analytical chemistry

PERSPECTIVE

pubs.acs.org/ac

Perspectives of Comprehensive Phosphoproteome Analysis Using Shotgun Strategy

Fangjun Wang, Chunxia Song, Kai Cheng, Xinning Jiang, Mingliang Ye, and Hanfa Zou*

CAS Key Laboratory of Separation Sciences for Analytical Chemistry, National Chromatographic R&A Center, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China

ABSTRACT: Protein phosphorylation is a ubiquitous post-translational modification that regulates almost all cellular processes. The analysis of protein phosphorylation is challenging due to the high dynamic range and low abundance natures of the analyte. Shotgun based proteomics has emerged as a very useful platform to achieve a comprehensive phosphoproteome analysis in considerable depth. In the past few years, significant breakthroughs on the large scale phosphorylation analysis have been witnessed along with the great development of related technologies. The combination of effective enrichment materials, refined analysis workflows, new type of powerful mass spectrometers, and sophisticated bioinformatic tools greatly boost the performance of comprehensive phosphoproteome analysis. In this Perspective, we briefly reviewed recent technological developments on the enrichment materials, prefractionation workflows, and different mass spectrometry fragmentation modes as well as software tools for phosphoproteome identification and quantification. Then, we described the current challenges and potential directions for the future of comprehensive phosphoproteome analysis. We also provide perspectives on how to further improve the performance of related analysis methods and technologies.

As one of the most important post-translational modifications (PTMs), protein phosphorylation plays crucial roles in diverse biological processes such as signal transduction and cell cycle, and it is believed that more than 30% of the proteins can be phosphorylated.[1] Although protein phosphorylation was discovered over 100 years ago, few phosphorylation sites had been mapped until 20 years ago when various proteome analysis technologies were developed. Clearly, the number of reported studies on protein phosphorylations drastically increased from the 1990s onward (Figure 1). In particular, the shotgun proteomic strategy, based on liquid chromatography coupled with mass spectrometry (LC-MS), opened the door to the comprehensive analysis of the whole organisms' proteome[2−4] and, consequently, to the global phosphoproteome analysis. Since, multiple phosphoproteomic technologies and methods, primarily based on shotgun proteomics, have been reported. The progression of phosphoproteomic was driven by developments in peptide/protein fractionation, phosphopeptide/protein enrichment, MS fragmentation, and more recently phosphopeptide quantification (Figure 2 for progression and key milestones). Currently, thousands of phosphorylation sites can be routinely detected in just one phosphoproteomic experiment. We expect that the development and application of phosphoproteomics will continue to grow. Excellent review papers on protein phosphorylation studies and phosphoproteome analysis have been published recently.[5−8] Therefore, we will mainly focus this Perspective article on the recent developments of methods and technologies for shotgun based comprehensive phosphoproteome analysis. Moreover, remaining challenges and potential solutions to overcome will be also discussed.

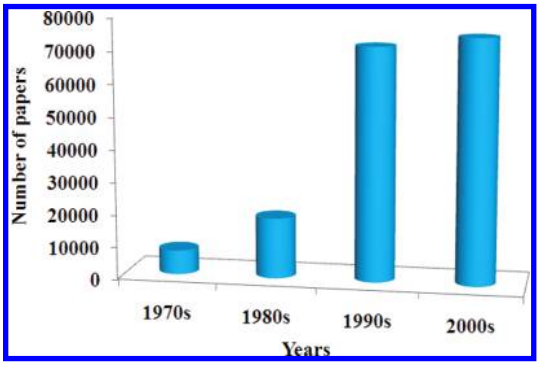

Figure 1. Number of regular publications concerning protein phosphorylation analysis in the last 40 years.

■ DEVELOPMENT OF METHODS AND TECHNOLOGIES FOR SHOTGUN BASED COMPREHENSIVE PHOSPHOPROTEOME ANALYSIS

The common workflow of the shotgun based phosphoproteome analysis is well established and usually contains a number of steps (Figure 3). Each of these steps is critical to obtain accurate results and comprehensive phosphoproteome analysis. In the past decade, the most prominent technological developments of phosphoproteome analysis were focused on the

Received: July 17, 2011
Accepted: September 19, 2011
Published: September 19, 2011

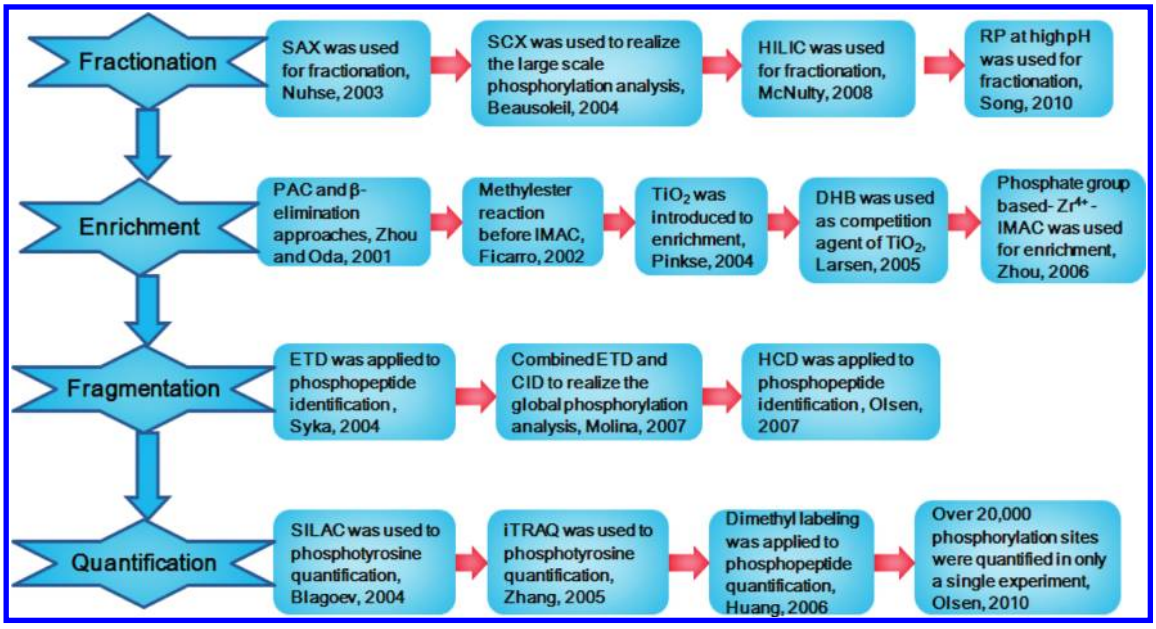

Figure 2. Progression and selected milestones in the development of comprehensive phosphoproteome analysis.

enrichment and fractionation of phosphopeptides, MS detection, large-scale identification, quantification, and bioinformatic analysis. Thus, we briefly discussed each of these aspects below.

PHOSPHOPEPTIDES ENRICHMENT

The high dynamic range and low abundance of both phosphoproteins and phosphopeptides present significant challenges for comprehensive phosphoproteome analysis. Phosphopeptide enrichment is the first and usually the most critical step to effectively decrease the complexity of proteome samples before the MS analysis. Currently, immobilized metal ion affinity chromatography (IMAC) and metal oxide affinity chromatography (MOAC) are most widely used to enrich phosphopeptides from complex matrixes.[9,10]

Unfortunately, nonspecific adsorption of peptides containing acidic residues or histidine was observed when the conventional IMAC adsorbent was used.[11] Instead, new IMAC adsorbents, which incorporate phosphonate as the ligand to immobilize the Ti^{4+} or Zr^{4+} ions, were demonstrated to improve the specificity of phosphopeptides enrichment significantly.[12−14] The mechanism of the phosphonate based-IMAC adsorbents for phosphopeptide enrichment is illustrated in Figure 4. The highly specific interaction of Ti^{4+} or Zr^{4+} ions for phosphate groups from the phosphopeptides leads to fewer nonspecific adsorption of nonphosphopeptides. Besides IMAC materials, the MOAC adsorbents have been also extensively investigated and used for the enrichment of phosphopeptides due to their high chemical stability and rigidity. Among the different types of metal oxides, TiO_2 microspheres exhibited the best performance for enrichment of phosphopeptides.[15]

Similarly to IMAC materials, nonspecific adsorption of acidic peptides onto the metal oxides was observed, which led to considerable efforts to improve their enrichment specificity. For example, 2,5-dihydroxybenzoic (DHB) and glutamic acid were used as competition agents to reduce the nonspecific adsorption of acidic peptides.[16,17] The strong nonspecific adsorption of TiO_2 microspheres caused by Lewis acid interaction is supported by recent observation on its maximum adsorption capacity of ammonia (Lewis base) at 71.6 μmol/g, which is much larger than IMAC adsorbents.[18] The comparison of Ti^{4+} or Zr^{4+}-IMAC adsorbents with TiO_2 and ZrO_2 microspheres were systematically studied. The Ti^{4+} or Zr^{4+}-IMAC adsorbents showed better performance than corresponding metal oxide adsorbents in the enrichment of phosphopeptides.[13] It might be because the hydrophilic polymer bead based Ti^{4+} or Zr^{4+}-IMAC adsorbents have better biological compatibility than metal oxides, and the space arms for the chelating of phosphonate on the IMAC adsorbents provide better flexibility to capture the phosphopeptides in solution.

Besides these conventional micrometer scale adsorbents, nanomaterial and magnetic nanomaterial based adsorbents have also been extensively studied recently.[19,20] These nanomaterial based adsorbents have a relatively large surface area and are expected to improve the enrichment sensitivity and selectivity of phosphopeptides. However, the chemical properties of the nanomaterials and magnetic nanomaterial based adsorbents' surface are usually similar to those of micrometer scale metal oxides adsorbents and have not outperformed conventional adsorbents up to now. Chemical derivatization strategies, such as β-elimination coupled with Michael addition and phosphoramidate chemistry (PAC), can also be applied for phosphopeptides enrichment, but the performance is usually inferior to IMAC and metal oxide.[9] Recently, nanosize polymer based metal ion affinity capture (PolyMAC) was developed, and high reproducibility, selectivity, and recovery for phosphopeptide enrichment were obtained.[21,22]

Phosphotyrosine dependent networks play a key role in transmitting signals, which is important for elucidating the regulatory mechanisms of each biological effect. For tyrosine phosphorylation analysis, antibody based enrichment is still the mainstream method due to its low abundance and the high specificity of antibody for phosphotyrosine.[23,24] Furthermore, the combination of antibody based enrichment with other phosphopeptide enrichment methods also shows a promising strategy for high sensitive tyrosine phosphorylation detection.[22]

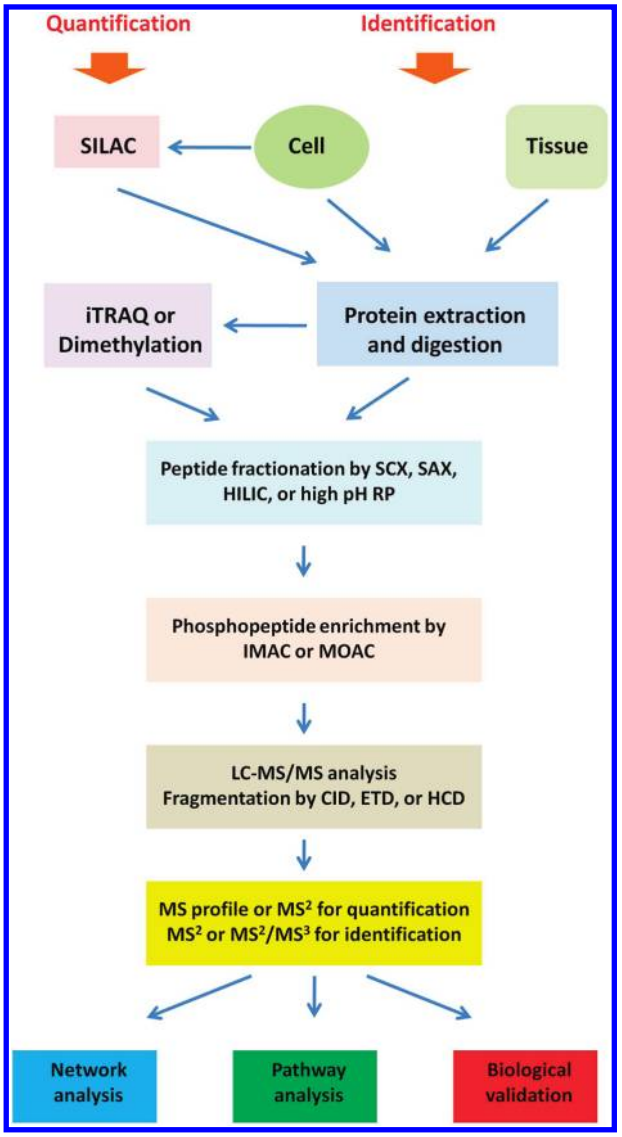

Figure 3. Typical workflow for qualitative and quantitative phosphoproteome analysis based on shotgun strategy.

Another approach to reduce the sample complexity is to use a multidimensional separation before or after phosphopeptide enrichment. The most commonly used multidimensional separation mode is the strong cation exchange combined with reversed phase liquid chromatography (SCX-RPLC).[25] Recent studies indicated that hydrophilic-interaction chromatography (HILIC)-RPLC[26] or RP-RPLC multidimensional separation mode at different pH[27] provided alternative strategies with comparable performance to conventional SCX-RPLC mode. Combinations of these effective multidimensional separation methods with the IMAC or TiO_2 enrichment strategies were demonstrated to be efficient strategies in increasing the phosphoproteome coverage.[28,29]

■ PHOSPHOPEPTIDE IDENTIFICATION AND PHOSPHORYLATION SITE LOCALIZATION

MS with collision induced dissociation (CID) is usually applied for the analysis of phosphopeptides. However, due to the labile phosphoester bond, considerable neutral loss (NL) of

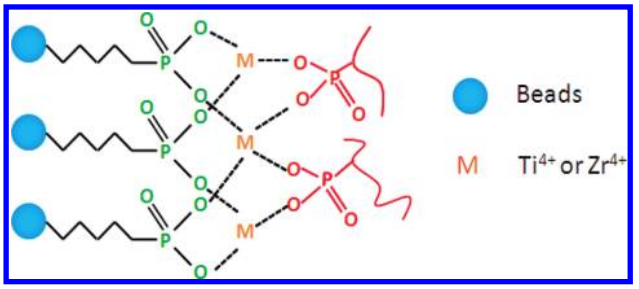

Figure 4. Mechanism of the phosphonate based-IMAC adsorbents for phosphopeptide enrichment.

phosphate group can be observed in the CID process, resulting in uninformative fragment ions, which greatly compromise phosphopeptides identification. However, since most of the phosphopeptides undergo NL of H_3PO_4 in CID, significant NL ions can be exploited as a marker of phosphopeptides.[14,30] To further fragment the phosphopeptide backbone and obtain fragment-rich mass spectra, a data dependent MS^3 strategy has been widely employed.[31,32] In this method, the significant NL ion which has been generated from MS^2 spectrum is isolated and subjected to an additional CID to generate an MS^3 spectrum. Instead of creating a separate MS^3 mass spectra, multistage activation (MSA or pseudo MS^3) is introduced to further fragment the NL ion in parallel during the generating of MS^2 fragments to yield a single spectrum.[33] Both methods showed improved backbone fragment information and, thus, improved the identification of phosphopeptides and assignment of phosphorylation sites. Electron transfer dissociation (ETD) and electron capture dissociation (ECD) are alternative fragmentation modes for the identification of phosphopeptides as the phosphoester bond that is labile under CID conditions remains intact during ECD/ETD, thereby greatly facilitating phosphorylation site assignment.[34] On the other hand, ETD/ECD generates *c*-, *z*-series ions in peptide fragmentations, which complements CID in phosphopeptides identification, as CID generates *b*-, *y*-series ions. Large-scale phosphoproteome analyses have been achieved with good complementary of CID and ETD modes.[35−37] Although ETD/ECD generally yield more confident phosphorylation site assignments, they nevertheless can lead to erroneous annotations,[38] whereas CID seems to be more sensitive and more efficient for phosphopeptide identification.[39] Therefore, especially in cases when limited sample amounts are available, the usage of a common CID based workflow is preferable. Recently, the high energy collision dissociation (HCD) showed good performances for the fragmentation of phosphopeptides, and more informative fragment ions with high mass accuracy were readily obtained.[40] However, the process of the ions transmission to the Orbitrap reduces the MS scan speed and phosphopeptide identification throughput. Furthermore, it has been reported that intramolecular transfer of phosphate may occur during CID fragmentation in ion-trap MS, which may invalidate the site localization in phosphoproteome analysis.[41] However, in some in-depth analyses, there were no evident signs of phosphate relocation in large scale phosphoproteome analysis observed.[42,43]

Because of the insufficient fragmentation of phosphopeptides and difficulties in phosphorylation site localizations, manual assessment of the spectra is commonly adopted to confirm the phosphopeptide identification.[14] Unfortunately, it is very time-consuming to perform manual validation for large-scale phosphoproteome analysis. In order to overcome the potential bias

and time required for manual validation, different bioinformatic tools have been developed. The first type is "automating expert validation" based methods, which simulate the approach for manual interpretation of phosphopeptide MS^2 spectra.[44] The basic concept of these methods is the selection of appropriate features and corresponding criteria. Since most of the phosphopeptides could generate neutral loss during CID fragmentation, the phosphate neutral loss in MS^2 spectrum can be used as a unique feature to confirm the presence of the phosphopeptides. The combination of the assignment information in both MS^2 and MS^3 spectra is another approach to validate the phosphopeptide identification;[45] it was also observed that the phosphoproteome coverage can be improved using the combination of MS^2 and MS^3 spectra instead of just using the MS^2 spectra. Different phosphopeptides have fragment spectra with different characteristics, and therefore, filtering all of the phosphopeptide identifications by setting global filtering criteria may be inappropriate and result in false positives/negatives. Recently, a classification filtering strategy was developed to improve the phosphopeptide identification and phosphorylation site localization. In this strategy, phosphopeptides were divided into four classes according to their different characteristics in fragment spectra, and the identifications from each class of mass spectra were processed and filtered separately using different filtering strategies.[46] Compared with MS^2 strategy and MSA strategy, the classification filtering strategy was demonstrated to provide better performances for the identification of phosphopeptides and for the localization of the phosphorylation to specific sites. Integrated with this classification filtering strategy, a publicly available pipeline containing many useful modules was also developed to facilitate a phosphoproteomic data process.[47] Alternatively, machine learning related algorithms also showed superior performance in discriminating between correctly and incorrectly assigned peptides[48] and can also be used to generate filters for phosphopeptide identifications.[49]

Unambiguous phosphorylation site localization is another challenge for current bioinformatic tools. As many phosphopeptides have multiple potential phosphorylation sites, in order to precisely localize the phosphorylation site to a specific amino acid, it is necessary to introduce another criterion to evaluate the probabilities of different phosphorylation localizations.[50] One of the most popular methods for phosphorylation site localization is the probability based scoring strategy, such as Ascore[51] and PTM-score.[52] Ascore was first developed to localize the phosphorylation site for data obtained from CID fragmentation, and it was modified and extended for ETD/ECD date with various data formats.[53] PTM-score was initially applied to evaluate the matches of MS^3 ions and then was used for the phosphorylation site localization.[54] It has been integrated and implemented in the quantitative platform MSQuant.[55] Even though the detailed methodologies were different, both of the methods shared a similar concept and the localization score was calculated on the basis of the binomial probability between the top 2 most probably localized phosphorylation sites. Other methods were also developed for site localization, such as a combination of MS^2 and MS^3 spectra,[56] using fragments from high level MS spectra (MS^3 or MS^4)[57] or applying Mascot Delta Score (MD score).[58]

A cutoff criterion was usually set for most of the localization score based methods to evaluate the confidence of the phosphorylation site localization. For example, Ascore >19 indicates >99% certainty for correct phosphorylation site localization; however, it was estimated that more than 30% of identified phosphorylation sites are below this threshold because of insufficient site-determining ions.[51] Therefore, improving the accuracy of phosphorylation site location from MS spectra with inferior quality remains the largest challenge for unambiguous site localization. It has been reported recently that high resolution mass spectrometers, such as LTQ-Orbitrap MS with fragmentation techniques, such as ETD/HCD, may provide more opportunity to decrease the false positive identification and locate the phosphorylation sites unambiguously.[34,40] However, corresponding database search methods need to be optimized for the phosphopeptide identification and site localization to fully benefit from these new techniques.

Table 1. Comparison of the Quantification Strategies Applied for the Comparative Phosphoproteome Analysis

		isotope labeling			
	label free	SILAC	iTRAQ/TMT	dimethyl	O^{18}
ability to minimize manual operation error	poor	good	medium	medium	medium
cost effectiveness	good	poor	poor	good	good
throughput for phosphoproteomics	good	good	medium	good	poor
applicability to different type of mass spectrometers	good	medium	poor	medium	medium
capability for analysis of multiple samples	good	medium	good	medium	poor
quantification precision	medium	good	good	good	medium

PHOSPHOPROTEOME QUANTIFICATION

Large-scale phosphoproteome quantification is insufficient to provide relevant information on the biological function of the phosphorylation processes in living organisms. Fortunately, the comparative quantification of phosphoproteome between different physiological or pathological conditions can give some clues to understand the biological functions of site-specific phosphorylations. Thus, the shotgun based quantitative phosphoproteome analysis attracts increasing attention.[7,59,60]

The two mainstream quantification methods are the label-free and the stable isotope labeling approaches. Both of these approaches have been successfully applied for the relative or absolute phosphoproteome quantification, and the strength and weakness of these quantification methods are listed in Table 1. The label-free strategy compares the signal intensity among different samples directly. It is a simple and efficient way to realize large-scale proteome quantification. However, variations in the signals among different runs produced by LC separation and MS detection can compromise the quantification accuracy. Moreover, the parallel phosphopeptide enrichment procedures may also introduce differences among different samples. Although acceptable quantitative reproducibility in label free phosphoproteomics can be obtained by optimizing different sample enrichment performance,[61] it is better to combine the enrichment procedures for different samples to eliminate the unexpected variations. Therefore, stable isotope labeling has emerged as the ideal approach for relative phosphoproteome quantification because different labeled samples can be mixed prior to the enrichment procedure. The most popular strategy for metabolic labeling in proteomics is the stable isotope labeling of amino acids in cell culture

(SILAC), which incorporates isotopes during cell culturing at protein level. Recently, SILAC was extended from cell culture to mammalian tissues for the large scale comparative phosphoproteome analysis.[29,62] In contrast, chemical isotope labeling strategies incorporate the isotope labeling at peptide level by chemical reactions. Although some variations might be introduced by the chemical labeling processes, these strategies still attract significant attention due to the ease of operation and good compatibility with any biologically derived sample. Not all of the proteomic labeling strategies employed with protein are performing as well for the study of phosphoproteins. For example, the statistical analysis of the comprehensive phosphoproteome data sets reveals that SILAC strategy appears to provide better results whereas the iTRAQ strategy is inferior to the other approaches.[29,50,63] This might be attributed to the decreased identification efficiency of phosphopeptides by the iTRAQ strategy as undesirable charge-enhancement is induced by the isobaric tags.[64] Alternatively, more economical chemical isotope reagents can be feasibly used for large-scale quantitative phosphoproteomic analysis such as an isotope dimethyl labeling approach, especially for tissue samples.[65,66]

One of the challenges in phosphoproteomics is to identify the same set of phosphopeptides across multiple samples. It is difficult to ensure that the same phosphopeptides are fragmented in different complex samples due to the somewhat random sampling in data dependent acquisition (DDA) mode. Instead, multiple reaction monitoring (MRM) or selected reaction monitoring (SRM) can be used to quantify a specific set of phosphopeptides across multiple samples due to its excellent selectivity and sensitivity.[67] Already, the combination of DDA identification with the MRM quantification to improve the phosphoproteome quantification reproducibility and sensitivity has been demonstrated.[68] However, some issues should be addressed to integrate the DDA identification results with MRM experiment design.[67] For example, the transition pair, which represents the selected specific pair of m/z values associated with the precursor and fragment ions, might be specific to the instrument used for DDA (for example, linear ion trap mass spectrometer). The same transition pair of ions might not be observed in the triple quadrupole mass spectrometer used for MRM due to the different fragmentation conditions.

PERSPECTIVES FOR THE FUTURE COMPREHENSIVE PHOSPHOPROTEOME ANALYSIS

As described above, the important progresses on enrichment adsorbents, refined workflows, and data processing software have made possible large-scale qualitative and quantitative phosphoproteome analyses. However, we are still far from performing exhaustive studies of the phosphoproteome. Here, we provide some perspectives on how to further improve the performance of these methods and technologies and their applications in comprehensive phosphoproteome.

ENRICHMENT SPECIFICITY AND RECOVERY

Single adsorbent cannot produce complete recovery in phosphopeptide enrichment due to the wide range of physical and chemical properties of different phosphopeptides.[69] Therefore, the combination of multiple enrichment techniques is a promising strategy to increase the phosphoproteome coverage. Alternatively, refining the enrichment workflow, such as selecting more appropriate sample loading and washing buffer or adding some competition agents, can also improve the enrichment specificity.[70,71] Thus, it is necessary to optimize the enrichment workflow in detail and standardize enrichment protocols for different adsorbents across laboratories.[8] Finally, it is still necessary to develop new adsorbents with different chemical properties to achieve even higher specificity and selectivity for the enrichment of phosphopeptides. Besides enrichment specificity, the phosphopeptide recovery produced by the enrichment adsorbents is of equal importance in achieving a comprehensive phosphoproteome analysis. However, the enrichment recovery is often overlooked due to the difficulty of evaluating from complex samples. Moreover, improving the enrichment specificity and the recovery is sometimes contradictory, as some buffers that can improve the enrichment specificity will also decrease the phosphopeptide recovery. It is clear that poor recovery affects the ability to identify low abundant phosphopeptides and might also affect the absolute quantification of specific phosphorylation sites. Therefore, more attention needs to be focused on the enrichment recovery.

DETECTION SENSITIVITY AND REPRODUCIBILITY

For most of the global phosphoproteome analyses, large amounts of starting samples, sufficient sample prefractionation, and ultralong analysis time are necessary.[28,29] However, the manually operating error and sample loss would be increased along with the tedious procedures, which will decrease the detection sensitivity and reproducibility significantly. Therefore, it is necessary to develop refined strategies to improve the phosphoproteome detection sensitivity and reproducibility, especially for minute samples, for example, clinical samples.

Developing the microminiaturization sample preparation platform is a useful way to improve the phosphopeptide detection sensitivity for limited samples. Packing the enrichment materials into tiny column or tip facilitates the automation of phosphopeptide enrichment and reduces sample loss. The miniaturized phosphoproteome reactor and microfluidic chip can be good choices as phosphopeptide enrichment and separation platforms for limited samples.[72] Furthermore, integrating all of the sample preparation procedures including enzyme digestion, phosphopeptide enrichment, and online multidimensional LC separation into one automated system can effectively eliminate sample loss and unexpected variations, which is extremely important for highly sensitive and reproducible phosphoproteome analysis. Alternatively, it was reported that the nano and monolithic materials exhibited excellent performance on simplifying the enrichment procedures and improving the detection sensitivity.[73,74] Recently, a multiprotease digestion strategy was demonstrated as another effective way to improve the phosphorylation detection sensitivity for individual phosphoproteins.[75] Such a strategy may be possibly applied to the in-depth analysis of complex protein samples. A miniaturized LC-MS system with a refined phosphopeptide enrichment protocol was developed, and 1011 phosphorylation sites could be identified from 10^4 HeLa cells.[76] Finally, changing the MS operation mode can also improve the phosphopeptide detection sensitivity,[77] such as when only the targeted precursor ions are selected for MS^2 detection. The high selectivity of this type of scanning mode can greatly improve the detection sensitivity and reproducibility of specific phosphoprotein.[7,77]

ACCURATE SITE-SPECIFIC QUANTIFICATION AND PHOSPHORYLATION OCCUPANCY

The same protein can be involved in multiple signaling pathways through different phosphorylation sites. Moreover, the level of phosphorylation and the total level of a protein are

not necessarily linked. For example, the phosphorylation level at a specific site might increase while the protein expression level may not change. The conventional methods used to improve the protein quantification accuracy, such as averaging several unique peptides' ratios, are not suitable for site specific phosphoproteome quantification,[78] and more intricate situations exist in the site-specific quantitative phosphoproteomics.[50] For example, the quantification results derived from different phosphopeptides may be attributed to the change of the same phosphorylation site due to the miscleavage during enzymatic digestion. As well, the quantification results of a specific phosphorylation site could be derived from the changes either at the phosphorylation level or the protein expression level. Therefore, it is crucial to properly relate the MS quantification results to the change of specific phosphorylation sites in order to correctly understand the actual status of the phosphorylation level within an organism.

Several factors should be addressed to obtain accurate phosphoproteome quantification. First, high quality mass spectra with good signal-to-noise are indispensable to obtain accurate quantification results.[79] The complexity of spectrum may distort the quantification results, if peptides with similar m/z coelute with the targeted isotope forms. Thus, the use of a new generation of mass spectrometers with high resolution and fast acquisition rate is important to guarantee the high quality of mass spectra. Second, decreasing the sample complexity by efficient enrichment and prefractionation strategies is another critical factor to obtain accurate quantification results, especially for the phosphopeptides whose ionization efficiency would be seriously suppressed by their nonphosphorylated analogues. Third, more sophisticated software tools are required for accurate phosphoproteome quantification, such as normalization of the experimental variables between runs in a label-free approach,[50] accurate discrimination of the target isotope envelope from the coeluted contamination, combination of the quantification results of phosphopeptides that are fractionated in two or more fractions, and so on.[79] Recently, a pseudo triplex stable isotope dimethyl labeling approach was developed to achieve replicated quantification in just one experiment. After control of the RSD criterion of the replicated quantification results, the quantification accuracy and phosphoproteome coverage were increased significantly.[80]

Another important issue is the stoichiometry or phosphorylation site occupancy in comprehensive phosphoproteome analysis.[29] Phosphorylation is just a transient modification on proteins, where only a fraction of a particular protein may be phosphorylated at a specific site at any time. For comprehensive phosphoproteome analysis, the phosphorylation occupancy of each site should be also ideally given. However, there is still a lack of efficient ways to calculate the phosphorylation occupancy using a shotgun phosphoproteome approach. In order to realize phosphorylation occupancy, some points should be addressed: (1) Usually, phosphopeptide and nonphosphopeptide quantifications are performed separately, such as the nonphosphorylated proteins are usually separated by SDS-PAGE followed with in-gel digestion and MS detection and phosphopeptides are fractionated by SCX chromatography followed with IMAC or TiO_2 enrichment and MS detection. These complicated workflows result in serious difficulties in obtaining accurate quantification for both phosphorylated and nonphosphorylated forms. (2) It is better to evaluate the enrichment recovery for individual phosphopeptide, which is crucial to calculate the phosphorylation occupancy accurately as part of the phosphopeptide may be lost during enrichment procedures. (3) Not only phosphorylation but also other PTMs, such acetylation, methylation, glycosylation, and so on, should be analyzed at the same time to evaluate the site occupancy accurately, which is also important for elucidating the biological function of a particular phosphorylation site because many phosphorylation processes are related to other PTMs to perform biological functions. (4) If the phosphopeptides and corresponding nonphosphopeptides are analyzed together for some simple proteins samples, the different ionization efficiency of phosphorylated and nonphosphorylated forms should be evaluated.

SIGNALING NETWORK ANALYSIS

Determination of an individual phosphorylation event is not enough to totally understand the actual state of phosphorylation and related biological functions within an organism. A signal network constructed from a comprehensive phosphoproteome data set is a promising way to systematically understand the roles of phosphorylation on different phosphoproteins.[54,81] It offers a new insight to find valuable information from such an enormous phosphoproteome analysis data set. For example, the network constructed by the phosphoproteins responding to a specific perturbation can clearly reveal the signal pathways related to the stimulation, which may elucidate the mechanism of some extraordinary biological statuses.[82] On the other hand, the incredible intricacy of a phosphorylation network provides more challenges for the researchers in this field. First, precise localization of the phosphorylation sites mentioned above is also critical to construct a credible signaling network;[83] second, the integrity of the comprehensive phosphoproteome data set is also important for a stable signal network with good reproducibility, especially some key regulators should be included. At last, a high quality of the kinase−substrate interaction information is indispensable to improve the accuracy of the signaling network.[50] In the future, a phosphorylation signaling network among different cell cycle phases, different subcellular organelles, and tissues in different disease stages will provide a better understanding of the physiological function of reversible phosphorylation in a biological system.[29,84,85] These temporal dynamic and spatial translocations of phosphorylated signaling cascade are valuable to correctly elucidate the progression mechanism of some diseases, such as cancers, and can give some important clues for early diagnosis and treatment.[86,87]

CONCLUSIONS

There is still a long way to reach the stage where we can perform exhaustive quantitative phosphoproteomics. The further improvement of the specificity and recovery of enrichment materials, the performance of analysis workflows, the resolution and scan speed of the mass spectrometer, and the capability of the bioinformatic tools will push the phosphoproteome analysis into a new era compatible with the requirement of biological and clinical researchers. Importantly, the quality of the comprehensive phosphoproteome analysis data set, such as the precise localization of the phosphorylation sites and accurate site-specific quantification, needs to be emphasized in the next few years rather than only increasing the phosphoproteome coverage. Finally, our ability to understand the function of the phosphorylation site has been outpaced by our ability to identify phosphorylation sites. We are lacking high-throughput functional assays for phosphoproteomics. Therefore, more attention should

be paid to bridging phosphoproteomic discovery with biological functions.

AUTHOR INFORMATION

Corresponding Author

*Phone: +86-411-84379610. Fax: +86-411-84379620. E-mail: hanfazou@dicp.ac.cn.

ACKNOWLEDGMENT

The first two authors contributed equally to this work. The authors greatly appreciate the precious suggestions from Dr. N. Dovichi and Dr. D. Figeys. Financial support is gratefully acknowledged from the National Natural Sciences Foundation of China (20735004), the Creative Research Group Project by NSFC (21021004), the China State Key Basic Research Program Grant (2007CB914102), and the Analytical Method Innovation Program of MOST (2009IM031800, 2010IM030500) to H.F.Z.

REFERENCES

(1) Kalume, D. E.; Molina, H.; Pandey, A. *Curr. Opin. Chem. Biol.* **2003**, *7*, 64–69.
(2) Eng, J. K.; McCormack, A. L.; Yates, J. R. *J. Am. Soc. Mass. Spectrom.* **1994**, *5*, 976–989.
(3) Washburn, M. P.; Wolters, D.; Yates, J. R. *Nat. Biotechnol.* **2001**, *19*, 242–247.
(4) Wolters, D. A.; Washburn, M. P.; Yates, J. R. *Anal. Chem.* **2001**, *73*, 5683–5690.
(5) Paradela, A.; Albar, J. P. *J. Proteome Res.* **2008**, *7*, 1809–1818.
(6) Macek, B.; Mann, M.; Olsen, J. V. *Annu. Rev. Pharmacol. Toxicol.* **2009**, *49*, 199–221.
(7) Nita-Lazar, A.; Saito-Benz, H.; White, F. M. *Proteomics* **2008**, *8*, 4433–4443.
(8) Ozlu, N.; Akten, B.; Timm, W.; Haseley, N.; Steen, H.; Steen, J. A. J. *Wiley Interdiscip. Rev.-Syst. Biol. Med.* **2010**, *2*, 255–276.
(9) Thingholm, T. E.; Jensen, O. N.; Larsen, M. R. *Proteomics* **2009**, *9*, 1451–1468.
(10) Han, G. H.; Ye, M. L.; Zou, H. F. *Analyst* **2008**, *133*, 1128–1138.
(11) Ficarro, S. B.; McCleland, M. L.; Stukenberg, P. T.; Burke, D. J.; Ross, M. M.; Shabanowitz, J.; Hunt, D. F.; White, F. M. *Nat. Biotechnol.* **2002**, *20*, 301–305.
(12) Zhou, H. J.; Xu, S. Y.; Ye, M. L.; Feng, S.; Pan, C.; Jiang, X. G.; Li, X.; Han, G. H.; Fu, Y.; Zou, H. *J. Proteome Res.* **2006**, *5*, 2431–2437.
(13) Zhou, H. J.; Ye, M. L.; Dong, J.; Han, G. H.; Jiang, X. N.; Wu, R. N.; Zou, H. F. *J. Proteome Res.* **2008**, *7*, 3957–3967.
(14) Feng, S.; Ye, M. L.; Zhou, H. J.; Jiang, X. G.; Jiang, X. N.; Zou, H. F.; Gong, B. L. *Mol. Cell. Proteomics* **2007**, *6*, 1656–1665.
(15) Pinkse, M. W. H.; Uitto, P. M.; Hilhorst, M. J.; Ooms, B.; Heck, A. J. R. *Anal. Chem.* **2004**, *76*, 3935–3943.
(16) Larsen, M. R.; Thingholm, T. E.; Jensen, O. N.; Roepstorff, P.; Jorgensen, T. J. D. *Mol. Cell. Proteomics* **2005**, *4*, 873–886.
(17) Wu, J.; Shakey, Q.; Liu, W.; Schuller, A.; Follettie, M. T. *J. Proteome Res.* **2007**, *6*, 4684–4689.
(18) Zhang, Y.; Chen, C.; Qin, H. Q.; Wu, R. A.; Zou, H. F. *Chem. Commun.* **2010**, *46*, 2271–2273.
(19) Zhou, H. J.; Tian, R. J.; Ye, M. L.; Xu, S. Y.; Feng, S.; Pan, C. S.; Jiang, X. G.; Li, X.; Zou, H. F. *Electrophoresis* **2007**, *28*, 2201–2215.
(20) Xu, X. Q.; Deng, C. H.; Gao, M. X.; Yu, W. J.; Yang, P. Y.; Zhang, X. M. *Adv. Mater.* **2006**, *18*, 3289–3293.
(21) Tao, W. A.; Wollscheid, B.; O'Brien, R.; Eng, J. K.; Li, X. J.; Bodenmiller, B.; Watts, J. D.; Hood, L.; Aebersold, R. *Nat. Methods* **2005**, *2*, 591–598.
(22) Iliuk, A. B.; Martin, V. A.; Alicie, B. M.; Geahlen, R. L.; Tao, W. A. *Mol. Cell. Proteomics* **2010**, *9*, 2162–2172.
(23) Boersema, P. J.; Foong, L. Y.; Ding, V. M.; Lemeer, S.; van Breukelen, B.; Philp, R.; Boekhorst, J.; Snel, B.; den Hertog, J.; Choo, A. B.; Heck, A. J. *Mol. Cell. Proteomics* **2010**, *9*, 84–99.
(24) Hinsby, A. M.; Olsen, J. V.; Bennett, K. L.; Mann, M. *Mol. Cell. Proteomics* **2003**, *2*, 29–36.
(25) Villen, J.; Gygi, S. P. *Nat. Protoc.* **2008**, *3*, 1630–1638.
(26) McNulty, D. E.; Annan, R. S. *Mol. Cell. Proteomics* **2008**, *7*, 971–980.
(27) Song, C. X.; Ye, M. L.; Han, G. H.; Jiang, X. N.; Wang, F. J.; Yu, Z. Y.; Chen, R.; Zou, H. F. *Anal. Chem.* **2010**, *82*, 53–56.
(28) Dephoure, N.; Zhou, C.; Villen, J.; Beausoleil, S. A.; Bakalarski, C. E.; Elledge, S. J.; Gygi, S. P. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 10762–10767.
(29) Olsen, J. V.; Vermeulen, M.; Santamaria, A.; Kumar, C.; Miller, M. L.; Jensen, L. J.; Gnad, F.; Cox, J.; Jensen, T. S.; Nigg, E. A.; Brunak, S.; Mann, M. *Sci. Signal* **2010**, *3*, ra3.
(30) Feng, S.; Pan, C. S.; Jiang, X. G.; Xu, S. Y.; Zhou, H. J.; Ye, M. L.; Zou, H. F. *Proteomics* **2007**, *7*, 351–360.
(31) Lee, J.; Xu, Y.; Chen, Y.; Sprung, R.; Kim, S. C.; Xie, S.; Zhao, Y. *Mol. Cell. Proteomics* **2007**, *6*, 669–676.
(32) Beausoleil, S. A.; Jedrychowski, M.; Schwartz, D.; Elias, J. E.; Villen, J.; Li, J. X.; Cohn, M. A.; Cantley, L. C.; Gygi, S. P. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 12130–12135.
(33) Schroeder, M. J.; Shabanowitz, J.; Schwartz, J. C.; Hunt, D. F.; Coon, J. J. *Anal. Chem.* **2004**, *76*, 3590–3598.
(34) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9528–9533.
(35) Molina, H.; Horn, D. M.; Tang, N.; Mathivanan, S.; Pandey, A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 2199–2204.
(36) Swaney, D. L.; Wenger, C. D.; Thomson, J. A.; Coon, J. J. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 995–1000.
(37) Boersema, P. J.; Mohammed, S.; Heck, A. J. *J. Mass Spectrom.* **2009**, *44*, 861–878.
(38) Beck, F.; Lewandrowski, U.; Wiltfang, M.; Feldmann, I.; Geiger, J.; Sickmann, A.; Zahedi, R. P. *Proteomics* **2011**, *11*, 1099–1109.
(39) Stingl, C.; Ihling, C.; Ammerer, G.; Sinz, A.; Mechtler, K. *Biochim. Biophys. Acta* **2006**, *1764*, 1842–1852.
(40) Olsen, J. V.; Macek, B.; Lange, O.; Makarov, A.; Horning, S.; Mann, M. *Nat. Methods* **2007**, *4*, 709–712.
(41) Palumbo, A. M.; Reid, G. E. *Anal. Chem.* **2008**, *80*, 9735–9747.
(42) Mischerikow, N.; Altelaar, A. F.; Navarro, J. D.; Mohammed, S.; Heck, A. J. *Mol. Cell. Proteomics* **2010**, *9*, 2140–2148.
(43) Aguiar, M.; Haas, W.; Beausoleil, S. A.; Rush, J.; Gygi, S. P. *J. Proteome Res.* **2010**, *9*, 3103–3107.
(44) Martin, D. M. A.; Nett, I. R. E.; Vandermoere, F.; Barber, J. D.; Morrice, N. A.; Ferguson, M. A. J. *Bioinformatics* **2010**, *26*, 2153–2159.
(45) Jiang, X. N.; Han, G. H.; Feng, S.; Jiang, X. G.; Ye, M. L.; Yao, X. B.; Zou, H. F. *J. Proteome Res.* **2008**, *7*, 1640–1649.
(46) Jiang, X. N.; Ye, M. L.; Han, G. H.; Dong, X. L.; Zou, H. F. *Anal. Chem.* **2010**, *82*, 6168–6175.
(47) Jiang, X.; Ye, M.; Cheng, K.; Zou, H. *J. Proteome Res.* **2010**, *9*, 2743–2751.
(48) Keller, A.; Nesvizhskii, A. I.; Kolker, E.; Aebersold, R. *Anal. Chem.* **2002**, *74*, 5383–5392.
(49) Lu, B. W.; Ruse, C.; Xu, T.; Park, S. K.; Yates, J. *Anal. Chem.* **2007**, *79*, 1301–1310.
(50) Choudhary, C.; Mann, M. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 427–439.
(51) Beausoleil, S. A.; Villen, J.; Gerber, S. A.; Rush, J.; Gygi, S. P. *Nat. Biotechnol.* **2006**, *24*, 1285–1292.
(52) Olsen, J. V.; Mann, M. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 13417–13422.
(53) Bailey, C. M.; Sweet, S. M. M.; Cunningham, D. L.; Zeller, M.; Heath, J. K.; Cooper, H. J. *J. Proteome Res.* **2009**, *8*, 1965–1971.
(54) Olsen, J. V.; Blagoev, B.; Gnad, F.; Macek, B.; Kumar, C.; Mortensen, P.; Mann, M. *Cell* **2006**, *127*, 635–648.
(55) Mortensen, P.; Gouw, J. W.; Olsen, J. V.; Ong, S. E.; Rigbolt, K. T. G.; Bunkenborg, J.; Cox, J.; Foster, L. J.; Heck, A. J. R.; Blagoev, B.; Andersen, J. S.; Mann, M. *J. Proteome Res.* **2010**, *9*, 393–403.

(56) Wan, Y. H.; Cripps, D.; Thomas, S.; Campbell, P.; Ambulos, N.; Chen, T.; Yang, A. *J. Proteome Res.* **2008**, *7*, 2803–2811.
(57) Ruttenberg, B. E.; Pisitkun, T.; Knepper, M. A.; Hoffert, J. D. *J. Proteome Res.* **2008**, *7*, 3054–3059.
(58) Savitski, M. M.; Lemeer, S.; Boesche, M.; Lang, M.; Mathieson, T.; Bantscheff, M.; Kuster, B. *Mol. Cell. Proteomics* **2011**, *10*, M110.003830.
(59) Palmisano, G.; Thingholm, T. E. *Expert Rev. Proteomics* **2010**, *7*, 439–456.
(60) Schreiber, T. B.; Mausbacher, N.; Breitkopf, S. B.; Grundner-Culemann, K.; Daub, H. *Proteomics* **2008**, *8*, 4416–4432.
(61) Soderblom, E. J.; Philipp, M.; Thompson, J. W.; Caron, M. G.; Moseley, M. A. *Anal. Chem.* **2011**, *83*, 3758–3764.
(62) Kruger, M.; Moser, M.; Ussar, S.; Thievessen, I.; Luber, C. A.; Forner, F.; Schmidt, S.; Zanivan, S.; Fassler, R.; Mann, M. *Cell* **2008**, *134*, 353–364.
(63) Wang, Z.; Gucek, M.; Hart, G. W. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 13793–13798.
(64) Thingholm, T. E.; Palmisano, G.; Kjeldsen, F.; Larsen, M. R. *J. Proteome Res.* **2010**, *9*, 4045–4052.
(65) Boersema, P. J.; Foong, L. Y.; Ding, V. M. Y.; Lemeer, S.; van Breukelen, B.; Philp, R.; Boekhorst, J.; Snel, B.; den Hertog, J.; Choo, A. B. H.; Heck, A. J. R. *Mol. Cell. Proteomics* **2010**, *9*, 84–99.
(66) Wu, C. J.; Chen, Y. W.; Tai, J. H.; Chen, S. H. *J. Proteome Res.* **2011**, *10*, 1088–1097.
(67) Lange, V.; Picotti, P.; Domon, B.; Aebersold, R. *Mol. Syst. Biol.* **2008**, *4*, 222.
(68) Wolf-Yadlin, A.; Hautaniemi, S.; Lauffenburger, D. A.; White, F. M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 5860–5865.
(69) Bodenmiller, B.; Mueller, L. N.; Mueller, M.; Domon, B.; Aebersold, R. *Nat. Methods* **2007**, *4*, 231–237.
(70) Thingholm, T. E.; Jorgensen, T. J. D.; Jensen, O. N.; Larsen, M. R. *Nat. Protoc.* **2006**, *1*, 1929–1935.
(71) Aryal, U. K.; Ross, A. R. S. *Rapid Commun. Mass Spectrom.* **2010**, *24*, 219–231.
(72) Raijmakers, R.; Kraiczek, K.; de Jong, A. P.; Mohammed, S.; Heck, A. J. R. *Anal. Chem.* **2010**, *82*, 824–832.
(73) Qiao, L.; Roussel, C.; Wan, J. J.; Yang, P. Y.; Girault, H. H.; Liu, B. H. *J. Proteome Res.* **2007**, *6*, 4763–4769.
(74) Dong, M. M.; Wu, M. H.; Wang, F. J.; Qin, H. Q.; Han, G. H.; Gong, J.; Wu, R. A.; Ye, M. L.; Liu, Z.; Zou, H. F. *Anal. Chem.* **2010**, *82*, 2907–2915.
(75) Han, G. H.; Ye, M. L.; Jiang, X. N.; Chen, R.; Ren, J.; Xue, Y.; Wang, F. J.; Song, C. X.; Yao, X. B.; Zou, H. F. *Anal. Chem.* **2009**, *81*, 5794–5805.
(76) Masuda, T.; Sugiyama, N.; Tomita, M.; Ishihama, Y. *Anal. Chem.* **2011**, DOI: 10.1021/ac201093g.
(77) Unwin, R. D.; Griffiths, J. R.; Leverentz, M. K.; Grallert, A.; Hagan, I. M.; Whetton, A. D. *Mol. Cell. Proteomics* **2005**, *4*, 1134–1144.
(78) Casado, P.; Cutillas, P. R. *Mol. Cell. Proteomics* **2011**, *10*, M110.003079.
(79) Mallick, P.; Kuster, B. *Nat. Biotechnol.* **2010**, *28*, 695–709.
(80) Song, C.; Wang, F.; Ye, M.; Cheng, K.; Chen, R.; Zhu, J.; Tan, Y.; Wang, H.; Figeys, D.; Zou, H. *Anal. Chem.* **2011**, DOI: 10.1021/ac201299j.
(81) Rikova, K.; Guo, A.; Zeng, Q.; Possemato, A.; Yu, J.; Haack, H.; Nardone, J.; Lee, K.; Reeves, C.; Li, Y.; Hu, Y.; Tan, Z. P.; Stokes, M.; Sullivan, L.; Mitchell, J.; Wetzel, R.; MacNeill, J.; Ren, J. M.; Yuan, J.; Bakalarski, C. E.; Villen, J.; Kornhauser, J. M.; Smith, B.; Li, D.; Zhou, X.; Gygi, S. P.; Gu, T. L.; Polakiewicz, R. D.; Rush, J.; Comb, M. J. *Cell* **2007**, *131*, 1190–1203.
(82) Matsuoka, S.; Ballif, B. A.; Smogorzewska, A.; McDonald, E. R., 3rd; Hurov, K. E.; Luo, J.; Bakalarski, C. E.; Zhao, Z.; Solimini, N.; Lerenthal, Y.; Shiloh, Y.; Gygi, S. P.; Elledge, S. J. *Science* **2007**, *316*, 1160–1166.
(83) Linding, R.; Jensen, L. J.; Ostheimer, G. J.; van Vugt, M. A.; Jorgensen, C.; Miron, I. M.; Diella, F.; Colwill, K.; Taylor, L.; Elder, K.; Metalnikov, P.; Nguyen, V.; Pasculescu, A.; Jin, J.; Park, J. G.; Samson, L. D.; Woodgett, J. R.; Russell, R. B.; Bork, P.; Yaffe, M. B.; Pawson, T. *Cell* **2007**, *129*, 1415–1426.
(84) Huttlin, E. L.; Jedrychowski, M. P.; Elias, J. E.; Goswami, T.; Rad, R.; Beausoleil, S. A.; Villen, J.; Haas, W.; Sowa, M. E.; Gygi, S. P. *Cell* **2010**, *143*, 1174–1189.
(85) Zhou, H. J.; Elisma, F.; Denis, N. J.; Wright, T. G.; Tian, R. J.; Zhou, H.; Hou, W. M.; Zou, H. F.; Figeys, D. *J. Proteome Res.* **2010**, *9*, 1279–1288.
(86) Blume-Jensen, P.; Hunter, T. *Nature* **2001**, *411*, 355–365.
(87) Harsha, H. C.; Pandey, A. *Mol. Oncol.* **2010**, *4*, 482–495.