Stefan Loroch[1]*
Clarissa Dickhut[1]*
René P. Zahedi[1]
Albert Sickmann[1,2]

[1]Leibniz-Institut für Analytische Wissenschaften – ISAS – e.V., Dortmund, Germany
[2]Medizinisches Proteom-Center (MPC), Ruhr-Universität, Bochum, Germany

Received December 31, 2012
Revised February 22, 2013
Accepted March 10, 2013

Review

Phosphoproteomics—More than meets the eye

PTMs enable cells to adapt to internal and external stimuli in the milliseconds to seconds time regime. Protein phosphorylation is probably the most important of these modifications as it affects protein structure and interactions, critically influencing the life cycle of a cell. In the last 15 years, new insights into phosphorylation have been provided by highly sensitive MS-based approaches combined with specific phosphopeptide enrichment strategies. Although so far research has mainly focused on the discovery and characterization of *O*-phosphorylation, this review also briefly outlines the current knowledge about *N*-phosphorylation depicting its ubiquitous relevance. Further, common pitfalls in sample preparation, LC-MS analysis, and subsequent data analysis are discussed as well as issues regarding quality and comparability of studies on protein phosphorylation.

Keywords:
Phosphopeptide enrichment / Phosphoproteomics / Protein phosphorylation
DOI 10.1002/elps.201200710

1 Introduction

Reaction to stimuli is one of life's most important principles. The resulting question is: How in particular is this process mediated? On the molecular level, a stimulus has to be recognized by a receptor and an effector has to react in response. Because regulation at genome and transcriptome level is too slow (within minutes to hours), a third mechanism for regulation is necessary: PTM of proteins. PTMs enable cells to dynamically adapt to internal and external stimuli in the milliseconds to seconds time regime, especially with respect to signal transduction [1]. Basically, PTMs directly affect protein structure and in consequence protein localization, activity, and interaction, thus representing essential regulatory switches in life and death [2].

One of the most important PTMs, especially with regard to signal transduction, is protein phosphorylation. Phosphorylation often serves as a molecular on/off switch and can be divided into four subclasses depending on the involved amino acids: (i) *O*-phosphomonoesters on the hydroxyl amino acids serine (phosphoserine, pSer), threonine (phosphothreonine, pThr), and tyrosine (phosphotyrosine, pTyr), (ii) *N*-phosphoramidates on arginine, lysine, and histidine, (iii) S-phosphothioesters on cysteine, and (iv) phosphoanhydrides on glutamic and aspartic acid, respectively [3]. In addition, in 2010, Kühlberg et al. provided evidence for the existence of proteinogenic phosphohydroxproline [4], a fourth *O*-phosphomonoester, so far only observed by enzymatic conversion of synthetic hydroxyproline-containing peptides [5]. Although phosphorylation of proteins occurs via the transfer of a phosphate group from adenosine triphosphate or guanosine triphosphate to the target protein by protein kinases, dephosphorylation is catalyzed by phosphatases releasing inorganic phosphate [6, 7]. Phosphorylation is estimated to be the most abundant PTM, with 3% of human genes encoding for proteins with kinase or phosphatase activity [8]. In addition, according to the phosphosite database (www.phosphosite.org) already more than 170 000 phosphorylation sites have been identified—mainly in *Homo sapiens*, where, for instance, 38 000 phosphosites were detected in HeLa cells. Despite this enormous amount of information, this number is constantly increasing due to the advancements in MS and phosphoproteomics. But although the quality of phosphorylation site annotation has improved, annotated phosphosites should not be considered to be highly confident without a closer look into the respective literature, as automatized database searches are susceptible not only to incorrect phosphosite localization but also to false peptide sequences. Compared to this vast, yet merely qualitative

Correspondence: Dr. Albert Sickmann, Leibniz-Institut für Analytische Wissenschaften – ISAS – e.V., Otto-Hahn-Str. 6b, D-44227 Dortmund, Germany
E-mail: albert.sickmann@isas.de
Fax: +49-231-1392-00

Abbreviations: **AC**, affinity chromatography; **ERLIC**, electrostatic repulsion hydrophilic interaction chromatography; **ETD**, electron transfer dissociation; **HCD**, higher energy collision dissociation; **IMAC**, immobilized metal ion affinity chromatography; **MD score**, Mascot Delta Score; **MOAC**, metal oxide affinity chromatography; **MRM**, multiple reaction monitoring; **pHis**, phosphohistidine; **pSer**, phosphoserine; **pThr**, phosphothreonine; **pTyr**, phosphotyrosine; **SCX**, strong cation exchange; **SIMAC**, sequential elution of immobilized metal ion affinity chromatography

*These authors contributed equally.

Colour Online: See the article online to view Fig. 3 in colour.

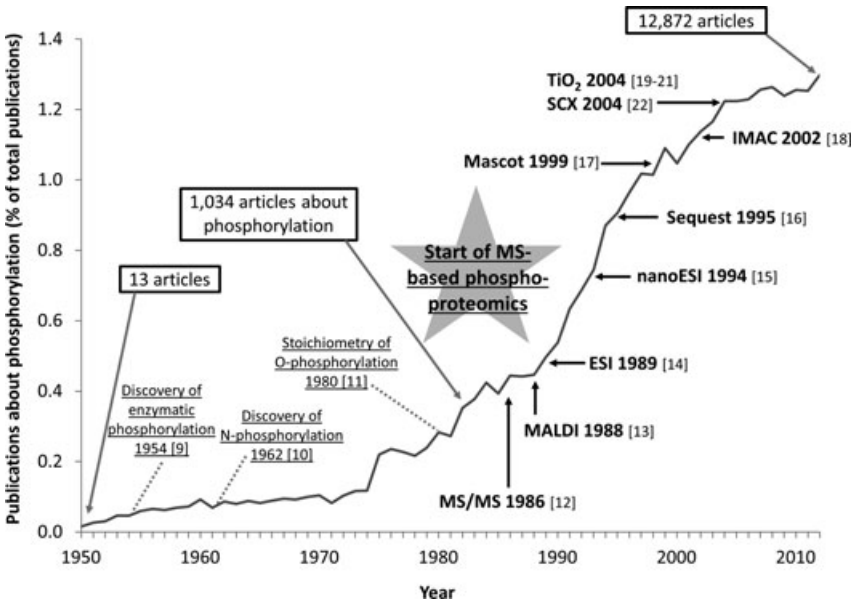

Figure 1. Share of NCBI-listed articles about phosphorylation in comparison to publications in total since 1950. In 2012, over 12 000 articles dealing with phosphorylation were published in scientific journals ("phosphorylation" in title/abstract searched with PubMed—NBCI). In the last 50 years, the share of publications dealing with phosphorylation increased by almost two orders of magnitude compared to the number of publications in total, illustrating the relevance of this research area.

knowledge about phosphorylation (Fig. 1), the precise role of specific phosphorylation sites, networks, and dynamics in the regulation of biochemical processes is still only sparsely understood and requires the simultaneous monitoring of different phosphoproteins in signaling networks to derive their contribution in the respective process. This gap of knowledge has to be filled by combining quantitative phosphoproteomics and systems biology modeling to address clearly defined scientific issues. Appropriate experimental setups will help to understand the roles of different types of phosphoproteins in biological systems and their dynamics with respect to changes in the cellular environment. This could lead to the development of powerful assays not only in life science, but also in clinical research for the characterization of cells and tissues during disease or drug treatment.

In this review, we will give a short overview of the history of phosphorylation research as well as current challenges and future tasks which make phosphorylation still one of the most exciting fields in proteome research.

2 The phosphorylation odyssey—eleven decades of research

In the early 20th century, first results indicated that some proteins do not only consist of nitrogen, oxygen, hydrogen, and sulfur, but also of phosphorus. By treating the chicken egg yolk protein vitellin with ammonia, Leven and Alsberg obtained a substance comprising about 10% phosphorus and 14% nitrogen [23]. Approximately 30 years later, in 1932, Lipmann and Levene described the discovery of serine phosphoric acid by the hydrolysis of vitellinic acid [24]. The phosphorylation of L-threonine in casein was discovered by de Verdier in 1952 using paper chromatography and X-ray crystallography on synthetic phosphoamino acids [25, 26]. Almost at the same time, in 1954, Burnett and Kennedy were able to provide the first evidence for enzymatic phosphorylation after incubating casein and liver mitochondria lysate with ^{32}P-labeled orthophosphate [9]. In 1979, Eckhart et al. discovered the presence of pTyr during their studies on the human papilloma virus [27]. Only one year later, Hunter and Sefton determined the stoichiometry of the *O*-phosphoamino acids in mammalian cells using phosphoamino acid analysis, yielding a ratio of 90:10:0.05 for pSer:pThr:pTyr [11]. Current literature still refers to these ratios although with the advent of MS-based analyses in the late 80s [28, 29], these values shifted toward a higher share (up to 2%) of tyrosine phosphorylation [30, 31]. However, MS-based approaches might bias these ratios since certain regions of proteins are more or less inaccessible to proteases and phosphopeptide enrichment methods might differ in selectivity.

3 How to find the Promised Land? Methods for phosphopeptide enrichment

Protein phosphorylation usually occurs in a substoichiometric manner, making its detection and identification a challenging task. In the past 60 years, a variety of techniques have been developed to address this issue and to enrich the small fraction of phosphorylated peptides/proteins from the non-phosphorylated matrix. These techniques include antibody-based affinity chromatography (AC), immobilized metal ion AC (IMAC), metal oxide AC (MOAC), and HPLC-based fractionation (Fig. 2). Although antibody-based AC is mostly utilized to study pTyr signaling networks [1, 30, 37–39], global phosphoproteome studies often employ techniques such as IMAC or MOAC to target O-phosphorylated peptides.

IMAC was introduced in the late 1980s by Porath using immobilized metal cations such as iron or gallium (non-covalently bound to nitrilotriacetic acid agarose or other

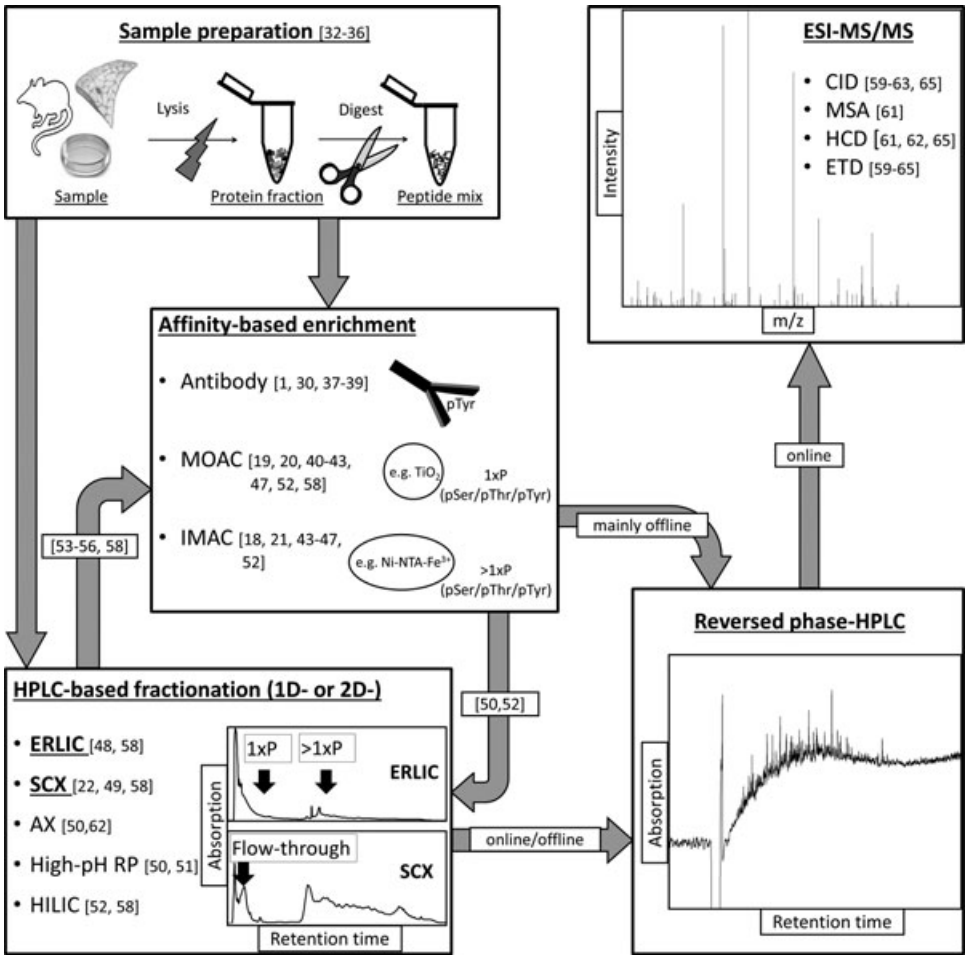

Figure 2. Strategies in phosphoproteomics. Current phosphoproteomic studies mostly employ techniques such as MOAC, IMAC, or antibody-based affinity chromatography combined with HPLC-based fractionation via ERLIC, SCX, anion exchange (AX), high-pH RP, or HILIC for the enrichment of phosphopeptides. Subsequently, nano-LC-MS/MS is performed applying different fragmentation methods such as CID, HCD, or ETD depending on the aims of the study and the instrument used.

matrices) for AC [44, 45]. Originally used for the enrichment of proteins carrying histidyl epitopes, IMAC became an extensively used technique for the enrichment of phosphopeptides from complex peptide samples in the past 15 years [18, 21, 46, 47, 66].

Instead of noncovalently bound metal cations, metal oxides such as TiO_2 and ZrO_2 are utilized in MOAC. The covalent character of the metal oxides enables the enrichment of phosphopeptides under harsh conditions, wherein additives such as dihydroxybenzoic, phthalic, or glycolic acid can be used to displace less affine compounds such as nonphosphorylated acidic peptides. Originally described for the trapping of phosphopeptides on HPLC-precolumns [19, 20, 40], MOAC became very popular as an off-line method using microcolumns [41]. A broad variety of different experimental conditions have been published in recent years, particularly for TiO_2. Optimized protocols combined with state-of-the-art LC-MS can yield thousands of phosphorylation sites with specificities of >90% from less than 1 mg of cellular or tissue protein [42, 43, 50]. As in most areas of proteome research, a combination of strategies can lead to a higher coverage of the phosphoproteome through the enrichment of different subsets of the proteome [67]. A prominent example is sequential elution of IMAC (SIMAC), which combines TiO_2-MOAC and IMAC for the separation of mono- and multiphosphorylated peptide species [47].

HPLC-based phosphopeptide enrichment strategies separate phosphopeptides based on their physicochemical properties such as net charge (IEC) or hydrophilicity (HILIC).

At low pH (e.g., pH 2), tryptic peptides normally have a net charge of 2+ and can easily be retained using strong cation exchange (SCX) chromatography. Introduction of phosphate moieties leads to net charge reduction resulting in early elution of phosphopeptides. In 2004, Beausoleil et al. demonstrated the applicability of SCX for the enrichment of phosphopeptides from a complex human sample by retaining nonphosphorylated peptides on the column, while phosphorylated peptides elute in the flow-through [22]. The retention of peptides can be further enhanced by the use of other proteases leading to higher charge states as demonstrated by Gauci et al. using Lys-N and Lys-C instead of trypsin [49].

In contrast, anion exchange chromatography tends to retain phosphopeptides more effectively than nonphosphorylated peptides [62]. To enhance separation efficiency, high amounts of organic solvent can be used to retain phosphopeptides not only by net charge but also by hydrophilicity. This mode of chromatography was introduced by Alpert in 2008

and termed as electrostatic repulsion hydrophilic interaction chromatography (ERLIC) [48]. However, thus far most studies based on ERLIC have employed relatively large amounts of starting material (≫1 mg) [68–70], currently rendering this method inappropriate for the analysis of small amounts of sample.

Recent studies often employ HPLC-based phosphopeptide enrichment as part of strategies for comprehensive analysis of the phosphoproteome. Often, these multidimensional approaches include off-line AC and in the last dimension, RP chromatography prior to MS analysis (RP-LC-MS/MS).

4 Strategies for large-scale phosphoproteomics

To get a deeper insight into the phosphorylation network of biological systems, it is necessary to combine enrichment techniques that cover complementary parts of the phosphoproteome (Fig. 2). Common combinations are SCX fractionation prior to IMAC [53–55] or MOAC (mainly TiO_2) [56,57] leading to increased phosphopeptide identification rates. A comparison of HPLC-based fractionation methods (ERLIC, HILIC, SCX) prior TiO_2-MOAC was published by Zarei et al. in 2011, demonstrating that SCX-MOAC was the most effective technique for the enrichment of singly phosphorylated peptides, whereas ERLIC-MOAC was the best for the enrichment of multiphosphorylated peptides [58]. Recently, Larsen and coworkers published a combination of TiO_2-IMAC-HILIC as a robust method for global phosphoproteome analysis and demonstrated the benefit of a TiO_2 pre-enrichment prior to SIMAC [52]. In addition, RP-orthogonal HILIC fractionation after MOAC/SIMAC-approaches is a feasible method for the fractionation of the phosphopeptides leading to a significant increase in phosphopeptide identifications [52,71].

In the past few years, 2D- and 3D-LC-MS setups have become more popular, often using combinations of SCX, strong anion exchange, HILIC, or RP at different pHs. In a recent publication, Ficarro et al. identified more than 7000 unique phosphopeptides by high-pH RP, strong anion exchange, and low-pH RP in an online 3D-LC-MS/MS setup [50]. With a prior IMAC enrichment, they identified more than 450 phosphopeptides per μg of cell lysate, but the huge expenditure of time has to be taken into account in multidimensional approaches.

In addition to the above-mentioned approaches, numerous novel but also pseudo-novel methods/protocols for an "improved" enrichment of phosphopeptides have been published throughout the past few years; yet when examined thoroughly only few are clearly beneficial. Currently, the real challenge for phosphoproteomics is not the mere enrichment of phosphopeptides, but (i) the reproducibility of the entire workflow from sample to analysis, (ii) the validity of phosphorylation site localization by existing algorithms, and (iii) the interpretation of large-scale quantitative datasets—especially in the context of the previously mentioned issues.

5 Behind detection there is localization—How to discover the dark side of the moon?

Current literature often uses the number of identified phosphopeptides as a benchmark for comparison, though at least as important as the identification of a phosphopeptide is the correct phosphorylation site localization within the sequence. In proteomics, low-energy CID is often the method of choice when analyzing doubly and triply charged peptides with IT instruments, leading to high-quality MS/MS spectra. When analyzing phosphopeptides, fragmentation via low-energy CID often generates dominant ions derived from neutral losses of phosphoric acid (pSer, pThr), metaphosphoric acid (mostly pTyr), and water, as well as combinations of these losses, leading to spectra which contain little information about the peptide sequences [63, 65]. As an alternative, when using IT instruments, multistage activation excites and fragments neutral loss-derived ions immediately after their generation, thus leading to hybrid MS/MS spectra containing fragment ions derived from the "intact" precursor as well as its neutral loss precursors. Notably, Reid and co-workers showed that IT-based CID (and multistage activation) can suffer from rearrangement of phosphate groups when multiple potential donor and acceptor amino acids are present, thus potentially leading to flawed phosphorylation site localizations [72, 73].

A third technique, beam-type CID (also referred to as higher energy collision dissociation, HCD, in Orbitraps) reduces the probability of neutral loss-derived ion generation compared to IT CID. Currently, it is controversially discussed whether high mass accuracy Orbitrap HCD outperforms low mass accuracy Orbitrap CID for phosphopeptide identification [74]. In general, IT CID is more sensitive and requires less ions to generate high-quality MS/MS spectra, yet the higher mass accuracy of HCD MS/MS data facilitates the confident identification of fragment ions and consequently (phospho)peptides.

When using electron transfer dissociation (ETD), neutral losses can be completely avoided and fragmentation occurs predominantly along the peptide's backbone. In contrast to CID, efficient ETD fragmentation requires at least triply charged precursors. Therefore, alternative proteases generating peptides with higher charge states (compared to tryptic peptides) have to be employed or decision tree-driven methods can be used for fragmentation choosing either CID or ETD depending on the charge state and mass of the precursor [60]. ETD is one of the most promising methods for confident phosphosite localization, although common database search algorithms might be inadequate for this type of MS/MS spectra [75]. Notably, ETD is less sensitive, often shows lower MS/MS sequence coverage and is 10 to 20 times slower than a normal CID (10 ms vs. 100–200 ms, depending on the number and charge states of the precursor ions and the precursor-anion ratio in the trap) leading to the prerequisite of reduced sample complexity or adequate peptide separation prior to MS analysis [64].

6 Confident phosphosite annotation requires an appropriate algorithm

In large-scale experiments, confident phosphorylation site annotation not only depends on the fragmentation method, but also on the validation algorithm. There are several validation algorithms available, calculating probabilities for correct site assignment or probabilities for each potential site to be phosphorylated in a given peptide. However in principle, the performance of such algorithms cannot be readily benchmarked. Savitski et al. utilized a set of 180 synthetic phosphopeptides, analyzed by different MS instruments and fragmentation techniques to validate phosphosite assignment based on the difference between the first and second best Mascot hit for a given phosphorylated sequence [74]. In this study, their Mascot Delta Score (MD score), a modified version of a long-established strategy to assess phosphosite localization, outperforms a self-made version of the Ascore by Gygi and co-workers [76]—the first software for validating phosphorylation site assignment—in the analysis of synthetic tyrosine phosphorylated peptides.

In 2011, Taus et al. introduced phosphoRS as a novel algorithm for phosphosite localization and evaluated its performance using a set of synthetic phosphopeptides [61]. The algorithm was implemented in the widely used commercial Proteome Discoverer Software (PD 1.3) and yields higher numbers of identified phosphorylation sites at a given confidence compared to Ascore and MD score. We have recently developed a novel approach utilizing a combination of X!Tandem, Mascot, and OMSSA [77] leading to a substantial improvement (up to 25%) of correct phosphosite localization, when searching the dataset of Savitski et al., compared to the MD score alone [78, 79].

Notably, these studies are all based on an idealized sample consisting of a relatively low number of synthetic phosphopeptides without background and thus do not necessarily reflect real life samples which are much more complex and often suffer from co-elution of different phospho-isoforms as well as very low signal intensities and poor spectrum quality. Especially in quantitative phosphoproteomics, the presence of different phospho-isoforms of the same peptide, sometimes even varying in the number of phosphorylation sites, is a big challenge which can easily interfere with accurate quantitation.

Thus, a more thorough evaluation of site localization in quantitative studies would require a more realistic phosphorylation standard which also takes into account the sample complexity usually encountered in phosphoproteomics, easily exceeding thousands of phosphopeptides [80].

In addition to the choice of search and phosphosite validation algorithms, confident site localization also depends on the choice of MS-instrumentation, fragmentation method, as well as on the (structural) properties of the precursors.

7 Does it really matter? Approaches for relative and absolute quantitation

Quantitative phosphoproteomics is an extremely challenging field. The mostly low abundance of phosphopeptides and accordingly (i) the need to enrich for phosphopeptides, (ii) the difficulty in confidently localizing the correct phosphorylation site, and (iii) the potential presence of multiple phosphoisoforms of the same peptide, complicate accurate quantitation even more than in global proteomic studies—especially when considering typically encountered levels of biological and technical variation. In principle, chemical [81] and metabolic stable isotope labeling techniques [82] in conjunction with phosphopeptide enrichment enable relative quantification of phosphopeptides between different samples, but cannot be used to determine phosphorylation site stoichiometry. However, multiple reaction monitoring (MRM) with spiked in isotopically labeled peptides can be used for absolute quantification of peptides and phosphopeptides and thus phosphorylation stoichiometry [83, 84]. As shown by Borchers and co-workers, phosphatase-based dephosphorylation of phosphopeptides and subsequent quantitation using MRM enables precise determination of stoichiometries in a targeted setup [85]. Measuring samples with and without phosphatase treatment, they described the phosphorylation stoichiometries of proteins involved in cancer signaling, demonstrating the power of this tool for understanding biological processes and pathways. Still the occupancy of different phosphorylation sites within the same peptide cannot be assessed using this approach. However, in contrast to discovery approaches, targeted MRM allows to monitor and differentiate even co-eluting phosphopeptide isomers based on isoform-specific fragment ions [78].

Although far from covering the "complete" phosphoproteome, targeted MRM and targeted MS/MS approaches might play an increasingly powerful role in understanding cellular signaling networks in the future.

8 Dos and don'ts—Commandments for phosphopeptide-specific sample preparation

From cell lysis to protein digest, phosphopeptide identification critically relies on sample preparation techniques. Disruption of cell membranes, isolation of proteins, and setting up the right conditions for protein digestion are only some of the important tasks. If not carefully chosen, the lysis conditions can lead to unanticipated peptide modifications, such as extensive carbamylation (Lys residues and N-termini) of up to 20% of peptides, in presence of urea [86]. Protein samples containing detergents such as SDS or NP-40 need to be purified, for example, by protein precipitation techniques or filter-aided sample preparation, a method introduced in 2005 by the Liebler group and later modified by Mann and co-workers [32, 87].

Furthermore, setting up the right digest conditions to avoid unspecific, semitryptic, or missed cleavage peptide generation is still one of the remaining challenges in the field of quantitative (phospho-) proteomics. Recently, a comprehensive investigation about trypsin quality and efficiency was published by Burkhart et al. [33]. As shown in this study, HPLC-based monolithic systems can be used as powerful tools for quality control with respect to digestion efficiency and reproducibility, as wells as for the correction of variability in sample amounts prior to quantitative analyses.

Quality control in the sample preparation process is crucial for successful and reproducible research and becomes even more important in the strong emerging field of quantitative phosphoproteomics [34, 35, 80]. Equally important, but often neglected, is the choice of consumables and solvents for sample preparation and LC-MS analysis [88]. Inappropriate storage of samples can lead to plasticizer contaminations in the presence of organic solvents and to the adsorption of peptide species to surfaces leading to systematic bias in quantification [36]. Thus, it is mandatory to evaluate the analysis and sample preparation strategies by using reference samples, for example, synthetic phosphopeptides or phosphoproteins such as casein, spiked into a background, to optimize the conditions for all of the important steps in the analytical workflow.

9 There is more—Phosphorylation of basic amino acids

Apart from the phosphorylation of the hydroxyl groups of the amino acids serine, threonine, and tyrosine, protein phosphorylation can also occur on nitrogen atoms in the side chains of the basic amino acids lysine, arginine, and histidine. First reports of phosphorylated lysine in mammalian nuclear protein histone H1 go back to the early 70s [89, 90], but so far the chemical properties of phospholysine, its biological function, and the underlying protein kinases and phosphatases have not yet been clearly revealed [91]. In contrast, at least in prokaryotes, the role of phosphoarginine is well characterized: phosphoarginine is involved in the stress response and protein degradation system of *Bacillus subtilis*, wherein the arginine kinase McsB controls the expression of stress response genes via phosphorylation of the heat-shock regulator protein CtsR [92].

As far back as 50 years ago, protein histidine phosphorylation was discovered in rat liver mitochondria by Boyer et al. [10]. Phosphohistidine (pHis) is the best studied basic phosphoamino acid and was shown to represent 6% of total protein phosphorylation in lower eukaryotes (e.g., amoebae), and at least ten times more abundant than pTyr [93]. The phosphorylation of histidine residues plays a major role in two- and multicomponent phosphosignaling pathways in prokaryotes, plants and fungi [94, 95]. Nevertheless, its relevance for mammalian cellular processes was questionable for many years. The first mammalian phosphohistidine phosphatase, phosphohistidine phosphatase 1, was not identified until 2002 [96, 97] and, apart from nucleoside diphosphate kinase, its associated kinases have not been well characterized [91]. Recently, the potential role of pHis in basic mammalian cellular processes, such as G-protein signaling [98], chromatin assembly [90], and ion conduction [99], is becoming clearer due to new sophisticated biochemical and analytical techniques.

10 Chemical properties of phosphoramidates

The structures of the basic phosphoamino acid residues are shown in Fig. 3. Notably, two biologically relevant isomers of pHis, 1- pHis, and 3-pHis, exist, named according to the position of the phosphorylation on the imidazole ring [100]. Our limited knowledge about *N*-phosphorylated amino acids is mostly due to their chemical instability. Compared to the P-O bond in phosphohydroxyamino acids, the P-N bond in basic phosphoamino acids is thermodynamically less stable due to the higher Gibbs free energy of hydrolysis [101]. In addition, phosphoramidates are also much less kinetically stable as the protonation of the nitrogen of the P-N bond under acidic conditions facilitates the dissociation of the phosphoryl group [102]. The acid lability, in particular, complicates the detection and analysis of phosphoramidates in biological samples, since techniques for the analysis of protein phosphorylation such as phosphoamino acid analysis and RP LC traditionally involve acidic conditions and therefore have to be modified.

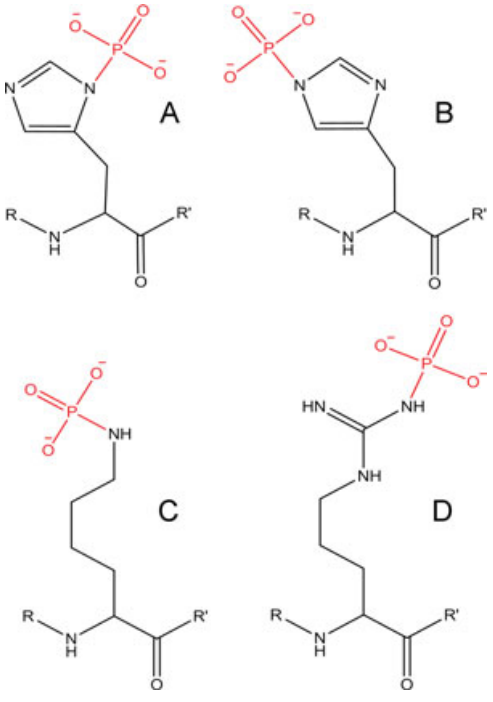

Figure 3. Structures of the basic phosphoamino acid residues. A characteristic of phosphohistidine is the existence of two biologically relevant structures: 1-phosphohistidine (A) and 3-phosphohistidine (B), besides phospholysine (C), and phosphoarginine (D) are illustrated.

11 Detection and analysis of basic phosphoamino acids

^{31}P NMR spectroscopy [103] and other biochemical methods have been applied for the detection of basic phosphoamino acids such as phosphatase and kinase activity assays [94], in-gel kinase assays [104], phosphoamino acid analysis [105, 106] and immunoprecipitation with phospho-specific antibodies [107].

Especially pTyr-specific antibodies have been extremely useful for determining the abundance and characteristics of pTyr in biological systems [37]. However, thus far, attempts to generate pHis-specific antibodies using pHis as an antigen have failed due to weak immune responses [108]. The lack of immune response might be attributable to the rapid dephosphorylation of phosphoramidates. Nonhydrolyzable analogs, at least for pHis, have been successfully used to raise antibodies, which selectively recognize the phosphorylation of His18 residue in recombinant protein histone H4 [107]. However, the cross-reactivity of this antibody toward pTyr has not yet been elucidated.

Recently, more and more MS-based approaches are being applied, although standard proteomic workflows utilizing acidic conditions for sample preparation and LC separation cannot be directly transferred to the analysis of *N*-phosphorylated peptides. Several studies have demonstrated the elimination of the phosphoryl moiety of pHis when acidic eluents are used in LC separations [51, 109]. Therefore, Kleinnijenhuis et al. [51] evaluated different buffer systems for the analysis of pHis peptides, stating that a neutral pH (pH 6.7) or a short acidic gradient over 7 min were the optimal HPLC conditions.

To overcome the substoichiometric level of phosphorylated peptides, different strategies for the enrichment of *N*-phosphorylated peptides prior to LC-MS analysis have been evaluated.

Napper et al. [110] tested three metal coatings of IMAC resins for their selectivity toward pHis peptides from proteolytic digests. Only Cu(II)-IMAC resin led to the selective detection of pHis peptides of the histidine-containing protein (HPr) from *Escherichia coli*. Alternatively, Kleinnijenhuis et al. [51] used Strata-X material, which retains nonphosphorylated peptides by hydrophobic interactions, to reduce the complexity of a tryptic digest of EnvZc from *E. coli*. However, with both strategies phosphopeptides were still less abundant than their nonphosphorylated counterparts, indicating the urgent need for new enrichment strategies, as none of the strategies for the enrichment of *O*-phosphorylated peptides seem to be applicable. Moreover, database search parameters critically influence the outcome of *N*-phosphorylation analyses as flawed settings can lead to incorrect identifications and phosphorylation site localizations: If only pHis is chosen as variable modification, a database search of a normal phospho-dataset (HeLa after TiO_2-HILIC enrichment, searched with Mascot and validated with phosphoRS) resulted in the detection of approximately 100 peptides containing pHis, whereas the same dataset shows no confident hits (phosphoRS site probability >80%) for pHis when considering pSer/pThr/pTyr as well.

Peptides containing basic phosphoamino acids have been successfully detected by MS using MALDI and ESI ionization techniques [111]. For pHis and phosphoarginine peptides, losses of phosphoric acid and less frequently metaphosphoric acid have been observed upon CID. Hence, Schmidt et al. [59] reported a higher number of incorrect phosphorylation site localizations utilizing CID and HCD compared to ETD due to the occurrence of these neutral losses. In conclusion, electron-based fragmentation methods such as ETD seem to be the method of choice for MS/MS fragmentation of *N*-phosphorylated peptides, as this method yields high-quality spectra, with few incorrect localizations and a high number of identified phosphopeptides. In addition, this method also avoids misinterpretation of actual *O*-phosphorylated peptides.

Notably, all these studies are based on in vitro phosphorylated samples containing relatively high numbers of phosphoamino acids. The detection and analysis of phosphoramidates in real biological samples remains extremely challenging.

12 Concluding remarks

With continuous advancements in MS-based phosphoproteomics, highly specific and sensitive strategies for the enrichment and analysis of phosphopeptides have enabled the identification of thousands of unique phosphorylation sites in prokaryotes and eukaryotes. Despite this, researchers still have to overcome pitfalls that occur during sample preparation and MS analysis, especially in respect to fragmentation techniques and subsequent search and validation algorithms, which critically influence the reliability of the experiment. Furthermore, like in global proteomics, a major problem is the poor reproducibility of phosphoproteomic studies, even with regard to technical replicates.

Regarding the analysis of *N*-phosphorylated peptides, there is still a long way to go, although with the adaption of traditional LC-MS strategies to the requirements of these acid-labile phosphoamino acids, further roles in the cellular processes of mammalian will surely be revealed.

Quantitative analysis of transient phosphorylation events in signal transduction pathways is the most challenging task in the area of phosphoproteomics as phosphorylation occurs within milliseconds to seconds. Moreover, when processing biological samples, the activities of phosphatases and proteases can easily bias the results of quantitative studies. Therefore, setting up standardized procedures for sample preparation, protein labeling, and data processing as well as introducing rules for reliable research based on quality control of spectra might lead to a higher comparability of studies performed in different laboratories. Furthermore, developing reliable approaches for the assessment of false localization rates similar to false discovery rates could enhance quality of phosphorylation site databases.

The financial support by the Ministerium für Innovation, Wissenschaft und Forschung des Landes Nordrhein-Westfalen and by the Bundesministerium für Bildung und Forschung (DYNAMO) is gratefully acknowledged.

The authors have declared no conflict of interest.

13 References

[1] Dengjel, J., Akimov, V., Olsen, J. V., Bunkenborg, J., Mann, M., Blagoev, B., Andersen, J. S., *Nat. Biotechnol.* 2007, *25*, 566–568.

[2] Eyrich, B., Sickmann, A., Zahedi, R. P., *Proteomics* 2011, *11*, 554–570.

[3] Sickmann, A., Meyer, H. E., *Proteomics* 2001, *1*, 200–206.

[4] Kuhlberg, A., Haid, M., Metzger, S., *J. Biol. Chem.* 2010, *285*, 31484–31490.

[5] Feramisco, J. R., Kemp, B. E., Krebs, E. G., *J. Biol. Chem.* 1979, *254*, 6987–6990.

[6] Linn, T. C., Pettit, F. H., Reed, L. J., *Proc. Natl. Acad. Sci. USA* 1969, *62*, 234–241.

[7] Linn, T. C., Pettit, F. H., Hucho, F., Reed, L. J., *Proc. Natl. Acad. Sci. USA* 1969, *64*, 227–234.

[8] Fardilha, M., Esteves, S. L., Korrodi-Gregorio, L., da Cruz e Silva, O. A., da Cruz e Silva, F. F., *Curr. Med. Chem.* 2010, *17*, 3996–4017.

[9] Burnett, G., Kennedy, E. P., *J. Biol. Chem.* 1954, *211*, 969–980.

[10] Boyer, P. D., Deluca, M., Ebner, K. E., Hultquist, D. E., Peter, J. B., *J. Biol. Chem.* 1962, *237*, PC3306–PC3308.

[11] Hunter, T., Sefton, B. M., *Proc. Natl. Acad. Sci. USA* 1980, *77*, 1311–1315.

[12] Hunt, D. F., Yates, J. R. 3rd, Shabanowitz, J., Winston, S., Hauer, C. R., *Proc. Natl. Acad. Sci. USA* 1986, *83*, 6233–6237.

[13] Karas, M., Hillenkamp, F., *Anal. Chem.* 1988, *60*, 2299–2301.

[14] Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., Whitehouse, C. M., *Science* 1989, *246*, 64–71.

[15] Wilm, M. S., Mann, M., *Int. J. Mass Spectrom. Ion Process.* 1994, *136*, 167–180.

[16] Yates, J. R. 3rd, Eng, J. K., McCormack, A. L., Schieltz, D., *Anal. Chem.* 1995, *67*, 1426–1436.

[17] Perkins, D. N., Pappin, D. J., Creasy, D. M., Cottrell, J. S., *Electrophoresis* 1999, *20*, 3551–3567.

[18] 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.

[19] Pinkse, M. W., Uitto, P. M., Hilhorst, M. J., Ooms, B., Heck, A. J., *Anal. Chem.* 2004, *76*, 3935–3943.

[20] Kuroda, I., Shintani, Y., Motokawa, M., Abe, S., Furuno, M., *Anal. Sci.* 2004, *20*, 1313–1319.

[21] Brill, L. M., Salomon, A. R., Ficarro, S. B., Mukherji, M., Stettler-Gill, M., Peters, E. C., *Anal. Chem.* 2004, *76*, 2763–2772.

[22] Beausoleil, S. A., Jedrychowski, M., Schwartz, D., Elias, J. E., Villen, J., Li, J., Cohn, M. A., Cantley, L. C., Gygi, S. P., *Proc. Natl. Acad. Sci. USA* 2004, *101*, 12130–12135.

[23] Levene, P. A., Alsberg, C. L., *J. Biol. Chem.* 1906, *2*, 127–133.

[24] Lipmann, F. A., Levene, P. A., *J. Biol. Chem.* 1932, *98*, 109–114.

[25] de Verdier, C. H., *Nature* 1952, *170*, 804–805.

[26] de Verdier, C.-H., *Acta Chemica Scandinavica* 1953, *7*, 196–200.

[27] Eckhart, W., Hutchinson, M. A., Hunter, T., *Cell* 1979, *18*, 925–933.

[28] Holmes, C. F., Tonks, N. K., Major, H., Cohen, P., *Biochimica et biophysica acta* 1987, *929*, 208–219.

[29] Erickson, A. K., Payne, D. M., Martino, P. A., Rossomando, A. J., Shabanowitz, J., Weber, M. J., Hunt, D. F., Sturgill, T. W., *J. Biol. Chem.* 1990, *265*, 19728–19735.

[30] Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., Mann, M., *Cell* 2006, *127*, 635–648.

[31] Zhang, G., Neubert, T. A., *J. Proteome Res.* 2011, *10*, 5454–5462.

[32] Manza, L. L., Stamer, S. L., Ham, A. J., Codreanu, S. G., Liebler, D. C., *Proteomics* 2005, *5*, 1742–1745.

[33] Burkhart, J. M., Schumbrutzki, C., Wortelkamp, S., Sickmann, A., Zahedi, R. P., *J. Proteomics* 2012, *75*, 1454–1462.

[34] Beck, F., Burkhart, J. M., Geiger, J., Zahedi, R. P., Sickmann, A., *Methods Mol. Biol.* 2012, *893*, 101–113.

[35] Burkhart, J. M., Vaudel, M., Zahedi, R. P., Martens, L., Sickmann, A., *Proteomics* 2011, *11*, 1125–1134.

[36] Kraut, A., Marcellin, M., Adrait, A., Kuhn, L., Louwagie, M., Kieffer-Jaquinod, S., Lebert, D., Masselon, C. D., Dupuis, A., Bruley, C., Jaquinod, M., Garin, J., Gallagher-Gambarelli, M., *J. Proteome Res.* 2009, *8*, 3778–3785.

[37] Rush, J., Moritz, A., Lee, K. A., Guo, A., Goss, V. L., Spek, E. J., Zhang, H., Zha, X. M., Polakiewicz, R. D., Comb, M. J., *Nat. Biotechnol.* 2005, *23*, 94–101.

[38] Dengjel, J., Kratchmarova, I., Blagoev, B., *Mol. Biosyst.* 2009, *5*, 1112–1121.

[39] Lee, C. F., Griffiths, S., Rodriguez-Suarez, E., Pierce, A., Unwin, R. D., Jaworska, E., Evans, C. A., S, J. Gaskell, Whetton, A. D., *Proteomics* 2010, *10*, 3321–3342.

[40] Sano, A., Nakamura, H., *Anal. Sci.* 2004, *20*, 565–566.

[41] Larsen, M. R., Thingholm, T. E., Jensen, O. N., Roepstorff, P., Jorgensen, T. J., *Mol. Cell. Proteomics* 2005, *4*, 873–886.

[42] Engholm-Keller, K., Larsen, M. R., *J. Proteomics* 2011, *75*, 317–328.

[43] Thingholm, T. E., Jorgensen, T. J., Jensen, O. N., Larsen, M. R., *Nat. Protoc.* 2006, *1*, 1929–1935.

[44] Yip, T. T., Tam, Y. Y., Kong, Y. C., Belew, M. C., Porath, J., *FEBS Lett.* 1985, *187*, 345–348.

[45] Porath, J., *Protein Expr. Purif.* 1992, *3*, 263–281.

[46] Posewitz, M. C., Tempst, P., *Anal. Chem.* 1999, *71*, 2883–2892.

[47] Thingholm, T. E., Jensen, O. N., Robinson, P. J., Larsen, M. R., *Mol. Cell. Proteomics* 2008, *7*, 661–671.

[48] Alpert, A. J., *Anal. Chem.* 2008, *80*, 62–76.

[49] Gauci, S., Helbig, A. O., Slijper, M., Krijgsveld, J., Heck, A. J., Mohammed, S., *Anal. Chem.* 2009, *81*, 4493–4501.

[50] Ficarro, S. B., Zhang, Y., Carrasco-Alfonso, M. J., Garg, B., Adelmant, G., Webber, J. T., Luckey, C. J., Marto, J. A., *Mol. Cell. Proteomics* 2011, *10*, O111 011064.

[51] Kleinnijenhuis, A. J., Kjeldsen, F., Kallipolitis, B., Haselmann, K. F., Jensen, O. N., *Anal. Chem.* 2007, *79*, 7450–7456.

[52] Engholm-Keller, K., Birck, P., Storling, J., Pociot, F., Mandrup-Poulsen, T., Larsen, M. R., *J. Proteomics* 2012, *75*, 5749–5761.

[53] Endler, A., Reiland, S., Gerrits, B., Schmidt, U. G., Baginsky, S., Martinoia, E., *Proteomics* 2009, *9*, 310–321.

[54] Villen, J., Beausoleil, S. A., Gerber, S. A., Gygi, S. P., *Proc. Natl. Acad. Sci. USA* 2007, *104*, 1488–1493.

[55] Wagner, P. D., Vu, N. D., *Biochem. J.* 2000, *346*(Pt 3), 623–630.

[56] Wu, J., Shakey, Q., Liu, W., Schuller, A., Follettie, M. T., *J. Proteome Res.* 2007, *6*, 4684–4689.

[57] Pinkse, M. W., Mohammed, S., Gouw, J. W., van Breukelen, B., Vos, H. R., Heck, A. J., *J. Proteome Res.* 2008, *7*, 687–697.

[58] Zarei, M., Sprenger, A., Metzger, F., Gretzmeier, C., Dengjel, J., *J. Proteome Res.* 2011, *10*, 3474–3483.

[59] Schmidt, A., Ammerer, G., Mechtler, K., *Proteomics* 2013, *13*, 945–954.

[60] Swaney, D. L., McAlister, G. C., Coon, J. J., *Nat. Methods* 2008, *5*, 959–964.

[61] Taus, T., Kocher, T., Pichler, P., Paschke, C., Schmidt, A., Henrich, C., Mechtler, K., *J. Proteome Res.* 2011, *10*, 5354–5362.

[62] Hennrich, M. L., Groenewold, V., Kops, G. J., Heck, A. J., Mohammed, S., *Anal. Chem.* 2011, *83*, 7137–7143.

[63] Boersema, P. J., Mohammed, S., Heck, A. J., *J. Mass Spectrom.* 2009, *44*, 861–878.

[64] Good, D. M., Wirtala, M., McAlister, G. C., Coon, J. J., *Mol. Cell. Proteomics* 2007, *6*, 1942–1951.

[65] Frese, C. K., Altelaar, A. F., Hennrich, M. L., Nolting, D., Zeller, M., Griep-Raming, J., Heck, A. J., Mohammed, S., *J. Proteome Res.* 2011, *10*, 2377–2388.

[66] Thingholm, T., Jensen, O., Larsen, M., in: Graauw, M. (Ed.), *Phospho-Proteomics*, Humana Press *Methods Mol. Biol.* 2009, *527*, 67–78.

[67] Bodenmiller, B., Mueller, L. N., Mueller, M., Domon, B., Aebersold, R., *Nat. Methods* 2007, *4*, 231–237.

[68] Gan, C. S., Guo, T., Zhang, H., Lim, S. K., Sze, S. K., *J. Proteome Res.* 2008, *7*, 4869–4877.

[69] Hao, P., Guo, T., Sze, S. K., *PLoS One* 2011, *6*, e16884.

[70] Zarei, M., Sprenger, A., Gretzmeier, C., Dengjel, J., *J. Proteome Res.* 2012, *11*, 4269–4276.

[71] McNulty, D. E., Annan, R. S., *Mol. Cell. Proteomics* 2008, *7*, 971–980.

[72] Palumbo, A. M., Reid, G. E., *Anal. Chem.* 2008, *80*, 9735–9747.

[73] Cui, L., Reid, G. E., *Proteomics* 2013, *13*, 964–973.

[74] Savitski, M. M., Lemeer, S., Boesche, M., Lang, M., Mathieson, T., Bantscheff, M., Kuster, B., *Mol. Cell. Proteomics* 2011, *10*, M110 003830.

[75] Kim, S., Mischerikow, N., Bandeira, N., Navarro, J. D., Wich, L., Mohammed, S., Heck, A. J., Pevzner, P. A., *Mol. Cell. Proteomics* 2010, *9*, 2840–2852.

[76] Beausoleil, S. A., Villen, J., Gerber, S. A., Rush, J., Gygi, S. P., *Nat. Biotechnol.* 2006, *24*, 1285–1292.

[77] Vaudel, M., Breiter, D., Beck, F., Rahnenfuhrer, J., Martens, L., Zahedi, R. P., *Proteomics* 2013, *13*, 1036–1041.

[78] Beck, F., Lewandrowski, U., Wiltfang, M., Feldmann, I., Geiger, J., Sickmann, A., Zahedi, R. P., *Proteomics* 2011, *11*, 1099–1109.

[79] Vaudel, M., Burkhart, J. M., Breiter, D., Zahedi, R. P., Sickmann, A., Martens, L., *J. Proteome Res.* 2012, *11*, 5065–5071.

[80] Ivanov, A., Colangelo, C., Dufresne, C., Friedman, D., Lilley, K., Mechtler, K., Phinney, B., Rose, K., Rudnick, P., Searle, B., Shaffer, S., Weintraub, S., *Proteomics* 2013, *13*, 904–909.

[81] Pflieger, D., Junger, M. A., Muller, M., Rinner, O., Lee, H., Gehrig, P. M., Gstaiger, M., Aebersold, R., *Mol. Cell. Proteomics* 2008, *7*, 326–346.

[82] Hegeman, A. D., Harms, A. C., Sussman, M. R., Bunner, A. E., Harper, J. F., *J. Am. Soc. Mass Spectrom.* 2004, *15*, 647–653.

[83] Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., Gygi, S. P., *Proc. Natl. Acad. Sci. USA* 2003, *100*, 6940–6945.

[84] Ciccimaro, E., Hanks, S. K., Yu, K. H., Blair, I. A., *Anal. Chem.* 2009, *81*, 3304–3313.

[85] Domanski, D., Murphy, L. C., Borchers, C. H., *Anal. Chem.* 2010, *82*, 5610–5620.

[86] Kollipara, L., Zahedi, R., *Proteomics* 2013, *13*, 941–944.

[87] Wisniewski, J. R., Zougman, A., Nagaraj, N., Mann, M., *Nat. Methods* 2009, *6*, 359–362.

[88] Winter, D., Seidler, J., Ziv, Y., Shiloh, Y., Lehmann, W. D., *J. Proteome Res.* 2009, *8*, 418–424.

[89] Sikorska, M., Whitfield, J. F., *Biochimica et biophysica acta* 1982, *703*, 171–179.

[90] Chen, C. C., Smith, D. L., Bruegger, B. B., Halpern, R. M., Smith, R. A., *Biochemistry* 1974, *13*, 3785–3789.

[91] Attwood, P. V., *Biochimica et biophysica acta* 2013, *1834*, 470–478.

[92] Fuhrmann, J., Schmidt, A., Spiess, S., Lehner, A., Turgay, K., Mechtler, K., Charpentier, E., Clausen, T., *Science* 2009, *324*, 1323–1327.

[93] Pesis, K. H., Wei, Y., Lewis, M., Matthews, H. R., *FEBS Lett.* 1988, *239*, 151–154.

[94] Matthews, H. R., *Pharmacol. Ther.* 1995, *67*, 323–350.

[95] Schaller, G. E., Shiu, S. H., Armitage, J. P., *Curr. Biol.* 2011, *21*, R320–R330.

[96] Ek, P., Pettersson, G., Ek, B., Gong, F., Li, J. P., Zetterqvist, O., *Eur. J. Biochem.* 2002, *269*, 5016–5023.

[97] Klumpp, S., Hermesmeier, J., Selke, D., Baumeister, R., Kellner, R., Krieglstein, J., *J. Cereb. Blood Flow Metab.* 2002, *22*, 1420–1424.

[98] Cuello, F., Schulze, R. A., Heemeyer, F., Meyer, H. E., Lutz, S., Jakobs, K. H., Niroomand, F., Wieland, T., *J Biol. Chem.* 2003, *278*, 7220–7226.

[99] Srivastava, S., Li, Z., Ko, K., Choudhury, P., Albaqumi, M., Johnson, A. K., Yan, Y., Backer, J. M., Unutmaz, D., Coetzee, W. A., Skolnik, E. Y., *Mol. Cell* 2006, *24*, 665–675.

[100] Hultquist, D. E., Moyer, R. W., Boyer, P. D., *Biochemistry* 1966, *5*, 322–331.

[101] Stock, J. B., Stock, A. M., Mottonen, J. M., *Nature* 1990, *344*, 395–400.

[102] Duclos, B., Marcandier, S., Cozzone, A. J., *Methods Enzymol.* 1991, *201*, 10–21.

[103] Besant, P. G., Byrne, L., Thomas, G., Attwood, P. V., *Anal. Biochem.* 1998, *258*, 372–375.

[104] Besant, P. G., Attwood, P. V., *Int. J. Biochem. Cell Biol.* 2000, *32*, 243–253.

[105] Tan, E., Lin Zu, X., Yeoh, G. C., Besant, P. G., Attwood, P. V., *Anal. Biochem.* 2003, *323*, 122–126.

[106] Zetterqvist, O., *Biochimica et biophysica acta* 1967, *136*, 279–285.

[107] Kee, J. M., Villani, B., Carpenter, L. R., Muir, T. W., *J. Am. Chem. Soc.* 2010, *132*, 14327–14329.

[108] Attwood, P. V., Piggott, M. J., Zu, X. L., Besant, P. G., *Amino Acids* 2007, *32*, 145–156.

[109] Lapek, J. D. Jr., Tombline, G., Friedman, A. E., *J. Proteome Res.* 2011, *10*, 751–755.

[110] Napper, S., Kindrachuk, J., Olson, D. J., Ambrose, S. J., Dereniwsky, C., Ross, A. R., *Anal. Chem.* 2003, *75*, 1741–1747.

[111] Besant, P. G., Attwood, P. V., *Mol. Cell. Biochem.* 2009, *329*, 93–106.

