



# A systematic approach to the analysis of protein phosphorylation

Huilin Zhou<sup>1,2</sup>, Julian D. Watts<sup>1,2</sup>, and Ruedi Aebersold<sup>1,2\*</sup>

Reversible protein phosphorylation has been known for some time to control a wide range of biological functions and activities<sup>1-3</sup>. Thus determination of the site(s) of protein phosphorylation has been an essential step in the analysis of the control of many biological systems. However, direct determination of individual phosphorylation sites occurring on phosphoproteins *in vivo* has been difficult to date, typically requiring the purification to homogeneity of the phosphoprotein of interest before analysis<sup>4-6</sup>. Thus, there has been a substantial need for a more rapid and general method for the analysis of protein phosphorylation in complex protein mixtures. Here we describe such an approach to protein phosphorylation analysis. It consists of three steps: (1) selective phosphopeptide isolation from a peptide mixture via a sequence of chemical reactions, (2) phosphopeptide analysis by automated liquid chromatography–tandem mass spectrometry (LC-MS/MS), and (3) identification of the phosphoprotein and the phosphorylated residue(s) by correlation of tandem mass spectrometric data with sequence databases. By utilizing various phosphoprotein standards and a whole yeast cell lysate, we demonstrate that the method is equally applicable to serine-, threonine- and tyrosine-phosphorylated proteins, and is capable of selectively isolating and identifying phosphopeptides present in a highly complex peptide mixture.

The sequence of chemical reactions for selectively isolating phosphopeptides from a peptide mixture consists of six steps (summarized in Fig. 1A). (1) Amino protection: Peptide amino groups are protected using *t*-butyl-dicarbonate (tBoc) chemistry<sup>7</sup> to eliminate the potential for intra- and intermolecular condensation in subsequent reactions. (2) Condensation reaction<sup>8</sup>: Carbodiimide catalyzes condensation reactions between the peptides and excess amine to form amide and phosphoramidate bonds at the carboxylate and phosphate bonds of the peptides, respectively. (3) Phosphate regeneration: Free phosphate groups are regenerated by brief acid hydrolysis of the phosphoramidate bonds, and excess amine is removed by washing the peptides on reverse-phase resin. (4) Condensation and reduction<sup>9</sup>: A carbodiimide-catalyzed condensation reaction attaches cystamine to the regenerated phosphate group(s). Reduction of the internal disulfide of cystamine next generates a free sulfhydryl group for every phosphate group of a phosphopeptide. (5) Solid-phase capture: Excess cystamine, reducing agent, and by-products are removed by washing the peptides on reverse-phase resin. The recovered phosphopeptides are attached to a solid phase by reacting the free sulfhydryl groups in the peptides with iodoacetyl groups immobilized on glass beads. (6) Phosphopeptide recovery: Following stringent washing of the resin, phosphopeptides are recovered by cleavage of phosphoramidate bonds using trifluo-

oroacetic acid (TFA) at a concentration that also removes the tBoc protection group, thus regenerating peptides with free amino and phosphate groups. The carboxylate groups, however, remain blocked from step 2.

To demonstrate this method, we analyzed a tryptic digest of phosphoprotein  $\beta$ -casein by microcapillary LC-MS/MS. As shown in Figure 1B, numerous peptides were present in the untreated  $\beta$ -casein digest. The peptide indicated in Figure 1B was a doubly charged ion at mass-to-charge ratio ( $m/z$ ) = 1031.6. When selected for fragmentation using collision-induced dissociation (CID) (Fig. 1D)<sup>10</sup>, its fragment ion spectrum exhibited mostly the  $y$ -ion series typical for low-energy peptide fragmentation and an additional major signal at  $m/z$  = 983.0, corresponding to loss of 98 Da due to loss of the  $H_3PO_4$  group from the parent ion<sup>11,12</sup>. Database searching of this CID spectrum identified a peptide with sequence FQS\*EEQQQTEDELQDK (asterisk \* denotes a phosphate group). The mass difference between the  $y_{13}$  and  $y_{14}$  ions corresponded to that of phosphoserine, confirming Ser50 of this protein as the known site of  $\beta$ -casein phosphorylation.

Subjecting the same  $\beta$ -casein digest to the phosphopeptide isolation procedure reduced sample complexity considerably, yielding only one important doubly charged peptide ion at  $m/z$  = 1182.5 (Fig. 1C). The CID spectrum of this peptide showed a clear fragment ion series and a major signal at  $m/z$  = 1133.6 due to the loss of  $H_3PO_4$  (Fig. 1E). Database searching of this spectrum identified the same peptide as that in Figure 1C. The increase in apparent mass for the same peptide (compare Fig. 1B with 1C, and 1D with 1E) is due to quantitative modification on all seven carboxylates (aspartic acid, glutamic acid, and the C terminus) with ethanolamine during the isolation procedure.

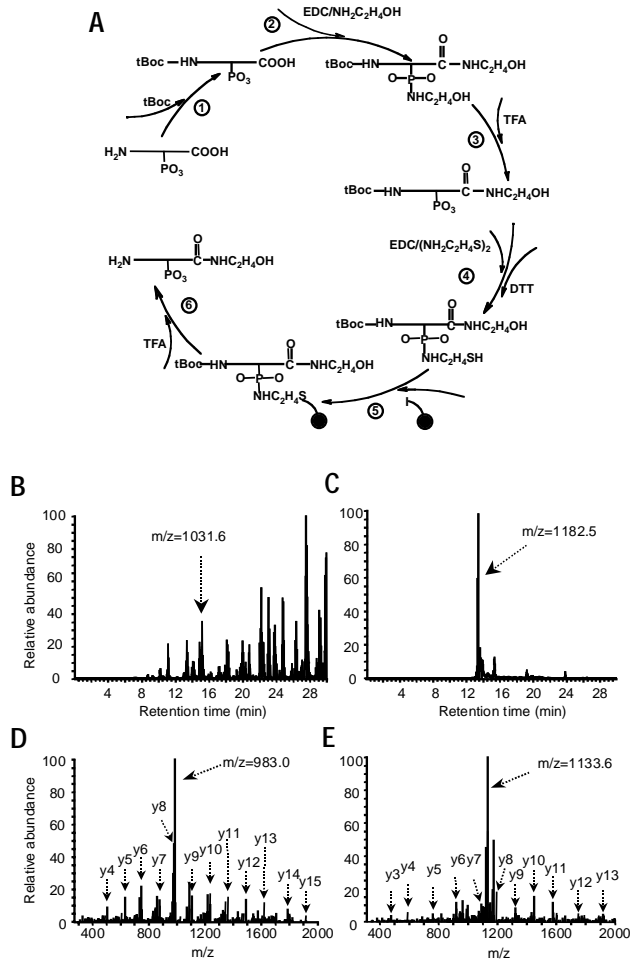
We next performed an experiment to determine sample recovery efficiencies for each step of the procedure and to demonstrate applicability to phosphotyrosine-containing peptides. To this end, we phosphorylated myelin basic protein (MBP) *in vitro* with the catalytic domain of the tyrosine kinase LCK, and radiolabeled ATP at a known specific activity. After trypsin digestion of the mixture, 5 pmol of phosphopeptides were isolated as before except that the carboxylate groups were blocked with *d*4-ethanolamine. The ion chromatogram for the isolated phosphopeptides is illustrated in Figure 2A. The most prominent ion at  $m/z$  = 630.1 (2+) was chosen for fragmentation, and its CID spectrum is shown in Figure 2B. This unambiguously identified the expected tyrosine-phosphorylated peptide TTHY\*GSLPQK from MBP (ref. 13). We determined phosphopeptide recovery efficiency over the entire six-step procedure by measuring radioactive counts recovered after each step of the procedure, with a final yield that consistently was ~20% of the starting material on multiple repetitions of this experiment (data not shown).

Several features of the method are immediately apparent. First, the isolation procedure yielded phosphopeptides essentially free of contamination from other peptides, and it is equally applicable to phosphoserine-, and phosphotyrosine-containing peptides (Figs 1 and 2). In addition, the CID spectra obtained from isolated phosphopeptides were of sufficient quality to allow unambiguous identification by sequence database searching. Also, only phosphopeptides with fully blocked carboxylates were recovered. Quantitative blocking of carboxylates would thus allow for incorporation of stable isotope tags that later can be differentiated and quantified by MS, much as we recently demonstrated the monitoring of protein expression levels by MS through isotopically tagging cysteine

<sup>1</sup>Department of Molecular Biotechnology, University of Washington, Seattle, WA 98195. <sup>2</sup>Current address: Institute for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA 98105. \*Corresponding author (raebersold@systemsbiology.org).



## TECH REPORT



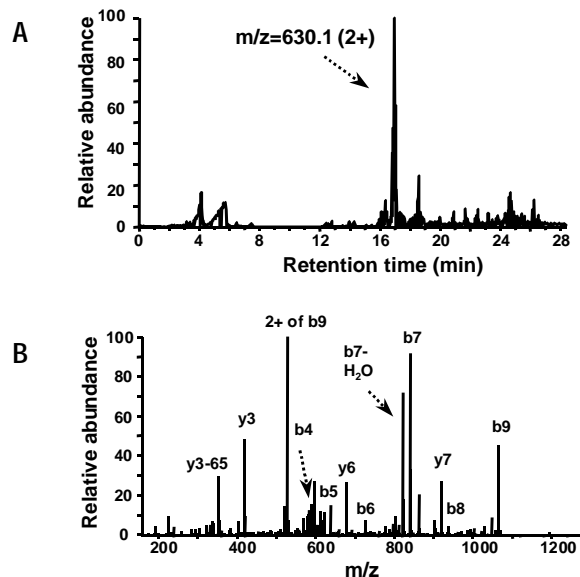
**Figure 1.** Phosphopeptide isolation strategy and validation with phosphoprotein β-casein. (A) Schematic illustration of the chemistry involved in selective phosphopeptide isolation. Full details are given in the text. (B–E) Phosphopeptide isolation from β-casein. A tryptic digest of β-casein was analyzed by LC-MS/MS both before (B, D) and after (C, E) phosphopeptide isolation according to the procedure in (A). The starting material for phosphopeptide isolation is 10 pmol. (B) Ion chromatogram of 1 pmol of β-casein digest before phosphopeptide isolation. The peak at  $m/z = 1031.6$  represents the doubly charged form of the expected tryptic phosphopeptide from β-casein. (C) Ion chromatogram of the isolated phosphopeptides of β-casein digest. The peak at  $m/z = 1182.5$  represents the doubly charged form of the same tryptic phosphopeptide from β-casein indicated in (B), additionally modified on its seven carboxylate groups with ethanolamine. (D) CID spectrum of β-casein digest in (B). The peak at  $m/z = 983.0$  represents the doubly charged form of the selected parent ion ( $m/z = 1031.6$ ) minus the  $H_3PO_4$  group. (E) CID spectrum isolated phosphopeptides of β-casein digest in (C). Again, the peak at  $m/z = 1133.6$  represents the doubly charged form of the selected parent ion ( $m/z = 1182.5$ ) minus  $H_3PO_4$ , and the y-ion series used for peptide identification is indicated. The b-ion series are much less intense and are omitted for clarity.

ent from the finding that >80% of the CID spectra that led to identification were derived from phosphopeptides. Additionally, CID spectra derived from the few nonphosphorylated peptides identified generally resulted from lower intensity precursor ions. Thus even with a highly complex starting material, only low levels of nonphosphorylated peptide background carried through the isolation procedure to the MS, affirming its selectivity. In the example shown, the ion at  $m/z = 1032.7$  in Figure 3B was selected for CID, with this spectrum shown in Figure 3C. In addition to a clear fragment ion series, a major signal corresponding to the doubly charged parent ion after undergoing the loss of  $H_3PO_4$  is apparent at  $m/z = 983.8$ . After database searching, the peptide was identified as being from enolase and having the indicated sequence. This peptide contained three potential threonine phosphorylation sites, and the parent ion mass indicated that the peptide contained a single phosphate group. The y7 to y13 ions confirmed that the phosphate was not on the N-terminal threonine. Two possible pairs of the y5 and y6 ions correspond to phosphorylation of either one of the other two threonine residues. Thus we could not definitively assign the exact site of phosphorylation for this peptide, neither can we preclude the possibility that a mixture of both possible monophosphorylated species of this peptide may have co-eluted from the LC column.

residues<sup>14</sup>. Finally, it is well known that phosphoserine and phosphothreonine readily lose the  $H_3PO_4$  group on CID<sup>11,12</sup>. The fragment ion that corresponds to the loss of 98 Da thus provides an immediate indication of a phosphoserine- or phosphothreonine-containing peptide. Phosphotyrosine is more stable and does not lose its phosphate group during fragmentation (see Fig. 2).

We next evaluated this procedure for the identification of multiple phosphopeptides present in a highly complex mixture. To this end, we chose to analyze a lysate of the yeast *Saccharomyces cerevisiae*. Phosphopeptides were isolated by this method and again analyzed by LC-MS/MS, with CID spectra recorded and searched against the yeast sequence database. Figure 3A shows the total ion intensity recorded with respect to retention time on the column, indicating the complexity of the sample. Figure 3B shows the  $m/z$  values obtained integrated over the time window delineated in Figure 3A. The major peptide peaks observed that also displayed loss of 98 Da during CID are labeled with an asterisk (\*), confirming that the majority of the peptides detected were indeed phosphorylated. Furthermore, the selectivity of the method was appar-

**Figure 2.** Phosphopeptide isolation from LCK-MBP kinase reaction mixture. (A) LC-MS ion chromatogram of the phosphopeptides isolated from a tryptic digest of protein mixture, generated from *in vitro* kinase reaction between LCK and MBP. (B) The most intense ion ( $m/z = 630.1$ , 2+ ion) was subjected to CID analysis and database searching, which identified this peptide as TTHY\*GSLPQK, with the tyrosine residue being phosphorylated.

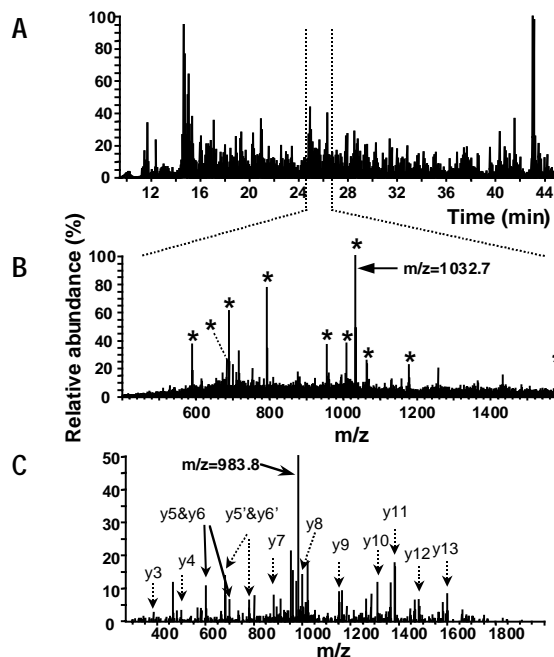




**Figure 3.** Phosphopeptide isolation from yeast cell lysate. (A) LC-MS chromatogram of the phosphopeptides isolated from a tryptic digest of whole yeast cell lysate. (B) Integrated  $m/z$  spectrum of ions eluting from the LC column with retention times between 24.7 and 26.5 min, as indicated in (A). Major ion peaks that additionally exhibited a loss of 98 Da on CID, indicating that they are phosphopeptides, are annotated with an asterisk (\*). (C) The CID spectrum recorded for the peptide peak indicated in (B) at  $m/z = 1032.7$  was sufficient to identify the phosphopeptide as TAGIQVADDLT\*VT\*NPAR from enolase. However, the exact site of threonine phosphorylation was not unambiguously defined because of the difficulty in assigning  $y_5$  and  $y_6$  ions. Both potential locations for the phosphate have thus been indicated (\*), although the parent ion mass confirms the peptide as a singly phosphorylated species.

Table 1 lists the proteins identified following database searching of the additional CID spectra obtained in the same manner, along with the sequences of the phosphopeptides determined. All peptides positively identified were singly phosphorylated species, and they were phosphorylated at serine or threonine residues. Table 1 also indicates the location of the phosphorylation site within the peptide when this could be unequivocally determined, or gives the possible phosphorylation sites in cases in which the observed CID data could not distinguish between two or more phosphorylation sites. As shown in Figure 3C, such occasions did not interfere with the identification of the phosphopeptide, and the site of phosphorylation could typically be confined to a cluster of hydroxyl amino acids.

We did not identify any peptides with multiple serine or threonine phosphorylation sites. In many cases, ions corresponding to loss of  $H_3PO_4$  dominated the fragmentation process, resulting in insufficient fragmentation at peptide bonds for sequencing. This effect would be compounded by multiple phosphoserine or phosphothreonine sites in a single peptide. Peptides that are too large or too small are generally not suitable for MS sequencing; besides, such peptides could be lost during the desalting steps in this



method. Whether this method could determine all the phosphorylation sites of a given protein therefore depends on whether phosphorylation sites are contained in peptides of suitable sizes and/or hydrophobicities for MS analysis, a limitation common to all MS-based methods. In such cases, alternative proteolytic enzymes can be considered. In the experiment on yeast lysate, tyrosine-phosphorylated peptides were not identified, probably because of their significantly lower abundance.

Most of the proteins identified were found to be glycolytic enzymes, including enolase, glyceraldehyde 3-phosphate dehydrogenase, phosphoglycerate kinase, and pyruvate kinase. Because the cells from which the proteins were isolated utilized glucose as a carbon source, the identification of phosphorylation sites on glycolytic enzymes as the major species present in this sample is perhaps not surprising. Phosphopeptides from other highly expressed proteins, such as pyruvate decarboxylase isozyme 1, ribosomal proteins, and heat shock proteins were also identified. Notably, most of the proteins listed in Table 1 were not annotated in the database as known phosphoproteins. However, previous studies from our group and others identified many of the proteins listed in Table 1 in multiple two-dimensional gel spots<sup>15,16</sup>, consistent with differentially phosphorylated forms of the same protein. These two-dimensional gel data are thus consistent with the identifications made here, supporting the assertion that these proteins are indeed phosphorylated *in vivo*. Although we did not identify phosphopeptides from lower-abundance regulatory proteins in this experiment, the method itself nonetheless is compatible with larger-scale sample preparation, or for analysis of enriched protein complexes of interest. Fractionation either before or after this isolation should greatly facilitate identification of low-abundance proteins.

**Table 1. Phosphopeptide profile from yeast grown with glucose as a carbon source**

Entry name <sup>a</sup>	Protein name	Phosphopeptide identified <sup>b</sup>
ENO1_yeast & ENO2_yeast	Enolase	TAGIQVADDLT*VT*NPAR <sup>c</sup> IGLDCAS*S*EFFK <sup>c</sup> SGET*EDT*FIADLVVGLR <sup>c</sup>
G3P1_yeast	Glyceraldehyde 3-phosphate dehydrogenase	LVSWYDNEYGYST*R <sup>c</sup> VIS*NASCTTNCLAPLAK VISNASCT*T*NCLAPLAK <sup>c</sup> TAGSNIIPSST*GAAK NPVILADACCS*R
DCP1_yeast	Pyruvate decarboxylase isozyme 1	TPANAAPVAPAS*T*PLK <sup>c</sup>
KPY1_yeast	Pyruvate kinase 1	GVNLPGTDVDPALPS*EK GVNLPGT*DVDLPALSEK
PGK_yeast	Phosphoglycerate kinase	DVT*FLNDCVGPVEAAVK VLENT*EIGDSIFDK EGIPAGWQGLDNGPES*R ASAPGS*VILLENLR ELPGVAFLS*EK
PGM1_yeast	Phosphoglycerate mutase 1	SFDVPPIDASSPFS*QK VYPDVLYT*S*K <sup>c</sup>
ALF_yeast	Aldolase	FAIPAINVT*S*S*S*AVAALAAAR <sup>c</sup>
G6PI_yeast	Glucose-6-phosphate isomerase	EANVT*GLR
HS75_yeast	Heat shock protein	SQIDEVVLVGGSS*TT*R <sup>c</sup>
HS72_yeast	Heat shock protein	TTPSFVGFDTT*ER
RL11_yeast	60S ribosomal protein	VLEQLSGQTPVQS*K
R141_yeast	40S ribosomal protein	IEDVTPVPS*DS*T*R <sup>c</sup>

<sup>a</sup>Entry names are according to Swiss-Prot nomenclature ([www.expasy.ch](http://www.expasy.ch)).

<sup>b</sup>Sequence and site of phosphorylation was identified by SEQUEST<sup>19</sup> (see text). Asterisks indicate site of phosphorylation at the serine or threonine residue to its left.

<sup>c</sup>Multiple asterisks within a sequence indicate ambiguity on the exact site of phosphorylation. All peptides are singly phosphorylated.



In conclusion, this method appears to be well suited to rapid, sensitive, and automated analysis of phosphoproteins on a large scale, an undertaking not practical in anything other than a purely descriptive sense until now. Using a reversible covalent linkage to capture phosphopeptides to a solid support resulted in a high degree of phosphopeptide enrichment before MS analyses. Because this method does not require metabolic  $^{32}\text{P}$  labeling, it is also amenable to the study of human tissues and tumors, considering the important role played by aberrant phosphorylation in oncogenesis and immune disorders. With its additional utility for describing quantitative changes in the phosphorylation state of proteins in different biological samples through the use of stable isotope tag, the method could thus enhance the traditional objectives of protein phosphorylation research and expand its scope in the near future.

### Experimental protocol

**In vitro kinase reaction.** A baculovirally expressed LCK kinase domain–glutathione *S*-transferase (GST) fusion protein (17  $\mu\text{g}$ ), 20  $\mu\text{g}$  MBP, and 10  $\mu\text{Ci}$  of  $^{32}\text{P}$ -containing ATP were incubated at 30°C for 1 h in 40  $\mu\text{l}$  of buffer containing 25 mM Tris (pH 7.5), 10 mM  $\text{MnCl}_2$ , 0.25 mM ATP. After 1 h, 6 M urea was added to stop the reaction. Reduction and alkylation was performed by the addition of dithiothreitol (DTT) to 10 mM for 30 min, followed by 2 h of incubation at 50 mM iodoacetamide. Samples were diluted with water to 1 ml, and 1  $\mu\text{g}$  of trypsin (Promega, Madison, WI) added for 4 h at 37°C. Peptides were then desalted on a  $\text{C}_{18}$  column (Waters Associates, Milford, MA, cat. no. WAT023590) and recovered by elution with 80% acetonitrile, 0.1% TFA. Recovered radiolabeled peptides were quantified by Cerenkov counting. From this sample, an estimated 5 pmol of phosphorylated peptides were taken for isolation of phosphopeptides and evaluation of recovery efficiencies. Deuterated d4-ethanolamine (Isotec, Miamisburg, OH) was used to block carboxylate groups in this experiment.

**Phosphoprotein profiling in yeast.** Yeast *S. cerevisiae* strain (BWG1-7A) was grown until mid-log phase on YPD (1% yeast extract and 2% peptone; DIFCO Laboratories, Detroit, MI) media with 2% glucose as a carbon source and harvested by centrifugation. Protein extract was prepared by the glass beads method as described<sup>17</sup>. A mixture of DNase 1 (20 U/ml) and RNase (10  $\mu\text{g}/\text{ml}$ ) was added for 30 min on ice (Worthington Biochemical, Lakewood, NJ). Protein concentration was determined using Bio-Rad protein assay (Bio-Rad Laboratories, Hercules, CA), and 500  $\mu\text{g}$  of the protein extract was then denatured in 0.1 M potassium phosphate buffer (pH 8.0) with 6 M urea. Proteins were reduced and alkylated as described above. Sample was then dialyzed before digestion overnight with trypsin at 37°C. The resulting peptide mixture was desalted by reverse-phase  $\text{C}_{18}$  column as described above.

**Phosphopeptide isolation procedure.** Peptide samples were dried, and then subjected to the method diagrammed in Figure 1A according to the following steps. Peptide mixture was resuspended in 50% (vol/vol) of 0.1 M phosphate buffer (pH 11)/acetonitrile. 0.1 M tBoc was added for 4 h at room temperature. Acetonitrile was then removed under reduced pressure. Samples were made to 1 M ethanolamine, 25 mM *N*-hydroxysuccinimide (NHS), and 0.5 M of *N,N'*-dimethylaminopropyl ethyl carbodiimide HCl (EDC) and incubated for 2 h at room temperature. Then, 10% TFA was added for 30 min at room temperature. Samples were then desalted on and recovered from a  $\text{C}_{18}$  column as described above. After the peptides were dried and redissolved in 1 M imidazole (pH 6.0), 0.5 M EDC was added for 3 h at room temperature. Samples were loaded onto a  $\text{C}_{18}$  column, washed, and treated with 1 M cystamine (pH 8.0) for 2 h at 50°C on the column. Peptides were washed with water and reduced with 10 mM DTT to generate free sulfhydryl groups. After washing to remove DTT, peptides were eluted with 80% acetonitrile, 0.1% TFA and incubated with 20 mg beads with immobilized iodoacetyl groups for at least 2 h at pH 8.0 (titrated with 1 M Tris pH 8.0, 50 mM ethylenediamine tetraacetic acid). Beads with immobilized iodoacetyl groups were prepared by a 2 h reaction between 3 equivalents of iodoacetic anhydride and 1 equivalent of amino beads (G4643; Sigma Chemical Co., St. Louis, MO) with 3.3 equivalents of diisopropylethylamine in dimethylformide. The formation of a tyrosine adduct with

carbodiimide is a possible side reaction<sup>18</sup>. Such an adduct is unstable against nucleophiles such as hydroxylamine. Therefore, after attachment of phosphopeptide to the beads, 1 M hydroxylamine (pH 10) was used to incubate beads for 2 h at room temperature. This restored tyrosine residues. Beads were then washed sequentially with 2 M NaCl, methanol, and water to remove nonspecifically bound molecules. Finally, the beads were incubated with 100% TFA for 30 min to recover phosphopeptides. Concurrently, tBoc protection was removed. The recovered sample was dried under reduced pressure and resuspended in water for LC-MS/MS analysis.

**LC-MS/MS and database analysis.** An LCQ ion trap mass spectrometer (Finnigan MAT, San Jose, CA) was used with a HP1100 solvent delivery system (Agilent, Palo Alto, CA). Peptides were pressure-loaded onto the column, then eluted and analyzed by microcapillary LC-MS/MS as described previously<sup>14</sup>. The collision energy for the LCQ was set at 30%. SEQUEST (ref. 19) was used for searching yeast YPD database for peptide sequence and identification of site(s) of phosphorylation. Search parameters included differential mass modification to serine, threonine, and tyrosine due to phosphorylation, static mass modification to aspartate, glutamate, and the C terminus due to the ethanolamine tag, and static mass modification to cysteine due to alkylation by iodoacetamide. Differential mass modification means both possibilities of modified or unmodified amino acid residues were used in database search, whereas static mass modification means that only the modified amino acid residue was used.

### Acknowledgments

The National Science Foundation Science and Technology Center for Molecular Biotechnology, the National Institutes of Health (RO1 A141109 and 1R33 CA84698), the NIH Research Resource Center (RR11823), and the Merck Genome Research Institute provided support for this work. We thank Drs. Beate Rist, Steven P. Gygi, and David R. Goodlett for helpful discussions.

1. Graves, J.D. & Krebs, E.D. Protein phosphorylation and signal transduction. *Pharmacol. Ther.* **82**, 111–121 (1999).
2. Koch, C.A., Anderson, D., Moran, M. F., Ellis, C. & Pawson, T. SH2 and SH3 domains: elements that control interactions of cytoplasmic signaling proteins. *Science* **252**, 668–674 (1991).
3. Hunter, T. 1001 protein kinases redux—towards 2000. *Semin. Cell Biol.* **5**, 367–376 (1994).
4. Verma, R. *et al.* Phosphorylation of Sic1p by G1 Cdk required for its degradation and entry into S phase. *Science* **278**, 455–460 (1997).
5. Watts, J.D. *et al.* Identification by electrospray ionization mass spectrometry of the sites of tyrosine phosphorylation induced in activated Jurkat T cells on the protein tyrosine kinase ZAP-70. *J. Biol. Chem.* **269**, 29520–29529 (1994).
6. Gingras, A.C. *et al.* Regulation of 4E-BP1 phosphorylation: a novel two-step mechanism. *Genes Dev.* **13**, 1422–1437 (1999).
7. Bodanszky, A.B.M. (ed.) *The practice of peptide synthesis*, Vol. 21. (Springer-Verlag, New York; 1984).
8. Hoare, D.G. & Koshland, D.E. Jr. A method for the quantitative modification and estimation of carboxylic acid groups in proteins. *J. Biol. Chem.* **242**, 2447–2453 (1967).
9. Chu, B., Wahl, G.M. & Orgel, L.E. Derivatization of unprotected polynucleotides. *Nucleic Acids Res.* **11**, 6513–6529 (1983).
10. Papayannopoulos, I.A. The interpretation of collision-induced dissociation tandem mass spectra of peptides. *Mass Spectrom. Rev.* **14**, 49–73 (1995).
11. Jonscher, K.R. & Yates, J.R. Matrix-assisted laser desorption/ionization/quadrupole ion trap mass spectrometry of peptides. Application to the localization of phosphorylation sites on the P protein from Sendai virus. *J. Biol. Chem.* **272**, 1735–1741 (1997).
12. Qin, J. & Chait, B.T. Identification and characterization of posttranslational modifications of proteins by MALDI ion trap mass spectrometry. *Anal. Chem.* **69**, 4002–4009 (1997).
13. Aebersold, R., Watts, J.D., Morrison, H.D. & Bures, E.J. Determination of the site of tyrosine phosphorylation at the low picomole level by automated solid-phase sequence analysis. *Anal. Biochem.* **199**, 51–60. (1991).
14. Gygi, S.P. *et al.* Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17**, 994–999 (1999).
15. Gygi, S.P., Rochon, Y., Franza, B.R. & Aebersold, R. Correlation between protein and mRNA abundance in yeast. *Mol. Cell. Biol.* **19**, 1720–1730 (1999).
16. Fletcher, B., Latter, G.I., Monardo, P., McLaughlin, C.S., Garrels, J.I. A sampling of the yeast proteome. *Mol. Cell. Biol.* **19**, 7357–7368 (1999).
17. Moore, D.D. *et al.* (eds). *Current protocols in molecular biology*. (Wiley, New York; 1987).
18. Carraway, K.L. & Koshland, D.E. Jr. Reaction of tyrosine residues in proteins with carbodiimide reagents. *Biochim. Biophys. Acta* **160**, 272–274 (1968).
19. Eng, J., McCormack, A.L. & Yates, J.R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **5**, 976–989 (1994).