RESEARCH ARTICLE

Studying the fragmentation behavior of peptides with arginine phosphorylation and its influence on phospho-site localization

Andreas Schmidt[1,2], *Gustav Ammerer*[1] *and Karl Mechtler*[2,3]

[1] Christian-Doppler-Laboratory for Proteome Analysis, Department of Biochemistry and Microbiology, Max F. Perutz Laboratories, University of Vienna, Vienna, Austria
[2] Institute for Molecular Pathology (IMP), Vienna, Austria
[3] Institute for Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), Vienna, Austria

Received: June 12, 2012
Revised: September 10, 2012
Accepted: October 1, 2012

Phospho-proteomic studies opened a broad view onto the main mechanisms of regulating cellular processes. Our recent discovery of a protein arginine kinase and its target in bacteria added a previously undescribed type of phosphorylation to control protein activity. Several challenges arise from large in vivo studies of this and other types of phosphorylations. The main factors impeding correct localization are low spectral quality, neutral loss of phosphoric acid, and gas-phase rearrangements, which have recently been described for phospho-serine, -threonine, and -tyrosine. Studies on histidine-phosphorylated peptides, a nitrogen-bound phosphorylation, also reported loss of phosphoric acid upon collision-induced dissociation. We were interested in studying the behaviour of arginine phosphorylation under different fragmentation conditions and its influence on site localization. First, we determined the percentage of false localizations obtained by three different search engines and a software tool dedicated for phospho-site determination. Next, we demonstrate that application of collisional activation for analysis of arginine-phosphorylated peptides leads to extensive elimination of phosphoric acid and increases the numbers of false localizations, while the modification is maintained on the arginine side chain upon electron-transfer dissociation. Furthermore, we also observed a rearrangement of the phosphorylation onto serine and glutamic acid side chains upon collisional activation.

Keywords:
MS/MS / Nanoproteomics / Peptide sequencing / Phosphopeptides / Phospho-site localization

Additional supporting information may be found in the online version of this article at the publisher's web-site

1 Introduction

In all domains of life, protein phosphorylation has been established as a key regulator of many cellular processes [1]. The human genome encodes for more than 500 protein kinases and 200 protein phosphatases, which conjointly control levels of protein phosphorylation [2, 3]. Innovations in selective enrichment, chromatographic separation, and subsequent MS/MS analysis enhanced the discovery of thousands of novel protein phosphorylation sites on serine (pSer), threonine (pThr), and tyrosine (pTyr) (Fig. 1) [4–8].

Nevertheless, correctly localizing the phosphorylated amino acid within the identified phosphopeptide sequence represents one of the most challenging tasks in such studies [6, 9]. The huge number of phosphopeptides identified in single experiments creates the need for automation of

Correspondence: Dr. Karl Mechtler, Institute for Molecular Pathology Vienna, IMP, Dr. Bohr-Gasse 7, Vienna A-1030, Austria
E-mail: mechtler@imp.ac.at
Fax: +43 (1) 790 44-110

Abbreviations: ETD, electron transfer dissociation; **FA**, formic acid; **FDR**, false discovery rate; **HCD**, higher-energy collision-induced dissociation; **MSA**, multi-stage activation; **NCE**, normalized collision energy

© 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

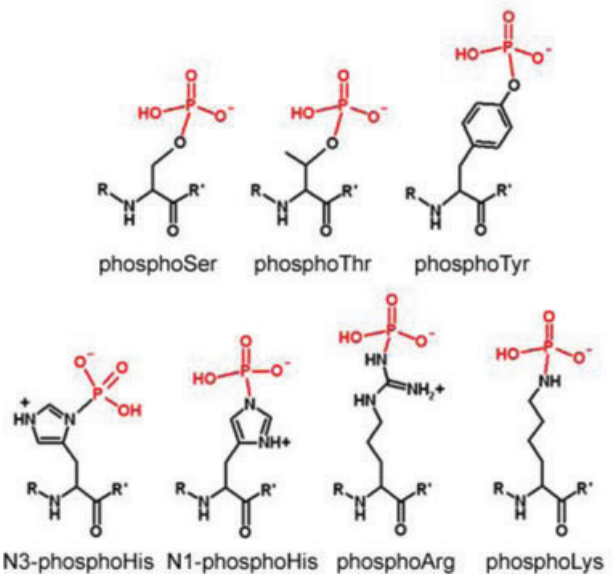

Figure 1. Oxygen- and nitrogen-bound protein phosphorylations. The upper row shows the phospho-esters of Ser, Thr, and Tyr that have been found as PTM in all kingdoms of life. *N*-bound protein phosphorylations of His (two isoforms), Arg, and Lys are displayed in the second row.

such analysis. Accordingly, additional software, detecting site-determining ions, must be employed in addition to the commonly used database search engines [6, 9, 10].

An important issue for misinterpretations arises from sequence-independent elimination of small molecules upon collisional activation [11]. Eliminations of water, carbon monoxide, or ammonia are well studied and generally not considered for sequence identification, since they occur in parallel to the intact fragment ions and often provide no relevant sequence information [12–14]. Recent studies on phosphoric acid (H_3PO_4) loss of phosphopeptides upon collisional activation reported that this reaction can occur via different mechanisms [11, 15, 16]. In case of peptides with pSer and pThr, loss of phosphoric acid (H_3PO_4) is usually explained by a charge-directed intramolecular S_N2 reaction involving the neighbouring *N*-terminal peptide bond [11, 16]. Furthermore, gas-phase rearrangements of the phospho-moiety onto an unmodified Ser or Thr sidechain have been reported upon CID, the commonly applied fragmentation technique for peptide sequence analysis [17, 18]. Those reactions were reported to occur with low frequency [18, 19]. As a result, phosphorylated peptide fragments, which do not comprise the original phosphorylation site, are detected, possibly leading to false phospho-site assignment, according to Palumbo and Reid [18]. In contrast, the studies by Aguiar et al. and Mischerikov et al. reported that the identification of the correct phosphorylation site is not impaired by the observed low levels of gas-phase rearrangements in large-scale experiments [19, 20]. False phospho-site assignments were only observed if low intensity signals were included for peptide identification [19]. pTyr is quite stable under CID fragmentation conditions, so that signals resulting from neutral loss processes usually have low intensity. Nevertheless, presence of acidic amino acids and arginine residues significantly increase the neutral loss of pTyr-containing peptides in MALDI-PSD fragmentation [21–23].

In bacteria, only very few proteins with phosphorylation of Ser, Thr, or Tyr have been reported and also the respective enzymes are not frequent in the genome [24]. However, phosphorylations of histidine (pHis), aspartate (pAsp), and arginine (pArg) (Fig. 1) play a major role as posttranslational modifications for signal transduction [25–27]. The properties of pHis or pArg bound to proteins and peptides, such as chemical stability and fragmentation behaviour, are barely understood. Phospho-His also undergoes a nominal elimination of H_3PO_4 upon thermal activation, despite the phospho-moiety being covalently attached to a nitrogen atom in the amino acid side chain [28]. Previously this reaction was explained by simultaneous loss of HPO_3 from the phosphorylated residues and a water molecule from a different amino acid [28]. Contrarily, a gas-phase rearrangement involving transfer of the phosphorylation onto an OH-moiety and thereby forming an isobaric intermediate followed by elimination of H_3PO_4 would also explain the observed neutral losses [29–31]. OH-groups that may serve as phospho-acceptors are present in the side chains of Ser, Thr, Tyr, the carboxylates of Asp, Glu as well as the free peptide *C*-terminus. Further protonated amide bonds can also act as phospho-acceptor sites [28, 29]. Upon H_3PO_4 loss, the unmodified His sidechain is restored and the information of the originally phosphorylated amino acid is lost [31].

Two novel MS/MS techniques, electron capture dissociation (ECD) and electron transfer dissociation (ETD), have been evaluated for phosphopeptide analysis [32, 33]. Both processes use thermal-energy electrons to dissociate peptides. The widely accepted theory of the fragmentation describes that the N-Cα bond is released in a radical mechanism, thereby forming c and z-type fragment ions [32]. CID-labile posttranslational modifications such as pSer/pThr remain stably attached during this fragmentation, which allows their exact localization within the peptide sequence from the resulting fragmentation spectrum [32, 33].

Our recent discovery of the first protein arginine kinase in *Geobacillus stearothermophilus* and upcoming studies to identify its targets in vivo, raise the question of how reliable is the identification and localization of pArg by database search engines and the additional phospho-site localization software, and consequently how well can they distinguish between pArg and other phosphorylations [25, 26]. First, we used the purified protein arginine kinase McsB to phosphorylate model proteins CtsR, ClpC, and McsB (by autophosphorylation) as well as synthetic oligopeptides on arginine. Employing different proteases, we prepared mixtures of arginine-phosphorylated peptides to evaluate the influence of

© 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

the fragmentation technique on phospho-site localization by database search engines and phospho-site localization tools in positive ionization mode [10, 34] (www.matrixscience.org). We further characterize the observed H_3PO_4 elimination with regard to peptide properties and amino acid composition employing high-resolution MS/MS.

2 Materials and methods

All reagents were from Sigma-Aldrich (Austria) in the highest available purity. 2,5-dihydroxybenzoic acid was from Bruker (Austria) in highest available purity. LC solvents were from Merck (Darmstadt, Germany) and VWR (Austria) in chromatographic purity and water was MilliQ grade purified by a Millipore system (Austria).

2.1 Model protein phosphorylation

McsB and CtsR (both *G. stearothermophilus*) and ClpC from *Bacillus subtilis* were expressed in His-tagged version in *Escherichia coli* cells and purified via Ni-affinity chromatography. Enzymatic phosphorylation and sample preparation for MS was carried out as described previously [26]. Trypsin (Promega), chymotrypsin, GluC (both Sigma), LysC (Wako), and LysN (Heck Laboratory, Utrecht) were applied as proteases to generate a variety of phosphopeptides with different properties. Enzymatic cleavage was stopped by freezing the sample after 4 h (chymotrypsin) or 14 h incubation at 37°C.

2.2 Phosphorylation of arginine and histidine and peptide modification

All peptides in Table 1, carrying the *N*-bound phosphorylation, were synthesized without the phospho-moiety on a peptide synthesizer (MultiSynTech, Germany) following standard Fmoc chemistry using specific resins for peptides with amidated *C*-terminus. The pArg was established via enzymatic phosphorylation with the purified protein arginine kinase McsB [26]. Histidine phosphorylation was achieved by chemical reaction of the peptide with phosphoric acid amidate (H_4PNO_3) at pH 7.5 in 20 mM HEPES [35]. H_4PNO_3 was synthesized from $POCl_3$ and NH_3 as described by Wei and Matthews and used without further purification [35]. Peptides with phospho-esters of Ser, Thr, or Tyr were synthesized with intact phosphorylation using the respective phospho-amino acid building block. All synthetic peptides were HPLC purified to more than 95%. Peptides with methylester modification of the COOH moieties were chemically modified with methanol and acetyl chloride following the protocol from Ficarro et al. [4].

Table 1. Synthetic phosphopeptides

Peptides	Sequence	Stable modification	
		N-term	*C*-term
PEP_01	K**pR**GGGGYIKIIKV	Acetyl	Free
PEP_02	K**pR**GGGGYIKIIKV	Acetyl	Methyl
PEP_03	K**pR**GGGGYIKIIKV	Acetyl	Amide
PEP_04	K**pH**GGGGYIKIIKV	Acetyl	Free
PEP_05	K**pS**GGGGYIKIIKV	Acetyl	Free
PEP_06	K**pR**GGGGFIKIIKV	Acetyl	Free
PEP_07	K**pR**GGGGFIKIIKV	Acetyl	Amide
PEP_08	YIVESK**pR**GGGGYIKIIK	Free	Free
PEP_09	YIVESK**pR**GGGGY	Free	Free
PEP_10	KAQPL**pR**AKFPAQPAK	Free	Free

S and Y phosphorylated peptides were synthesized with the intact phosphorylation; peptide with pR were enzymatically phosphorylated using purified protein arginine kinase McsB and H was modified with H_4PNO_3.

2.3 Peptide purification

All synthetic peptides were isolated from impurities by RP-HPLC at pH 9.0 on an inert Merck-Hitachi HPLC system equipped with a L6210 pump, D-6000 interface, L-4500 diode array detector (all Merck, Germany), using a Gemini NX column (15 cm × 4.2 mm, Phenomenex, Germany). A 10 mM NH_4HCO_3 buffer system without ACN in buffer A and 80% ACN in buffer B was used for HPLC purification. The applied gradient consisted of a prewash with 5% B for 10 min, a short gradient to 30% B in 7 min followed by a shallow gradient from 30%B to 60%B in 18 min to separate phosphorylated and unmodified peptide.

2.4 Fragmentation conditions

For ESI, MS/MS fragmentation studies were carried out on an LTQ-Orbitrap XL and an LTQ-Velos Orbitrap equipped with ETD reagent source (Thermo-Fisher, Bremen). MALDI was performed on a 4800 MALDI-TOF/TOF Analyzer (ABSciex, Darmstadt). MALDI conditions are described into detail in the corresponding Supporting Information section.

For LC-nanoESI-MS/MS-analysis of model protein digests, resulting peptide mixtures were separated on an Ultimate 3000 beta HPLC system (Dionex, Idstein) equipped with a C18 column (Dionex, AcclaimPepMap, 25 cm × 75 μm × 3 μm, 100 Å). A 15 min gradient from 0% to 100% B (30% ACN, 0.08% formic acid (FA), solvent A: 5% ACN, 0.1% FA) followed by high organic wash (C; 80% ACN, 10% trifluorethanol, 0.08% FA) was applied. The HPLC outlet was directly coupled to a nanoESI-source (Proxeon, Denmark) providing ionization of the peptides for MS detection in an LTQ OrbitrapXL. Detection and fragmentation of peptides of the CtsR/McsB sample was achieved by using a data-dependent method including one survey scan and five tandem MS scans with either CID, multistage-activation CID (MSA-CID),

© 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

higher-energy collision-induced dissociation (HCD), or ETD fragmentation. Peptides with charge 1+ were excluded from MS/MS analysis and doubly charged peptides were excluded for the ETD analysis only. The ClpC/McsB sample was analyzed employing MSA-CID- and ETD-MS/MS fragmentation in a data-dependent fragmentation method.

Synthetic peptides were diluted to 5 μM in 70% ACN, 0.1% FA, transferred to static nano-ESI-Emitters ES380 (Proxeon, Denmark) and directly infused into the LTQ-Velos Orbitrap mass spectrometer (Thermo-Fisher, Bremen) with a voltage of 800 V. Fragmentation experiments were carried out with the doubly and triply charged peptide ions. For low energy CID-MS/MS and MS^3 experiments the normalized collision energy (NCE) was set to 35% and the activation time to 30 ms, unless otherwise stated. For HCD fragmentation, NCE levels were adjusted to leave a precursor ion intensity of 15–30%. Prior to ETD experiments, the ETD source was tuned to optimal emission of fluoranthene anions and an activation time of 90–120 ms was adjusted. Fragment ion spectra were either analyzed in the ion trap for sensitivity reasons or in the Orbitrap mass analyzer at a resolution of 60 000 or 100 000 to allow unambiguous identification of fragment ion species. Hundred scans were accumulated for each average MS/MS spectrum.

2.5 Data interpretation

Orbitrap MS/MS spectra from direct infusion experiments were manually annotated using a mass accuracy of 0.015 Da, the maximal mass deviation for ion trap data was 0.4 Da. MALDI data were interpreted with a mass window of 0.1 Da. Fragment ions were calculated by the GPMAW software versus 7.0 (Lighthouse data, Denmark). Fragment ion intensities were normalized to the intensity of the isolated precursor at NCE 0%.

Enzymatic digests of McsB and CtsR were searched against a *B. subtilis* database containing McsB, CtsR and ClpC sequences of *G. stearothermopilus* (approximately 3500 different proteins) with the Protein Discoverer software versus 1.3 (Thermo-Fisher, Bremen) employing the Sequest, ZCore and Mascot search algorithm nodes [34, 36] (www.matrixscience.org). Identified phosphopeptides with a false discovery rate (FDR) of less than 5% were additionally manually evaluated. Search parameters were 10 ppm precursor mass accuracy, 0.6 Da fragment mass accuracy for ion trap data or a 0.025 Da mass for HCD-MS/MS data. Carbamido-methylation of Cys was selected as fixed modification, oxidation of methionine and phosphorylation of Ser, Thr, Tyr, and Arg as variable modifications. Phosphopeptide hits were extracted and the obtained result scores for each putative phospho-isoform were considered for the localization accuracy; peptides with only one putative phospho-site or without Ser, Thr, or Tyr were not considered. Peptide isoforms with different phospho-site locations were treated as individual hits when they exhibited baseline separation in the extracted ion chromatogram.

In addition, phosphopeptide hits from Mascot were reanalyzed by our in house phosphorylation site localization software phosphoRS [10]. This tool is implemented in the protein discoverer workflow and computes probabilities for individual phospho-sites within a peptide sequence independently from the Mascot peptide score. Prerequisite are the amino acid sequence of the best matching peptide and a peak list of the raw spectrum. Localization efficiency was calculated for all peptides identical with an FDR of 1%.

3 Results

3.1 Influence of phosphopeptide fragmentation on phospho-site localization

To produce a high number of chemically diverse arginine phosphorylated peptides, we performed proteolytic cleavage of in vitro phosphorylated model proteins using trypsin, chymotrypsin, LysC, LysN, and GluC. Analysis by LC-MS/MS was performed using CID, MSA-CID, HCD, and ETD fragmentation techniques for the same sample in individual experiments. MS/MS data were searched with the search algorithms Mascot and SEQUEST against the *B. subtilis* database within the Protein Discoverer software. ETD data were additionally analyzed with Zcore, a database search tool specifically designed for ETD fragmentation [36]. Different charge states were considered as unique phosphopeptide hits, since the difference in their MS/MS spectra also affects phosphorylation site localization. According to the phospho-site reported for the top hit, all identifications were grouped into three categories: "correct" pArg identification, "false" O-phosphorylation, and "ambiguous" assignment. Correct localizations represented the phosphopeptide sequence with pArg as rank 1 and a difference of either Mascot delta score >5, ΔXcorr >0.5, or ΔZCore probability >5 to the next O-phosphorylated isoform. These cut-off values should represent at least 75% localization probability and should provide a good measure for comparison of the different search algorithms. For false localizations, an O-phosphorylated isoform was rank 1 representing the same Δscore cut-offs to the correct arginine-phosphorylated isoform. Results from the phospho-site localization tool phosphoRS were catagorized by the individual site probabilities. To be correctly localized the site probability for pArg had to be 75% or higher, false localizations exhibited a $\geq$ 75% probability for pSer, pThr, or pTyr. Peptide hits with localization probabilities below 75% for any site were considered as ambiguous.

Figure 2 and Supporting Information Table S1 show the distribution of phospho-site assignments for combined phosphopeptide hits from the McsB/CtsR protein digests that were identified by Mascot and Sequest. Supporting Information Table S1 contains the list of identified phosphopeptides and detailed information on localization for each of the applied fragmentation techniques. For both CID

© 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

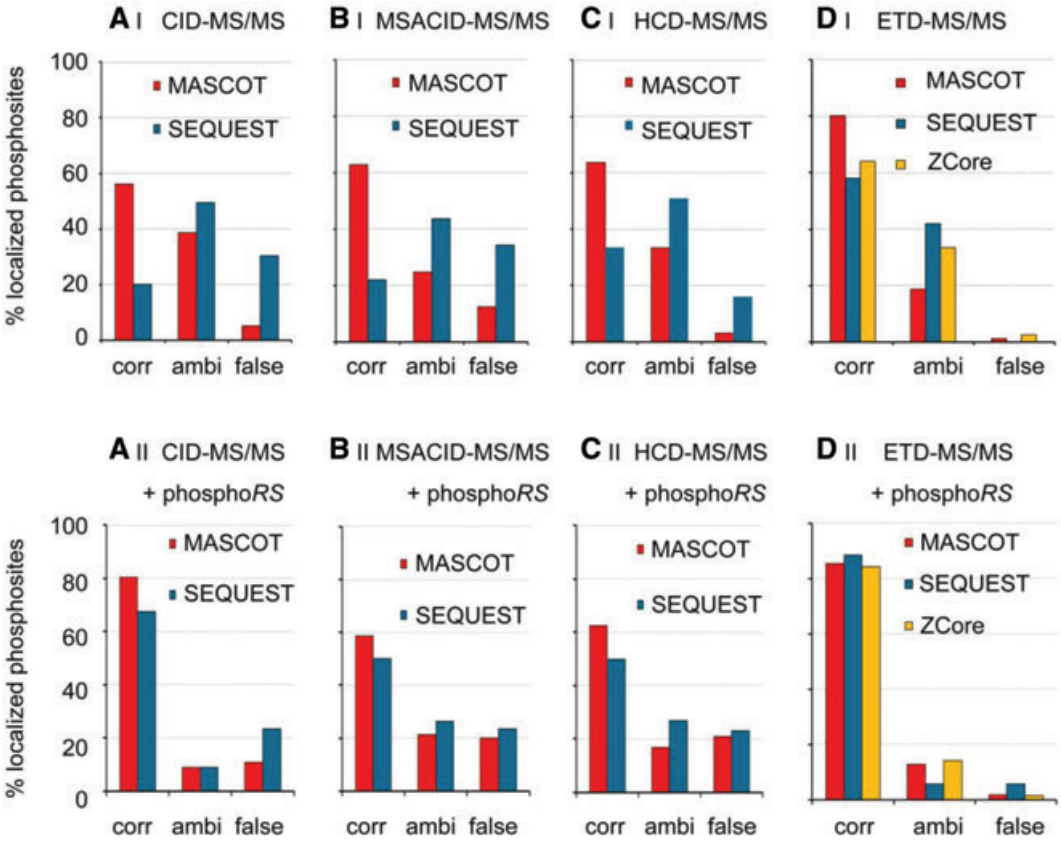

Figure 2. Probability of phospho-site localization from CID-, multi-stage activation-CID-, HCD-, and ETD-MS/MS data of arginine-phosphorylated peptides from McsB and CtsR. Phosphopeptides from digests with five different enzymes were analyzed by Mascot, SEQUEST, and ZCORE search algorithms within the protein discoverer software suite. Graphs A I to D I show the percentage of correct localizations of pArg by the respective search engine from MS/MS data generated by different fragmentation principles CID, MSA-CID, HCD, and ETD. Obviously, spectra from collision-based fragmentations have a lower success rate for correct pArg localization in comparison to ETD fragmentation. Site localization by the software tool phosphoRS after identification of phosphopeptides is shown in graphs A II to D II. Interestingly, single-stage CID reached up to 80% correct localizations with software-based phospho-site scoring while localization of MSA-CID and HCD data yielded many false results. Ambi stands for ambiguous localization.

datasets phospho-site localization by Mascot (overall, CID 63 unique peptides, MSA-CID 80 peptides) outperformed Sequest (CID 62 peptides, MSA-CID 101 peptides, Fig. 2AI and BI). Thereby, including neutral losses on fragment ions for all phospho-amino acids by Mascot, whereas Sequest allows neutral losses only for Ser, Thr, and Tyr, might partially contribute to the observed difference. Sequest reported more ambiguous or false localizations (35%) than Mascot (~5%). Although a higher percentage of correct identifications were reported by Mascot for MSA-CID data in comparison to normal CID fragmentation, the rate of false assignments was also doubled. Analysis of the second sample containing the chaperone ClpC together with the protein arginine kinase McsB confirmed the poor localization efficiency observed for MSA-CID fragmentation (Supporting Information Fig. S1). Interestingly, localization by the phospho-site localization software tool phosphoRS resulted in a much higher percentage of correct localizations for CID than for MSA-CID data (Fig. 2AII and BII) [10]. Elsholz et al. reported very good localization results from MSA-CID fragmentation on an Orbitrap using MaxQuant as for localization of pArg [25]. A closer look at the spectra and the applied acquisition method revealed that instead of fragmentation of the neutral loss product precursor of doubly and triply charged peptide ions at M-49 Da and M-32.7 Da, respectively, multistage activation was applied at M-98 Da and M-80 Da leading to MS/MS spectra which are closer to normal CID fragmentation. Therefore, good results for site localization were achieved; nevertheless, some of the identified Ser-, Thr-, or Tyr- phosphorylation sites might indeed be arginine phosphorylations, since the localization results were not further questioned.

HCD (Fig. 2CI) resulted in the lowest number of identified phosphopeptides (Mascot 37, Sequest 42) that might be due to the lower sensitivity of this technique when compared to CID. Results for phospho-site assignment were comparable to that of low-energy CID fragmentation in the ion trap, although HCD data are acquired with high resolution and mass accuracy. Software assisted phospho-site localization

© 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

only marginally improved the localization data obtained from the search algorithm (Fig. 2CII).

ETD generated data (Fig. 2DI) led to the highest percentage of correct assignments (Mascot 80 hits; Sequest 53; ZCore 73) for all applied search algorithms and had almost no false site localizations (Mascot 1, ZCore 3, Fig. 2DI). Most ambiguous localizations of ETD data arise from peptides with a Ser, Thr, or Tyr residue next to the originally pArg in combination with a low number of identified fragment ions. All observed and phosphorylated fragment ions carried the intact modification, thereby providing unequivocal assignment of the pArg for the majority of peptides. Application of special localization software slightly increased the number of correctly localized phosphopeptides without adding false localizations (Fig. 2DII). Since peptide ions with charge 2+ were excluded from the analysis due to low fragmentation efficiency, the influence of low spectral quality was minimized. Similarly, ETD outperformed MSA-CID also for the second sample, identifying arginine phosphorylation of ClpC and McsB with higher confidence.

Interestingly, application of the phosphoRS software improved the localization results for normal CID fragmentation significantly. In contrast, localization of MSA-CID and HCD data did not improve for this dataset. Highest reliability for phospho-site localization was obtained for ETD data where neutral losses were not present.

Next, we determined the influence of the charge state on phospho-site localization. Therefore, we counted the number of correct, ambiguous, and false localizations for charge states 2+, 3+, and higher (Supporting Information Table S2). As expected, CID-spectra generated in the ion trap provided the highest identification and localization rate for 2+ charged peptides, whereas higher charge states had a slightly worse localization rate for the CtsR/McsB sample but showed better localization in the ClpC/McsB sample. Only small number of ETD spectra for 2+ charged ions were identified in the CtsR/McsB dataset since this charge state was excluded from MS/MS analysis, due to the absence of many fragment ions, these spectra had the lowest localization success (Supporting Information Table S2). Phospho-Arg was localized with highest confidence in more than 80% of all spectra resulting from 3+ and higher charged peptides. This result was also observed for the second digest sample, where also 2+ charged ions were allowed for ETD analysis. Interestingly, those also reached a localization rate of 70% for pArg (Supporting Information Table S2).

Based on these results, it seemed that data for thermal activation-based fragmentation techniques (CID and HCD) possess a higher tendency to produce false localizations. Therefore, we assumed that the fragmentation products upon elimination of H_3PO_4 influence phospho-site determination by database search engines. A dominant loss of H_3PO_4 following CID was observed for 85% of the phosphorylated peptides (Supporting Information Fig. S2). In contrast, loss of the phospho-moiety (HPO_3) alone was almost not observed, but loss of water additionally to H_3PO_4 (–116 Da, M-H_5PO_5) occurred regularly. Under standard CID conditions small peptides revealed a higher tendency to form fragment ions with intact phosphorylation, while large peptides almost exclusively eliminated H_3PO_4. Furthermore, peptides with additional basic residues revealed a higher tendency for loss of H_3PO_4. Our data are in agreement with the reported MS/MS spectra of Elsholz et al. that also exhibit signals for H_3PO_4 elimination with rather high intensities for medium and large peptides [25].

3.2 H_3PO_4-elimination of pArg peptides and observed fragment ions

In order to determine the source of the poor localization efficiency, we decided to have a closer look at the generated fragment ions. To obtain pure MS/MS spectra of a single peptide species, we prepared arginine-phosphorylated model peptides for offline MS/MS analysis by CID, HCD, and ETD fragmentation (Table 1). CID-MS/MS of 2+-charged peptide ions with free *C*-terminus (PEP 01, 06, 08, and 10) resulted in dominant loss of H_3PO_4 while backbone fragments had less than 5% intensity of the M-H_3PO_4 signal as shown for peptide PEP_01 in Fig. 3A. The intensity of backbone fragments was significantly higher for peptides with a methyl esterified or amide-modified C-terminus (Supporting Information Fig. S3A). Nevertheless, fragmentation products generated by the elimination of H_3PO_4 were still the most abundant species. Apart from the precursor ion also fragment ions with neutral loss were identified (Fig. 3A). We could determine multiple b-ions comprising the putative pArg with an unmodified guanidyl group (b-80 Da) instead of a phosphorylation, although the corresponding ETD-MS/MS spectrum showed only phosphorylation of Arg and not a mixture of different peptides (Supporting Information Fig. S4). Those fragments demonstrate a cleavage of the N–P bond of the pArg and were found with high intensity. In contrast, the loss of HPO_3 from the precursor (M-80 Da) was almost absent in this spectrum. Elimination of H_3PO_4 from b-ions (b-98 Da) was also almost absent. Our observations were confirmed by MS^3-analysis of the dephosphorylated peptide ion, showing a series of b-80 Da ions as most abundant ions (Fig. 3B and C). In the MS^3-spectrum, all observed y-ions revealed a low-mass shift by 18 Da, pinpointing the site for water elimination to the peptide *C*-terminus (Fig. 3B and D). Since water elimination has only minor abundance for fragmentation of the corresponding unphosphorylated model peptide, this fragmentation pathway has to be favoured for the phosphorylated peptide. Furthermore, we did not observe individual losses of HPO_3 or water for the pArg peptides using energy resolved CID fragmentation (Supporting Information Fig. S5). Comparison of the fragmentation products of PEP_04 and 05, bearing pHis and pSer side chains instead of the pArg in PEP_01, respectively, revealed similar neutral loss behavior only for the pHis peptide (Supporting Information Fig. S6). CID-MS^3 spectra of pHis and pArg exhibited b-80 Da ions

© 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

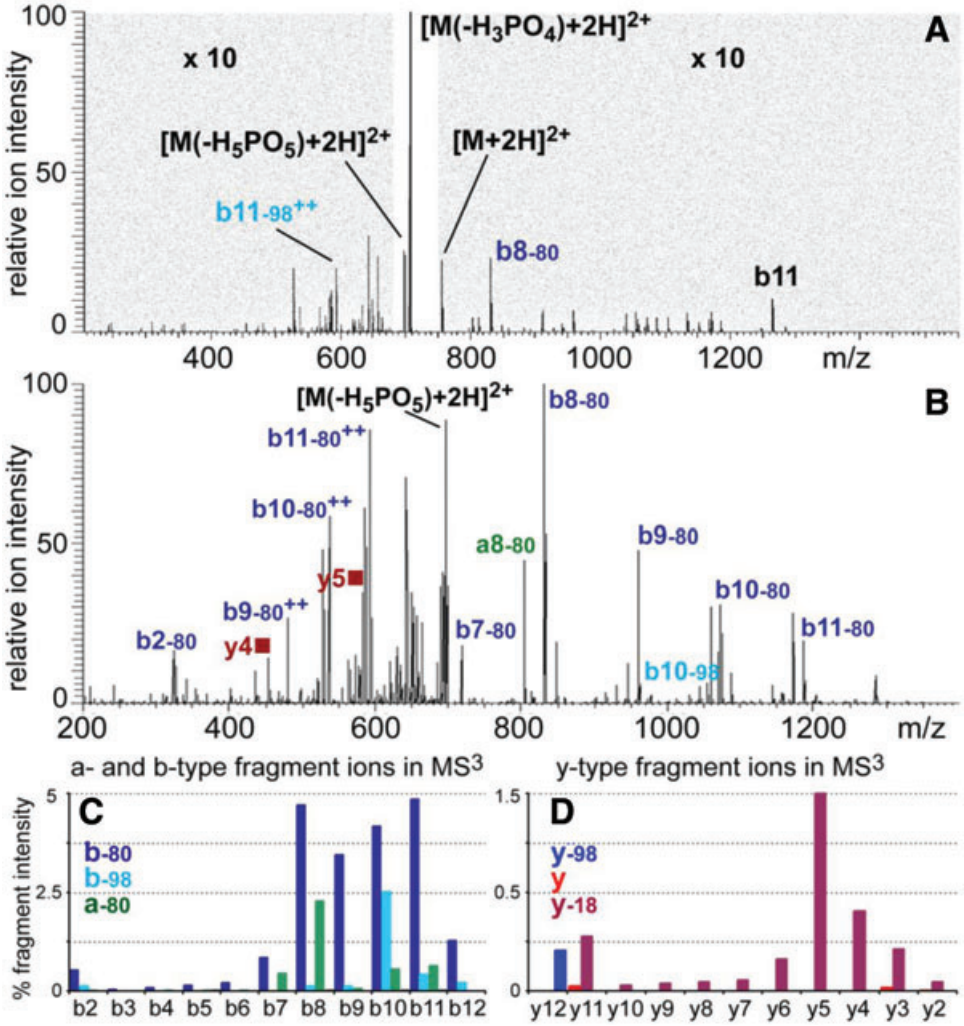

Figure 3. CID-MS/MS and -MS^3 of an arginine-phosphorylated model peptide. (A) The high-resolution CID-MS/MS spectrum of PEP_01 exhibits a dominant loss of H_3PO_4 (M-97.977 Da) and to a minor extent a combined loss of H_3PO_4 and H_2O (H_5PO_5, M-116 Da). Backbone fragments have only low intensity (gray boxes: intensity × 10). (B) High-resolution CID-MS^3 spectrum of PEP_01. The majority of b-ions exhibits loss of HPO_3 (79.9663 Da, b-80), while y-ions are exclusively found with loss of H_2O (y■). B-ions with H_3PO_4-loss (b-98) are present at low abundance. (C) The most intense fragment series of PEP_01 are b-HPO_3 ions (b-80) and corresponding a-HPO_3 ions (dark green). Only b10 and b11 have more intense signals for elimination of H_3PO_4. (D) The masses of all detected y-ions (y■) were reduced by 18 Da according to a water molecule (−18.0103 *m/z*), intact y-ions (red) were almost absent in the CID-MS^3 spectrum of PEP_01. Loss of H_3PO_4 was observed for y12 (y-98).

with high intensity and y-ions with water loss. In contrast, the serine phosphorylated peptide produced predominantly b-98 Da ions as expected from the respective charge-directed elimination mechanism.

Similarly, formation of "unmodified" b-ions was observed for many peptides upon CID, MSA-CID, and HCD activation from the initial LC-MS/MS dataset. These ions infer false phospho-site localizations when isoform sequences of the same peptide with Ser, Thr, or Tyr phosphorylation are explained by these fragments (Supporting Information Fig. S7).

3.3 Neutral loss via rearrangement of the phospho-moiety

So far, our experiments demonstrated that elimination of H_3PO_4 from arginine-phosphorylated peptides involves cleavage of two chemical bonds, the N–P bond of the phosphoguanidyl side chain, thereby reconstituting an unmodified arginine, and elimination of water from a different site in the sequence, predominantly from the *C*-terminus.

To study the reactivity of other OH-groups, a series of peptides (PEP_10–14) was designed containing one or no free OH-moiety as hydroxy or carboxy group in a peptide side chain per peptide. All peptides underwent elimination of H_3PO_4 as major fragmentation pathway. MS^3 analysis revealed that OH groups in Glu and Ser are involved in the neutral loss by elimination of water (Supporting Information Fig. S8). Moreover, the CID-MS/MS spectra of peptides 11, 13, and 14 revealed y6 and y4 ions with a positive mass shift of 79.966 Da indicating the formation of phosphoglutamate and phosphoserine on PEP_13 and 14, respectively (Fig. 4A and Supporting Information Fig. S9A). Identity of these in situ phosphorylated fragments was confirmed by MS^3 and MS^4 analysis (Fig. 4B and Supporting Information Fig. S9B). Detection of the phosphorylated y6 ion for PEP_11 is interesting, because the corresponding MS^3 spectrum allowed localization of the relocated phospho-moiety to the last three amino acids that have no OH-groups apart from

© 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

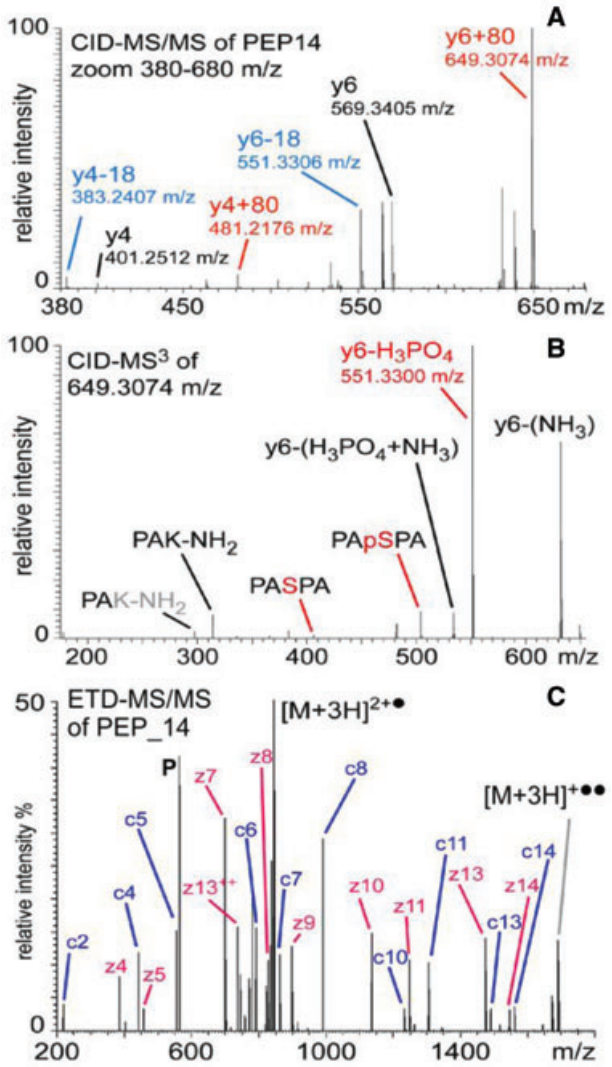

Figure 4. Rearrangement of the phosphorylation from Arg to Ser in CID-MS/MS of PEP_014. (A) The CID-MS/MS spectrum of PEP_14 (range 380–650 *m/z*) reveals newly phosphorylated y6 and y4 ions (red) without the original phospho-arginine. (B) The CID-MS^3 spectrum of the phosphorylated y6-ion verifies transfer of the phospho-moiety from R6 to S12. (C) ETD-MS/MS spectrum of PEP_14 corresponding fragments pointing to phosphorylation of Ser12 are absent in the ETD-MS/MS spectrum, confirming presence of only one precursor species with R6 phosphorylation. (P indicates the signal of the unfragmented triply charged precursor ion $[M+3H]^{3+}$).

protonated amide bonds. Subsequently, acquired ETD-MS/MS spectra of the same peptides did not show presence of mixed phosphopeptides populations but solely Arg6 phosphorylated precursors (Fig. 4C and Supporting Information Fig. S9C).

Fragmentation studies of peptides without free *C*-terminus revealed involvement of other side chain or intermediate OH-groups in the elimination of H_3PO_4 upon low energy collisional activation. Application of higher collision energies such as in beam type-CID (also HCD) generated multiple neutral loss species of the precursor and the fragment ions (Supporting Information Fig. S10). These conditions apply a high amount of energy within a short-time span and thereby also enable fragmentation pathways that are impaired upon low energy CID activation. Detection of fragment ions with loss of HPO_3 or H_3PO_4 and with intact phosphorylation in the same spectrum might cause false phospho-site localizations of HCD data, regardless of the high resolution and mass accuracy.

4 Discussion

Our study depicts, the difficulty of phospho-site localization of peptides with pArg—an important regulatory mechanism involved in heat-shock response in Gram-positive bacteria [25]. Peptides from the in vitro phosphorylated proteins were analyzed by LC-MS/MS and subsequent phosphopeptide identifications from database search engines and additional phospho-site scoring were compared to the correct phosphopeptide sequence. Thereby, CID or HCD had a significantly reduced success rate for localization of pArg in comparison to ETD (Fig. 2). Apart from spectral quality issues, the elimination of H_3PO_4 played a major role for false localizations. In the course of this elimination reaction, the N–P bond at the phosphoguanidyl group is cleaved with high frequency by all types of collisional activation, thereby generating fragment ions with an "unmodified" arginine side chain. For phospho-site assignment, these fragments point away from the arginine as phosphorylation site towards another amino acid. Simultaneously, water is released from an OH-group in the peptide, predominantly the *C*-terminus, giving rise to misinterpretation as putative H_3PO_4 elimination from pSer or pThr. Apart from the *C*-terminus also acidic amino acid side chains or Ser and Thr can participate in the H_3PO_4 elimination. Additionally, we hold direct evidence for a rearrangement of the phospho-moiety from the arginine sidechain onto Ser or Glu, when the *C*-terminus is not available. Since protease cleavage generates free peptide *C*-termini, rearrangement reactions are seldom observed for such samples and depend highly on the position of the phospho-site. Our data contribute to the understanding of MS/MS spectra of *N*-phosphorylated peptides generated by different fragmentation methods. Further, we explain the problem of false phospho-site assignment of pArg-peptides due to the specific neutral loss behaviour of such peptides upon collisional activation, involving losses from two different functional groups in the peptide ion.

The obtained data cannot unequivocally clarify the underlying reaction mechanism, a simultaneous elimination of HPO_3 and H_2O or rearrangement followed by H_3PO_4-elimination. Although reaction intermediates of the phospho-group rearrangement have been observed in CID-MS/MS spectra, it can not be excluded that the observed product ions are generated by a combination of both pathways.

© 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Fragmentation studies with low molecular weight phosphoarginine derivatives suggest that rearrangements are indeed the preferred pathway upon collisional activation [37]. We also propose an experimental outline for further large-scale studies for pArg in bacteria and other organisms, including ETD-fragmentation for reliable site localization.

This work was funded by Boehringer Ingelheim, the Christian Doppler Research Association, the Austrian Science Fund via the Special Research Programs Chromosome Dynamics (SFB-F3402) and P24685-B24 as well as the European Commission via the FP7 projects MeioSys and Prime XS. The authors thank Dr. J. Fuhrmann and Dr. T. Clausen from the IMP for sharing purified and phosphorylated proteins. Further, we acknowledge D. Broch Trentini and T. Köcher for critical comments on the manuscript and M. Madalinski for peptide synthesis. We also thank Prof. Dr. A. Heck for providing the LysC endoproteinase.

The authors have declared no conflict of interest.

5 References

[1] Hunter, T., Signaling—2000 and beyond. *Cell* 2000, *100*, 113–127.

[2] Cohen, P., The role of protein phosphorylation in neural and hormonal control of cellular activity. *Nature* 1982, *296*, 613–620.

[3] Manning, G., Whyte, D. B., Martinez, R., Hunter, T. et al., The protein kinase complement of the human genome. *Science* 2002, *298*, 1912–1934.

[4] Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., Burke, D. J. et al., Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat. Biotechnol.* 2002, *20*, 301–305.

[5] Huttlin, E. L., Jedrychowski, M. P., Elias, J. E., Goswami, T. et al., A tissue specific atlas of mouse protein phosphorylation and expression. *Cell* 2010, *143*, 1174–1189.

[6] Olsen, J. V., Blagoev, B., Gnad, F., Macek, B. et al., Global, in vivo, and site-specific phosphorylation dynamics in signalling networks. *Cell* 2006, *127*, 635–648.

[7] Pinkse, M. W. H., Uitto, P. M., Hilhorst, M. J., Ooms, B. et al., Selective isolation at the femtomol level of phosphopeptides from proteolytic digests using 2D-nanoLC-ESI-MS/MS and titanium oxide precolumns. *Anal. Chem.* 2004, *76*, 3935–3943.

[8] Schroeder, M. J., Shabanowitz, J., Schwartz, J. C., Hunt, D. F. et al., A neutral loss activation method for improved phosphopeptide sequence analysis by quadrupole ion trap mass spectrometry. *Anal. Chem.* 2004, *76*, 3590–3598.

[9] Beausoleil, S. A., Villen, J., Gerber, S. A., Rush, J. et al., A probability-based approach for high-throughput protein phosphorylation analysis. *Nat. Biotechnol.* 2006, *24*, 1285–1292.

[10] Taus, T., Koecher, T., Pichler, P., Paschke, C. et al., Universal and confident phosphorylation site localization using phosphoRS. *J. Proteome Res.* 2011, *10*, 5354–5362.

[11] Boersema, P. J., Mohammed, S., Heck, A. J. R., Phosphopeptide fragmentation and analysis by mass spectrometry. *J. Mass Spectrom.* 2009, *44*, 861–878.

[12] Ballard, K. D., Gaskell, S. J., Dehydration of peptide [M+H]+ ions in the gas phase. *J. Am. Soc. Mass Spectrom.* 1993, *4*, 477–481.

[13] Pingitore, F., Polce, M. J., Wang, P., Wesdemiotis, C. et al., Intramolecular condensation reactions in protonated dipeptides: carbon monoxide, water, and ammonia losses in competition. *J. Am. Soc. Mass Spectrom.* 2004, *15*, 1025–1038.

[14] Reid, G. E., Simpson, R., O'Hair, R. A. J., Leaving group and gas phase neighboring group effects in the side chain losses of protonated serine and its derivatives. *J. Am. Soc. Mass Spectrom.* 2000, *11*, 1047–1060.

[15] Gronert, S., Li, K. H., Horiuchi, M., Manipulating the fragmentation patterns of phosphopeptides via gas-phase boron derivatization: determining phosphorylation sites in peptides with multiple serines. *J. Am. Soc. Mass Spectrom.* 2005, *16*, 1905–1914.

[16] Palumbo, A. M., Jetze, J. T., Reid, G. E., Mechanistic insights into the multistage gas-phase fragmentation behavior of phosphoserine- and phosphothreonine-containing peptides. *J. Proteome Res.* 2008, *7*, 771–779.

[17] Froelich, J. M., Reid, G. E., The effect of post-translational and process-induced modifications on the multistage gas-phase fragmentation reactions of protonated peptides. *Comb. Chem. High Throughput Screen.* 2009, *12*, 175–184.

[18] Palumbo, A. M., Reid, G. E., Evaluation of gas-phase rearrangement and competing fragmentation reactions on protein phosphorylation site assignment using collision induced dissociation-MS/MS and MS^3. *Anal. Chem.* 2008, *80*, 9735–9747.

[19] Aguiar, M., Haas, W., Beausoleil, S. A., Rush, J. et al., Gas-phase rearrangements do not affect site localization reliability on phosphoproteomic data sets. *J. Proteome Res.* 2010, *9*, 3103–3107.

[20] Mischerikov, N., Altelaar, A. F. M., Navarro, J. D., Mohammed, S. et al., Comparative assessment of site assignments ion CID and electron transfer dissociation spectra of phosphopeptides discloses limited relocation of phosphate groups. *Mol. Cell. Proteomics* 2010, *9*, 2140–2148.

[21] Annan, R. S., Carr, S. A., Phosphopeptide analysis by matrix-assisted laser desorption time-of-flight mass spectrometry. *Anal. Chem.* 1996, *68*, 3413–3421.

[22] Metzger, S., Hoffmann, R., Studies on the dephosphorylation of phosphotyrosine-containing peptides during post-source decay in matrix-assisted laser desorption/ionization. *J. Mass Spectrom.* 2000, *35*, 1165–1177.

[23] Tholey, A., Reed, J., Lehmann, W. D., Electrospray tandem mass spectrometric studies of phosphopeptides and phosphopeptide analogues. *J. Mass Spectrom.* 1999, *34*, 117–123.

[24] Deutscher, J., Saier M. H. Jr., Ser/Thr/Tyr-phosphorylation in bacteria—for long time neglected, now well established. *J. Mol. Microbiol. Biotechnol.* 2005, *9*, 125–131.

[25] Elsholz, A. K., Turgay, K., Michalik, S., Hessling, B. et al., Global impact of protein arginine phosphorylation on the

© 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

physiology of *Bacillus subtilis*. *Proc. Natl. Acad. Sci. USA* 2012, *109*, 7451–7456.

[26] Fuhrmann, J., Schmidt, A., Spiess, S., Lehner, A. et al., McsB is a protein arginine kinase that phosphorylates and inhibits the heat-shock regulator CtsR. *Science* 2009, *324*, 1323–1327.

[27] Hoch, J. A., Two-component and phosphorelay signal transduction. *Curr. Opinion Microbiol.* 2000, *3*, 165–170.

[28] Medzihradszky, K. F., Phillipps, N. J., Senderowicz, L., Wang, P., Turck, C. W., Synthesis and characterization of histidine-phosphorylated peptides. *Protein Sci.* 1997, *6*, 1405–1411.

[29] Besant, P. G., Attwood, P. V., Detection and analysis of protein histidine phosphorylation. *Mol. Cell. Biochem.* 2009, *329*, 93–106.

[30] Kleinnijenhuis, A. J., Kjeldsen, F., Kallipolitis, B., Haselmann, K. F. et al., Analysis of histidine phosphorylation using tandem MS and ion-electron reactions. *Anal. Chem.* 2007, *79*, 7450–7456.

[31] Zu, X.-L., Besant, P. G., Axel, I., Attwood, P. V., Mass spectrometric analysis of protein histidine phosphorylation. *Amino Acids* 2007, *32*, 347–357.

[32] Stensballe, A., Jensen, O. N., Olsen, J. V., Haselmann, K. F. et al., Electron capture dissociation of singly and multiply phosphorylated peptides. *Rapid Comm. Mass Spectrom.* 2000, *14*, 1793–1800.

[33] Syka, J. E. P., Coon, J. J., Schroeder, M. J., Shabanowitz, J. et al., Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl. Acad. Sci. USA* 2004, *101*, 9528–9533.

[34] Ducret, A., Oostveen, I. v., Eng, J. K., Yates, J. R. III., Aebersold, R., High throughput protein characterization by automated reverse-phase chromatography/electrospray tandem mass spectrometry. *Protein Sci.* 1998, *7*, 706–719.

[35] Wei, Y.-F., Matthews, H. R., Identification of phosphohistidine in proteins and purification of protein-histidine kinases. *Methods Enzymol.* 1991, *200*, 388–414.

[36] Sadygov, R. G., Good, D. M., Swaney, D. L., Coon, J. J., A new probabilistic database search algorithm for ETD spectra. *J. Proteome Res.* 2009, *8*, 3198–3205.

[37] Gao, X., Ni, F., Bao, J., Liu, Y. et al., Formation of cyclic acylphosphoramidates in mass spectra of *N*-monoalkyloxyphosphoryl amino acids using electrospray ionization tandem mass spectrometry. *J. Mass Spectrom.* 2010, *45*, 779–787.

© 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim