goodman, B. Schulenberg, J. Hendrickson, J.M. Beechem, R.P. Haugland, W.F. Patton, Proteomics 3 (2003) 1128.
[9] K. Marcus, D. Immler, J. Sternberger, H.E. Meyer, Electrophoresis 21 (2000) 2622.
[10] S.P. Gygi, G.L. Corthals, Y. Zhang, Y. Rochon, R. Aebersold, Proc. Natl. Acad. Sci. U.S.A. 97 (2000) 9390.
[11] W.F. Patton, J.M. Beechem, Curr. Opin. Chem. Biol. 6 (2002) 63.
[12] S.L. Wu, G. Choudhary, M. Ramstrom, J. Bergquist, W.S. Hancock, J. Proteome Res. 2 (2003) 383.
[13] M.J. MacCoss, W.H. McDonald, A. Saraf, R. Sadygov, J.M. Clark, J.J. Tasto, K.L. Gould, D. Wolters, M. Washburn, A. Weiss, J.I. Clark, J.R. Yates III, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 7900.
[14] M. Mann, S.E. Ong, M. Gronborg, H. Steen, O.N. Jensen, A. Pandey, Trends Biotechnol. 20 (2002) 261.
[15] J.P. DeGnore, J. Qin, J. Am. Soc. Mass Spectrom. 9 (1998) 1175.
[16] J. Peng, D. Schwartz, J.E. Elias, C.C. Thoreen, D. Cheng, G. Marsischky, J. Roelofs, D. Finley, S.P. Gygi, Nat. Biotechnol. 21 (2003) 921.
[17] K.S. Aulak, M. Miyagi, L. Yan, K.A. West, D. Massillon, J.W. Crabb, D.J. Stuehr, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 12056.
[18] A. Pandey, A.V. Podtelejnikov, B. Blagoev, X.R. Bustelo, M. Mann, H.F. Lodish, Proc. Natl. Acad. Sci. U.S.A. 97 (2000) 179.
[19] A.R. Salomon, S.B. Ficarro, L.M. Brill, A. Brinker, Q.T. Phung, C. Ericson, K. Sauer, A. Brock, D.M. Horn, P.G. Schultz, E.C. Peters, Proc. Natl. Acad. Sci. U.S.A. 100 (2003) 443.
[20] M. Gronborg, T.Z. Kristiansen, A. Stensballe, J.S. Andersen, O. Ohara, M. Mann, O.N. Jensen, A. Pandey, Mol. Cell Proteomics 1 (2002) 517.

[21] G. Muszynska, L. Andersson, J. Porath, Biochemistry 25 (1986) 6850.
[22] M.C. Posewitz, P. Tempst, Anal. Chem. 71 (1999) 2883.
[23] S.B. Ficarro, M.L. McCleland, P.T. Stukenberg, D.J. Burke, M.M. Ross, J. Shabanowitz, D.F. Hunt, F.M. White, Nat. Biotechnol. 20 (2002) 301.
[24] H. Kaji, H. Saito, Y. Yamauchi, T. Shinkawa, M. Taoka, J. Hirabayashi, K. Kasai, N. Takahashi, T. Isobe, Nat. Biotechnol. 21 (2003) 667.
[25] J. Bunkenborg, B.J. Pilch, A.V. Podtelejnikov, J.R. Wisniewski, Proteomics 4 (2004) 454.
[26] L. Xiong, D. Andrews, F. Regnier, J. Proteome Res. 2 (2003) 618.
[27] S.R. Jaffrey, H. Erdjument-Bromage, C.D. Ferris, P. Tempst, S.H. Snyder, Nat. Cell Biol. 3 (2001) 193.
[28] H. Zhang, X.J. Li, D.B. Martin, R. Aebersold, Nat. Biotechnol. 21 (2003) 660.
[29] T. Mega, N. Nakamura, T. Ikenaka, J. Biochem. (Tokyo) 107 (1990) 68.
[30] K.A. Resing, R.S. Johnson, K.A. Walsh, Biochemistry 34 (1995) 9477.
[31] H. Jaffe, H.C. Veeranna, Pant, Biochemistry 37 (1998) 16211.
[32] K.D. Greis, B.K. Hayes, F.I. Comer, M. Kirk, S. Barnes, T.L. Lowary, G.W. Hart, Anal. Biochem. 234 (1996) 38.
[33] W.J. Qian, M.B. Goshe, D.G. Camp 2nd, L.R. Yu, K. Tang, R.D. Smith, Anal. Chem. 75 (2003) 5441.
[34] Y. Oda, T. Nagasu, B.T. Chait, Nat. Biotechnol. 19 (2001) 379.
[35] M.B. Goshe, T.P. Conrads, E.A. Panisko, N.H. Angell, T.D. Veenstra, R.D. Smith, Anal. Chem. 73 (2001) 2578.
[36] Z.A. Knight, B. Schilling, R.H. Row, D.M. Kenski, B.W. Gibson, K.M. Shokat, Nat. Biotechnol. 21 (2003) 1047.
[37] H.E. Meyer, E. Hoffmann-Posorske, L.M. Heilmeyer Jr., Methods Enzymol. 201 (1991) 169.
[38] F. Rusnak, J. Zhou, G. Hathaway, J. Biomol. Tech. 13 (2002) 228.
[39] H. Zhou, J.D. Watts, R. Aebersold, Nat. Biotechnol. 19 (2001) 375.