An Isotope Labeling Strategy for Quantifying the Degree of Phosphorylation at Multiple Sites in Proteins

Adrian D. Hegeman, Amy C. Harms, and Michael R. Sussman
Biotechnology Center, University of Wisconsin–Madison, Madison, Wisconsin, USA

Anne E. Bunner and Jeffrey F. Harper
Department of Cell Biology, The Scripps Research Institute, La Jolla, California, USA

A procedure for determining the extent of phosphorylation at individual sites of multiply phosphorylated proteins was developed and applied to two polyphosphorylated proteins. The protocol, using simple chemical (Fischer methyl-esterification) and enzymatic (phosphatase) modification steps and an accessible isotopic labeling reagent (methyl alcohol-d_4), is described in detail. Site-specific phosphorylation stoichiometries are derived from the comparison of chemically identical but isotopically distinct peptide species analyzed by microspray liquid chomatography-mass spectrometry (μLC-MS) using a Micromass Q-TOF2 mass spectrometer. Ten phosphorylation sites were unambiguously identified in tryptic digests of both proteins, and phosphorylation stoichiometries were determined for eight of the ten sites using the isotope-coded strategy. The extent of phosphorylation was also estimated from the mass spectral peak areas for the phosphorylated and unmodified peptides, and these estimates, when compared with stoichiometries determined using the isotope-coded technique, differed only marginally (within ~20%). (J Am Soc Mass Spectrom 2004, 15, 647–653) © 2004 American Society for Mass Spectrometry

Protein phosphorylation is an omnipresent and important dynamic phenomenon in living systems that affects protein structure, protein-protein interactions and catalytic activity during physiological processes [1]. Often, protein phosphorylation occurs simultaneously at many different sites in a single protein [2, 3]. In these cases it is desirable to learn the extent to which each site is phosphorylated, to begin to allow correlation of phosphorylation status to function.

While mass spectrometric (MS) techniques can be used to great effect in locating and sequencing phosphopeptides derived from proteolytic digests of polyphosphorylated proteins, differences in ionization and detection efficiencies, from peptide to peptide, make MS based quantification problematic [4]. Numerous isotopic-coding methodologies have been described for comparative proteomic applications in the past few years (reviewed in [5, 6]). A technique utilizing differential protease-dependent incorporation of $^{18}O/^{16}O$, followed by phosphopeptide enrichment, phosphatase treatment, and MALDI-MS analysis has been employed to examine changes in protein phosphorylation [7]. Zhang and coworkers described using differential phosphatase treatment and isotope coding to measure phosphorylation stoichiometry with MALDI-MS analysis [8]. Here we describe an alternative approach, which combines differential phosphatase treatment with a one step isotope coding strategy, to yield phosphorylation site stoichiometries. This method is coupled to microspray liquid chromatography/tandem mass spectrometry (μLC-MS/MS) to characterize phosphorylation sites in two poly-phosphorylated proteins from the *Arabidopsis thaliana* calcium dependent protein kinase super-family, CPK1 and CRK3 [9].

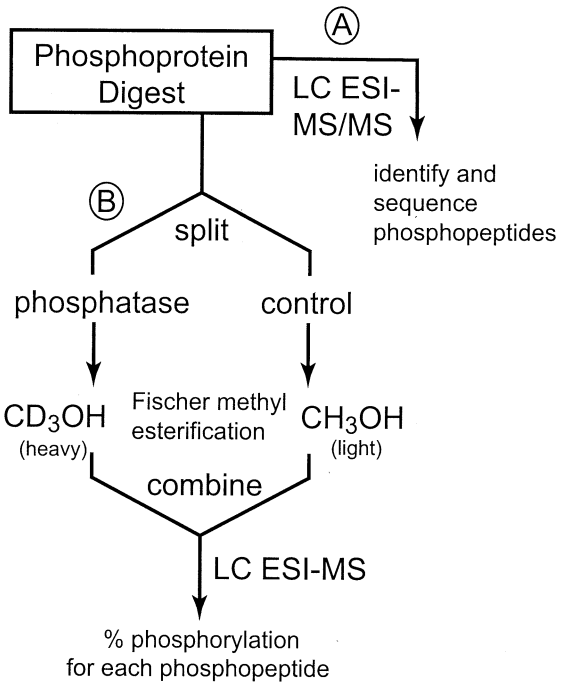
We have applied a simple isotopic-coding strategy to determine the phosphorylation stoichiometry of phosphopeptides derived from purified polyphosphorylated proteins. The two-part procedure is shown in Scheme **1**. First, a proteolyzed phosphoprotein is analyzed by μLC-MS/MS (Scheme **1a**) to identify and sequence phosphopeptides present in the digest. Second, either natural abundance or per-deuterated methanol is used in a bulk Fischer methyl-esterification of carboxylic acid residues of peptides in the digest to introduce an isotopic tag, which allows differentiation between phosphatase treated and untreated peptides (Scheme **1b**). Mass changes resulting from the derivati-

Published online February 27, 2004

Address reprint requests to Dr. A. D. Hegeman, Biotechnology Center, University of Wisconsin–Madison, 425 Henry Mall, Madison, WI 53706, USA. E-mail: hegeman@biochem.wisc.edu

1044-0305/04/$30.00
doi:10.1016/j.jasms.2003.12.019

Received July 9, 2003
Revised December 6, 2003
Accepted December 6, 2003



Scheme 1

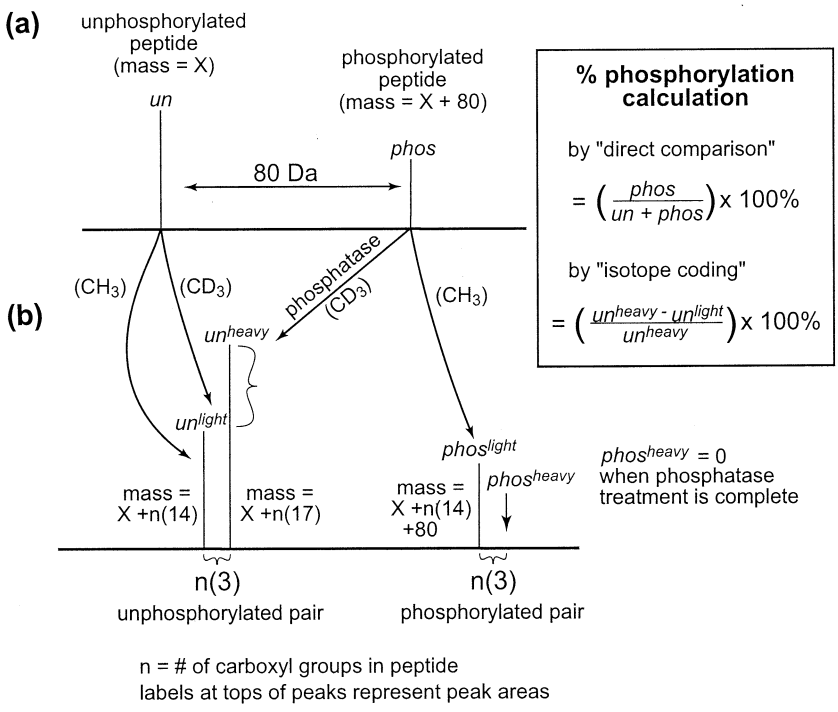
zation and enzymatic treatment are illustrated in Scheme **2**. Part A shows a phosphorylated and unphosphorylated peptide pair, and part B indicates the changes in mass that occur upon differential methylation and phosphatase treatments.

A simpler stoichiometry calculation may be made by using the areas of the phosphorylated and unphosphorylated peptides in Scheme **2a**, but one would need to assume that the two chemically distinct species have similar ionization and detection properties. Stoichiometries can be calculated from integrated heavy and light unphosphorylated peptide peaks in Scheme **2b** without making this assumption, by the comparison of chemically identical, but isotopically distinct (un^{light} and un^{heavy}) species. These measurements are used to test the validity of the assumption in the direct comparison of phosphorylated and unphosphorylated peptide peak volumes for underivatized samples (Scheme **2a**) as a method for estimating phosphorylation site stoichiometries.

Experimental

Protein Purification, Autophosphorylation, and Proteolysis

All reagents, except where otherwise noted, were obtained from Sigma-Aldrich (St. Louis, MO). *Arabidopsis thaliana* CRK3 (at2g46700) and CPK1 (at5g04870) were expressed in *Escherichia coli* (DH5α) as N-terminal glutathione S-transferase and C-terminal 6 × His fusion proteins to allow purification via the affinity sandwich method [10]. Briefly, bacterial lysate is passed over a nickel column (Probond, Invitrogen, Carlsbad, CA), washed, and eluted with imidazole. The eluent containing the over-expressed fusion protein is then further purified on a glutathione Sepharose 4B column, washed and eluted with glutathione. Proteins were concentrated and exchanged into storage buffer (20 mM TRIS-



Scheme 2

HCl pH 7.5, 50% glycerol, 5 mM EDTA, 100 μM NaCl, and 1 mM DTT, ~5 mM glutathione remains) by ultrafiltration and frozen prior to subsequent steps. Each fusion protein was allowed to autophosphorylate for 40 min at 25 °C in separate 1 mL reaction mixtures containing 40 μg of protein, 20 mM TRIS-HCl pH 7.5, 10 mM $MgCl_2$, 1 mM EGTA, 1.1 mM $CaCl_2$ and 300 μM ATP, additionally, 2–3% glycerol, 200–300 μM EDTA and 200–300 μM glutathione were carried over as components of the enzyme storage buffer (3% and 300 μM for CPK1 and 2% and 200 μM for CRK3). Protein was then precipitated by addition of 4 volumes of acetone and pelletted by centrifugation. The pellets were suspended each in 40 μL of 8 M urea, 1 mM DTT, and heated to 37 °C for 5 min prior to dilution in 280 μL of 50 mM ammonium bicarbonate pH 7.5, 1 mM DTT. Proteolysis was then initiated by addition of 0.8 μg of sequencing-grade modified trypsin (Promega, Madison, WI) and continued at 37 °C for 14 h. Half of the mixture was set aside for the phosphatase treatment/methyl esterification procedure (Scheme **1b**). The remaining portion of the digest was terminated by addition of formic acid to 0.5%, and peptides were purified by solid phase extraction using a C18-ZipTip (Millipore, Billerica, MA) for subsequent μLC-MS and μLC-MS/MS analyses.

μLC-MS and MS/MS

A portion of each purified digest, corresponding to 10 μg (120 pmoles) of pre-digested protein, was analyzed by μLC-MS/MS using a Micromass Q-TOF2 spectrometer to locate phosphorylation sites (Scheme **1a**). Chromatographic separation of peptides prior to mass spectral analysis was accomplished using C18 reverse phase HPLC columns made in-house from which eluted species are directly micro-electrosprayed [11]. Columns were made using lengths of fused silica tubing (365 μm o.d., 100 μm i.d.) with pulled tips (1 μm orifice) that were packed to 12 cm with Zorbax Eclipse XDB-C18, 5 μm, 300 Å pore size media. An Agilent 1100 series HPLC (Palo Alto, CA) delivered solvents A: 0.1% (vol/vol) formic acid in water, and B: 95% (vol/vol) acetonitrile, 0.1% (vol/vol) formic acid at either 1 μL/min for sample application, or 150–200 nL/min during a 200 min 2% (vol/vol) B to 70% (vol/vol) B gradient. Voltage was applied upstream of the column, by introduction of a platinum wire electrode into the fluid path via a PEEK T-junction. As peptides eluted from the HPLC-column/electrospray source, MS/MS spectra were collected over four channels from 400 to 1500 *m/z*; redundancy was limited by dynamic exclusion. Charge dependent collision energy profiles were empirically pre-determined. MS/MS data were converted to *pkl* file format using MassLynx version 3.5 (Micromass, Waters, Milford, MA). Resulting *pkl* files were used to search a modified *Arabidopsis thaliana* database containing both fusion-protein amino acid sequences using Mascot (Matrix Science, London, UK) with methionine oxidation and serine and threonine phosphorylation as variable modifications. Putative phosphopeptides identified by Mascot were confirmed and sequences were assigned to MS/MS spectra manually. The remaining 10 μg of each digest was analyzed by μLC-MS (same gradient and conditions as above) to allow integration of uninterrupted ion current for the direct comparison of phosphorylated and unphosphorylated peptide signals.

Phosphorylation Site Stoichiometry Determination

As outlined in Scheme **1b**, half of each autophosphorylated CPK1 and CRK3 fusion protein digest was split additionally into two tubes (each with 10 μg, 120 pmoles of digested protein). One of the two aliquots for each protein was treated with 10 units of calf intestinal alkaline phosphatase (New England Biolabs, Beverly, MA) with addition of the manufacturer's buffer, for 14 h at 37 °C while the other aliquot was not. Both aliquots for each protein were lyophilized and then differentially isotopically labeled by bulk Fischer esterification in either natural abundance- or deutero-methanol that is 2 M in anhydrous HCl. The esterification protocol is essentially that described by Ficarro and coworkers [12], where HCl is generated *in situ* by addition of acetylchloride to the methanol, on ice, with stirring, prior to its addition to the peptide mixture. The reaction was allowed to proceed for 2 h at room temperature at which point the methanol was removed under vacuum (Savant Speed Vac, GMI, Albertville, MN). As a modification of the Ficarro protocol, the sample was resuspended in methanol (using the same isotopic form as originally used) and evaporated again to remove as much HCl as possible. The samples were resuspended in methanol a second time, and sodium bicarbonate was added until CO_2 generation ceased, indicating neutralization of the HCl. (Removal of HCl is critical to prevent ester hydrolysis prior to addition of aqueous solvents.) Methanol was then removed and the residue was resuspended in 0.5% formic acid in water (formic acid should be added in excess of remaining bicarbonate to ensure sample acidification). Peptides were purified by sold phase extraction using C18-ZipTips (Millipore), and heavy and light methylated samples were combined prior to μLC-MS analysis as described above.

Data Analysis

Methylated and deuteromethylated peptide *m/z* values (for +1, +2, +3, and +4) were calculated for each phosphorylated and unphosphorylated peptide pair identified, and confirmed from the μLC-MS/MS analysis. The anticipated mass shifts resulting from phosphatase treatment and the isotopically coded (methyl) esterification are detailed in Scheme **2**. Signal intensities were determined by generating composite spectra dur-

Table 1. AtCPK1 and AtCRK3 phosphorylation sites and stoichiometry

Protein	Site in protein	Phosphopeptide	% Phosphorylation	
			Using isotopic-coding	Using direct comparison[b]
AtCPK1	S18	NGFLQSVpSAAMWRPR	none detected	90%
	T64	LSDEVQNKPPEQVpTMPKPGTDVETK	68 (±7)%[a]	79%
	S83	TEpSKPETLEEISLESKPETK	59 (±2)%	57%
	T107	SEpTKPESKPDPPAKPK	23 (±2)%	19%
	S137	TEpSVLQR	76 (±2)%	54%
	S151	KTENFKEFYpSLGR	62 (±3)%	63%
AtCRK3	S65	SSNPSPWFSPFPHGpSASPLPSGVSPAPAR	none detected	37%
	S132	GTEPEQSLDKpSFGYGK	80 (±3)%	100%
	T280, S281 or S282	DLKPENFLF*T*S*SR	60 (±7)%	24%
	T280, S281 or S282	DLKPENFLF*T*S*SREDSDLK	95 (±7)%	100%
	S311	LNDIVGpSAYYVAPEVLHR	59 (±3)%	63%
	S353	TEpSGIFR	96.8 (±0.3)%	93%

*The marked serine and threonine residues indicate ambiguity in the assignment of the phosphorylation site. These two peptides cover the same phosphorylated site in the intact AtCRK3 protein.
[a]This phosphopeptide was only partially dephosphorylated, and so the percentage listed was calculated using the combination of both phosphorylated and dephosphorylated pairs.
[b]Errors for percentages determined by direct comparison are at least 10%.

ing peptide elution, and transforming a region covering each charge state independently using the MaxEnt3 algorithm (Micromass) to deisotope, centroid and convert m/z to mass. Peak intensities were recorded for all observed calculated masses and isotopic pairs were corrected for mixing error as determined by the average ratio of 10 (CRK3) and 20 (CPK1) peptide pairs corresponding to unphosphorylated portions of each protein. The phosphorylation stoichiometry was calculated for each site using the corrected values for unphosphorylated peptide isotopic pair via the formula: % phosphorylation = $(un^{heavy} - un^{light})/un^{heavy} \times 100\%$. The percentages determined for each charge state were treated as independent measurements and averaged to generate the percentage and error values. The completeness of the phosphatase treatment was assessed for each site using the corrected values for phosphorylated peptide isotopic pair via the formula: % dephosphorylation = $phos^{light}/(phos^{heavy} + phos^{light}) \times 100\%$.

Results and Discussion

Two polyphosphorylated proteins, autophosphorylated CPK1 and CRK3, were trypsinized and subjected to μLC-MS/MS analysis resulting in the identification of six phosphopeptides in each protein. The exact phosphorylated residues were identified by manual assignment of MS/MS data for all six unique sites for CPK1 and four of five unique sites for CRK3. The peptide sequences and phosphorylation sites are shown in Table 1. Sequence coverage from the analysis of the trypsin digests of 10 μg (120 pmoles) of protein was 68% for both kinase constructs. Quality and extent of coverage decreases as the amount of analyte is reduced, with less than 40% coverage expected for 1 μg (approaching 10 pmoles) of the same sample providing a practical limit for sensitivity at this step. Once phosphorylation sites have been characterized, the sensitivity limits for detection of unfragmented derivatized and underivatized peptides are less restricted (realistically 1 pmole to 100 fmoles), as they do not rely on assignment of tandem mass spectral data, although we have not explicitly tested these limits in this initial study. Tandem MS data were obtained for each phosphopeptide identified and the fragmentation spectra were manually assigned (data not shown); in two of the peptides the precise location of the phosphorylation was ambiguous.

Phosphorylation stoichiometry was determined for five of the six peptides from each protein using the isotopic-coding technique described herein. Percent phosphorylation is listed for each site in the table under the "Using isotopic-coding" heading. The derivatization protocol apparently altered the physical properties of the two unanalyzed peptides so that they could not be detected under the analysis conditions. Figure 1 shows mass spectra of a CRK3 tryptic peptide (LNDIVGpSAYYVAPEVLHR) without (Figure 1a) and with (Figure 1b) the application of the isotopic-coding technique. Part A shows the underivatized phosphorylated and unphosphorylated peptide peaks. For each case in this study the phosphorylated peptide eluted before the unphosphorylated species. Figure 1b shows the same peptides after differential phosphatase treatment and derivatization. As shown in the inset elution profiles (Figure 1b), deuterium isotope effects observed with the unphosphorylated methyl esterified isotopic pairs, in all ten observed cases, resulted in the deuterated species eluting slightly prior to the proteated species. The absence of a heavy isotope species of the phosphorylated species (see arrow) shows that the phosphatase treatment was complete. This was the case for all of the peptides observed except for the LSDEVQNKPPEQVpTMPKPGTDVETK peptide for which dephos-

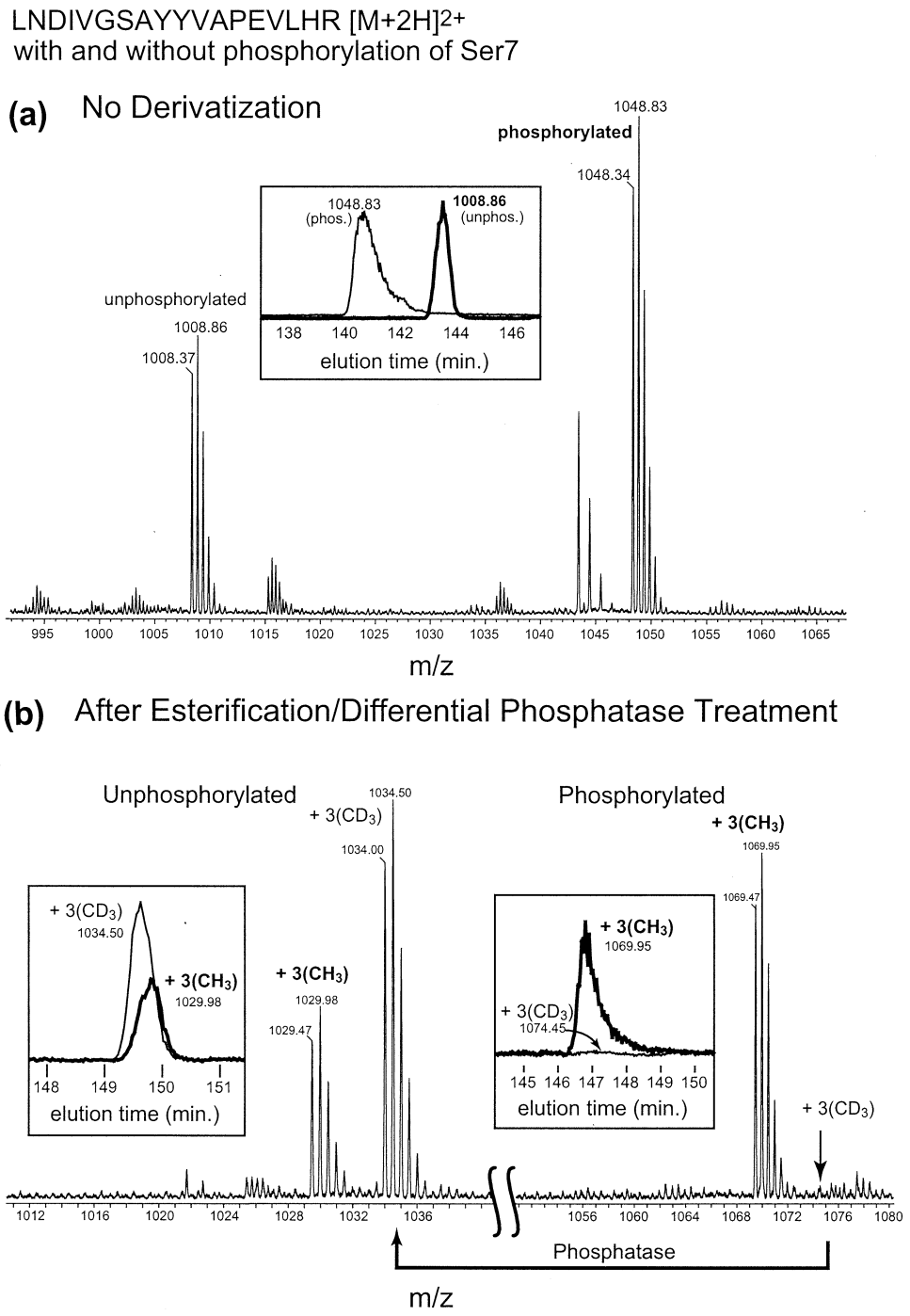


Figure 1. Two ways to determine the phosphorylation stoichiometry of a partially phosphorylated CRK3 tryptic peptide (LNDIVGpSAYYVAPEVLHR). (**a**) Direct comparison of peptide with and without phosphorylation. The spectrum is the sum of all spectra collected over both peaks in the inset selected-ion elution profile. (**b**) Differential phosphatase treatment followed by methyl-ester isotope coding. Methylation occurs at three carboxylic acid moieties in this peptide: the C-terminus; and two side chains [LND(CH_3)IVGpSAYYVAPE(CH_3)VLHR-(CH_3)]. Phosphorylation stoichiometry is determined from the difference in intensities of the unphosphorylated +3(CH_3) and +3(CD_3) derivatized peptide peaks. The modified peptides are isotopically distinct but chemically identical and have the same ionization and detection properties while being easily distinguished by mass difference. The phosphatase treatment has resulted in the complete conversion of the phosphorylated +3(CD_3) to unphosphorylated +3(CD_3) peptide as indicated by the arrow on the right side of the spectrum. Both spectra are the sum of all spectra over each inset selected-ion elution profile.

phorylation was only 54% complete. The value listed for this peptide in the table was derived by correcting for the measured percent dephosphorylation. This peptide was also only observed with a single methionine oxidation. The Fischer esterification was quantitative with no peaks assignable to partial derivatization products as was the case for all ten peptides observed using this method (data not shown). Complete removal of HCl is critical to prevent ester hydrolysis upon addition of aqueous solvent (see modifications to the procedure of Ficarro et al. [12] described in the Experimental section).

Two of the peptides with ambiguous phosphorylation sites overlap because of a missed cleavage. For both of these peptides, phosphorylation appears to occur at one of three adjacent residues in a "TSS" motif. As the digests were largely complete, it is likely that missed cleavage results from phosphorylation of the serine residue immediately prior to the arginine residue blocking proteolysis. A missed cleavage event will interfere with the ultimate goal of assigning a phosphorylation stoichiometry for a specific site in the protein. In these cases it is impossible to meaningfully combine the stoichiometries determined for the individual overlapping peptides to determine a stoichiometry for the site. Because of this, neither peptide is included in subsequent data analysis. All of the other stoichiometry measurements in this study were from unique and fully trypsinized peptides, which provided useful information about the site stoichiometry for the intact polypeptide chain.

If one assumes that the modification of a peptide by phosphorylation has no significant effect on its ionization and detection properties during MS analysis, it would then be reasonable to compare the two species directly to obtain phosphorylation stoichiometries [13]. To test this assumption, phosphorylated and unphosphorylated peptide peak areas were determined, and used to calculate the per cent phosphorylation for each phosphopeptide identified in the CPK1 and CRK3 digests. Phosphorylation stoichiometries determined using direct comparison are listed in the table under the "Using direct comparison" heading. When these measurements are compared to the values obtained using the isotopic coding strategy, (not including the two CRK3 peptides with ambiguous phosphorylation sites and partial proteolysis) all eight peptides were within 22% of the isotope coded values, and five were within 4%. No correlation was evident between the discrepancy in the two values and total ion current during phosphorylated and unphosphorylated peptide peak elution. Within the scope of this sample, the ionization and detection efficiencies are similar enough for a given phosphorylated and unphosphorylated species to allow the phosphorylation stoichiometry to be estimated to within ~20% of the values determined for chemically identical species.

Conclusions

An isotope-coding method for determining site specific phosphorylation stoichiometries for polyphosphorylated proteins has been described in detail and applied to two proteins. Tryptic digests of these proteins together contained twelve phosphopeptides, two of which could not be analyzed as the esterified peptides were not observed. While the 17% failure rate is a drawback for this technique, the relative ease of the procedure and availability (and affordability) of the isotopic and other reagents are an advantage. An even simpler approach, comparing the peak areas of a phosphopeptide and its unphosphorylated form, yielded stoichiometry estimates to within 20% of the isotope-coded results. This demonstrates that the direct comparison of phosphorylated and unphosphorylated peptide peak intensities is reasonable means of estimating phosphorylation stoichiometries. Differential phosphatase treatment followed by Fisher methyl esterification with CH_3OH or CD_3OH provides a simple technique for determining protein phosphorylation stoichiometries when more precision is required.

Acknowledgments

This work was supported by a grant from the National Science Foundation 2010 Project (MCB-0114769). The authors thank Drs. Catherine W. M. Chan, John C. Cushman, Alice C. Harmon, and Estelle M. Hrabak for helpful advice and valuable discussions.

References

1. Johnson, L. N.; Lewis, R. J. Structural Basis of Control by Phosphorylation. *Chem. Rev* **2001,** *101,* 2209–2242.
2. Chen, S. L.; Huddleston, M. J.; Shou, W.; Deshaies, R. J.; Annan, R. S.; Carr, S. A. Mass Spectrometry-Based Methods for Phosphorylation Site Mapping of Hyperphosphorylated Proteins Applied to Net1, a Regulator of Exit from Mitosis in Yeast. *Mol. Cell. Proteom* **2002,** *1,* 186–196.
3. Radimerski, T.; Mini, T.; Schneider, U.; Wettenhall, R. E. H.; Thomas, G.; Jenö, P. Identification of Insulin-Induced Sites of Ribosomal Protein S6 Phosphorylation in Drosophila melanogaster. *Biochemistry* **2000,** *39,* 5766–5774.
4. Mann, M.; Ong, S.-E.; Grønborg, M.; Steen, H.; Jensen, O. N.; Pandey, A. Analysis of Protein Phosphorylation Using Mass Spectrometry: Deciphering the Phosphoproteome. *Trends Biotechnol* **2002,** *20,* 261–268.
5. Goshe, M.B.; Smith, R. D. Stable Isotope-Coded Proteomic Mass Spectrometry. *Curr. Opin. Chem. Biol* **2003,** *14,* 101–109.
6. Tao, W. A.; Aebersold, R. Advances in Quantitative Proteomics via Stable Isotope Tagging and Mass Spectrometry. *Curr. Opin. Chem. Biol* **2003,** *14,* 110–118.
7. Bonenfant, D.; Schmelzle, T.; Jacinto, E.; Crespo, J. L.; Mini, T.; Hall, M. N.; Jenoe, P. Quantitation of Changes in Protein Phosphorylation: A Simple Method Based on Stable Isotope Labeling and Mass Spectrometry. *Proc. Nat. Acad. Sci. U.S.A* **2003,** *100(3),* 880–885.
8. Zhang, X.; Jin, Q. K.; Carr, S. A.; Annan, R. S. N-Terminal Peptide Labeling Strategy for Incorporation of Isotopic Tags:

A Method for the Determination of Site-Specific Absolute Phosphorylation Stoichiometry. *Rapid Commun. Mass Spectrom* **2002,** *16,* 2325–2332.

9. Harmon, A. C.; Gribskov, M.; Gubrium, E.; Harper, J. F. The CDPK Family of Protein Kinases. *New Physiolog* **2001,** *151,* 175–183.
10. Harper, J. F.; Huang, J.-F.; Lloyd, S. J. Genetic Identification of an Autoinhibitor in CDPK, a Protein Kinase with a Calmodulin-Like Domain. *Biochemistry* **1994,** *33,* 7267–7277.
11. Gatlin, C. L.; Kleemann, G. R.; Hays, L. G.; Link, A. J.; Yates, J. R. Protein Identification at the Low Femptomole Level from Silver Stained Gels Using a New Fritless Electrospray Interface for Liquid Chromatography Microspray and Nanospray Mass Spectrometry. *Anal. Biochem* **1998,** *263,* 93–101.
12. Ficarro, S. B.; McCleland, M. L.; Stukenberg, P. T.; Burke, D. J.; Ross, M. M.; Shabanowitz, J.; Hunt, D. F.; White, F. M. Phosphoproteome Analysis by Mass Spectrometry and its Application to Saccharomyces cerevisiae. *Nat. Biotechnol* **2002,** *20,* 301–305.
13. Resing, K. A.; Ahn, N. G. Protein Phosphorylation Analysis by Electrospray Ionization-Mass Spectrometry. *Methods Enzymol* **1997,** *283,* 29–44.