Anal. Chem. **2004,** *76,* 4472−4483

Analysis of Protein Phosphorylation by Hypothesis-Driven Multiple-Stage Mass Spectrometry

Emmanuel J. Chang,[†] Vincent Archambault,[‡] Derek T. McLachlin,[†,§] Andrew N. Krutchinsky,[†] and Brian T. Chait*[,†]

Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, and Laboratories of Cellular and Structural Biology and Yeast Molecular Genetics, The Rockefeller University, 1230 York Avenue, New York, New York 10021

We describe a strategy, which we term hypothesis-driven multiple-stage mass spectrometry (HMS-MS), for the sensitive detection and identification of phosphopeptides derived from enzymatic digests of phosphoproteins. In this strategy, we postulate that any or all of the potential sites of phosphorylation in a given protein may be phosphorylated. Using this assumption, we calculate the *m*/*z* values of all the corresponding singly charged phosphopeptide ions that could, in theory, be produced by the enzyme employed for proteolysis. We test ions at these *m*/*z* values for the presence of phosphoserine or phosphothreonine residues using tandem mass spectrometry (MS^2) in a vacuum MALDI ion trap mass spectrometer, where the neutral loss of the elements of H_3PO_4 (98 Da) provides a sensitive assay for the presence of phosphopeptides. Subsequent MS^3 analysis of the $(M + H - 98)^+$ peaks allows us to confirm or reject the hypotheses that the putative phosphopeptides are present in the sample. HMS-MS was successfully applied to the detection and identification of phosphopeptides from substrates of the *Saccharomyces cerevisiae* cyclin-dependent kinase (Cdk) Cdc28, phosphorylated in vitro (Ipl1) and in vivo (Orc6), basing hypothesis formation on the minimal Cdk consensus phosphorylation motif Ser/Thr-Pro. The method was also used to find in vitro phosphopeptides from a domain of the *Drosophila melanogaster* protein PERIOD, hypothesizing possible phosphorylations of all Ser/Thr residues without assuming a consensus motif. Our results demonstrate that HMS-MS is a sensitive, highly specific tool for systematically surveying proteins for Ser/Thr phosphorylation, and represents a significant step toward our goal of comprehensive phosphorylation mapping.

Phosphorylation of proteins at serine, threonine, and tyrosine residues regulates many cellular processes and pathways, such as cell cycle progression, signal transduction cascades, and gene expression.[1,2] To understand such processes, it is critical to have available effective methods for following phosphorylation. Although techniques such as immunochemistry, autoradiography, and Edman sequencing have been successfully used for this purpose,[3] the study of protein phosphorylation continues to present a formidable analytical challenge. Mass spectrometry (MS) is playing an increasingly important role in this task because the method is relatively fast and sensitive, typically does not require radioisotope labeling, and is often able to localize the covalent attachment of phosphate to specific residues.[4−6]

MS of intact phosphoproteins can determine the average number of attached phosphate groups, and given sufficient resolution can provide a measure of the distribution of the number of phosphates.[7−11] MS of phosphopeptides derived from enzyme digestion of phosphoproteins can localize the attachment of phosphate moieties to specific peptides.[12,13] In tandem MS (MS^2), phosphorylated peptide ions undergo preferential neutral loss of the elements H_3PO_4 or HPO_3 (98 and 80 Da, respectively), although they can also fragment along the backbone.[14−17] Thus, MS^2 can be used to confirm the phosphorylation of a peptide and

* To whom correspondence should be addressed: Telephone: (212) 327-8849. Fax: (212) 327-7547. E-mail: chait@rockefeller.edu.

[†] Laboratory of Mass Spectrometry and Gaseous Ion Chemistry.

[‡] Laboratories of Cellular and Structural Biology and Yeast Molecular Genetics.

[§] Current address: BioMedCom Consultants, inc., 6700 Côte de Liesse, Suite 105, Montreal, Quebec H4T 2B5, Canada.

(1) Hunter, T. *Cell* **2000**, *100*, 113−127.
(2) Chen, P. *Trends Biochem. Sci.* **2000**, *25*, 596−601.
(3) Yan, J. X.; Packer, N. H.; Gooley, A. A.; Williams, K. L. *J. Chromatogr., A* **1998**, *808*, 23−41.
(4) McLachlin, D. T.; Chait, B. T. *Curr. Opin. Chem. Biol.* **2001**, *5*, 591−602.
(5) Mann, M.; Ong, S.-E.; Gronborg, M.; Steen, H.; Jensen, O. N.; Pandey, A. *Trends Biotechnol.* **2002**, *20*, 261−268.
(6) Jensen, O. N. *Curr. Opin. Chem. Biol.* **2004**, *8*, 33−41.
(7) Carmel, G.; Leichus, B.; Cheng, X.; Patterson, S. D.; Mirza, U.; Chait, B. T.; Kruet, J. *J. Biol. Chem.* **1994**, *269*, 7304−7309.
(8) Ge, H.; Zhao, Y.; Chait, B. T.; Roeder, R. G. *Proc. Natl. Acad. Sci., U.S.A.* **1994**, *91*, 12691−12695.
(9) Hunter, A. P.; Games, D. E. *Rapid Commun. Mass Spectrom.* **1994**, *8*, 559−570.
(10) Weijland, A.; Neubauer, G.; Courtneidge, S. A.; Mann, M.; Wierenga, R. K.; Superti-Furga, G. *Eur. J. Biochem.* **1996**, *240*, 756−764.
(11) Wind, M.; Wesch, H.; Lehmann, W. D. *Anal. Chem.* **2001**, *73*, 3006−3010.
(12) Larsen, M. R.; Sorensen, G. L.; Fey, S. J.; Larsen, P. M.; Roepstorff, O. *Proteomics* **2001**, *1*, 223−238.
(13) Liao, P. C.; Leykam, J.; Andrews, P. C.; Gage, D. A.; Allison, J. *Anal. Biochem.* **1994**, *219*, 9−20.
(14) Bennett, K. L.; Stensballe, A.; Podtelejnikov, A. V.; Moniatte, M.; Jensen, O. N. *J. Mass Spectrom.* **2002**, *37*, 179−190.
(15) Ma, Y.; Lu, Y.; Zeng, H.; Ron, D.; Mo, W.; Neubert, T. A. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1693−1700.
(16) Zhang, X.; Herring, C. J.; Romano, P. R.; Szczepanowska, J.; Brzeska, H.; Hinnebusch, A. G.; Qin, J. *Anal. Chem.* **1998**, *70*, 2050−2059.
(17) Lee, C. H.; McComb, M. E.; Bromirski, M.; Jilkine, A.; Ens, W.; Standing, K. G.; Perrault, H. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 191−202.

10.1021/ac049637h CCC: $27.50
Published on Web 06/10/2004

under favorable conditions can also shed light on the specific site of phosphorylation.[18−20]

Although powerful, MS as currently applied to the analysis of phosphorylation still has a number of practical limitations. Thus, incomplete sequence coverage, an issue in most protein MS experiments, is exacerbated in phosphorylation analysis because the negative charge on phosphate moieties can reduce the MS response in the positive ion mode and can also hinder proteolytic digestion by some enzymes, including trypsin.[4,21] In addition, phosphopeptides are often present at lower abundance than their nonphosphorylated counterparts. The higher abundance nonphosphorylated peptides can suppress the MS response of the phosphopeptides. Therefore, truly complete analysis of protein phosphorylation requires technical improvements in the sensitivity and comprehensiveness of phosphopeptide detection. One approach is to selectively enrich the phosphorylated component of peptide mixtures. Thus, immobilized metal ion affinity chromatography (IMAC) has been used to selectively enrich phosphopeptides.[22−25] Although methyl esterification of acidic residues has proved useful for reducing the nonphosphopeptide background normally present after IMAC,[26] the technique is complicated by preferential recovery of multiply phosphorylated residues and variable phosphopeptide recovery under different binding and elution conditions. As an alternative to IMAC, a number of groups have reported the chemical conversion of phosphorylated residues to moieties that are readily purified from mixtures using affinity- or covalent chemistry-based methods. Currently, these chemistry-based methods are limited by side reactions and insufficient sensitivity and have not yet attained widespread use.[27−31]

Another major theme in phosphorylation analysis is the use of MS^2 for identifying phosphopeptides. Several groups have utilized single-stage mass spectrometric maps of proteolytic digests to identify potential phosphopeptides, followed by MS^2 interrogation of these candidates to test whether they are, indeed, phosphorylated and to confirm or determine their identity.[14−17] Limitations of this approach include incomplete sequence coverage; inadequate signal-to-noise in the single-stage mass spectrum; and in LC/MS^2 analyses, the relatively small time window available for the examination of each eluting peptide. In contrast, linked scanning techniques avoid reliance on single-stage mass spectra by performing MS^2 at every m/z value in the mass range of interest. Thus, low-level phosphopeptides that might not be observed in a single-stage spectrum can be detected by assaying for the loss of the elements of H_3PO_4 (98 Da, in a neutral loss linked-scan experiment) or PO_3^- (79 Da, in a precursor ion scan experiment).[14,32−35] Although extraordinarily powerful, such methods suffer from low efficiency of sample usage as a result of the need to scan the mass spectrometer over the full m/z range of interest.

Recently Krutchinsky et al.[36] have described a MALDI-ion trap mass spectrometer that can perform multiple-stage MS experiments with high speed and sensitivity and have demonstrated useful MS^2 and MS^3 mass spectra from peptide ions where the single-stage MS signals are masked by chemical noise.[37] In the present work, we utilize the unique characteristics of this instrument to develop an effective scheme (termed hypothesis-driven multiple-stage MS[38] (HMS-MS)) for the analysis of Ser/Thr phosphorylated proteins. This approach is capable of detecting phosphopeptides resulting from low stoichiometry phosphorylation and constitutes an important step toward our goal of comprehensive phosphorylation site detection. In this paper, we explore the sensitivity limits of HMS-MS for the detection of phosphopeptides and apply the technique to the analysis of both in vitro- and in vivo-phosphorylated proteins.

EXPERIMENTAL SECTION

Instrumentation. Mass spectra were collected using a commercial MALDI-TOF mass spectrometer (STR-Voyager DE Workstation, Applied Biosystems, Foster City, CA), an in-house modified prototype quadrupole-quadrupole-TOF mass spectrometer[39] (Centaur, Sciex, Concord, ON, Canada) and an in-house-assembled vacuum MALDI-ion trap mass spectrometer that incorporates a Finnigan LCQ DecaXP (ThermoElectron, San Jose, CA) mass analyzer.[36]

Calculation of Hypothetical Masses. Hypothetical masses of potential phosphopeptides were calculated using in-house software written in PERL. The software takes as input the sequence of the protein of interest and the phosphorylation site consensus motif (if any) and outputs a list of potential phosphopeptides and their corresponding molecular masses. The cyclin-dependent kinase (Cdk) minimal consensus motif (Ser/Thr followed by Pro)[40] was used as input along with the sequences of *Saccharomyces cerevisiae* Ipl1 (N-terminal domain) and Orc6. For the Per protein construct (residues 531−700) from *Drosophila melanogaster*, all Ser/Thr residues within the domain encompassing residues 531−640 were considered to be possible phosphorylation sites.

(18) Neubauer, G.; Mann, M. *Anal. Chem.* **1999**, *71*, 235−242.
(19) Annan, R. S.; Huddleston, M. J.; Verma, R.; Deshaies, R. J.; Carr, S. A. *Anal. Chem.* **2001**, *73*, 393−404.
(20) DeGnore, J. P.; Qin, J. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 1175−1188.
(21) Hynek, R.; Zuzalkova, L.; Sikac, Z.; Kas, J. *J. Liq. Chromatogr. Relat. Technol.* **2002**, *25*, 2307−2314.
(22) Stensballe, A.; Andersen, S.; Jensen, O. N. *Proteomics* **2001**, *1*, 207−222.
(23) Raska, C. S.; Parker, C. E.; Dominski, Z.; Marzluff, W. F.; Glish, G. L.; Pope, R. M.; Borchers, C. H. *Anal. Chem.* **2002**, *74*, 3429−3433.
(24) Xhou, W.; Merrick, B. A.; Khaledi, M. G.; Tomer, K. B. *Anal. Biochem.* **1999**, *274*, 174−180.
(25) Posewitz, M. C.; Tempst, P. *Anal. Chem.* **1999**, *71*, 2883−2892.
(26) Ficarro, S. B.; McCleland, M. L.; Stukenberg, P. T.; Burke, D. J.; Ross, M. M.; Shabanowitz, J.; Hunt, D. F.; White, F. M. *Nat. Biotechnol.* **2002**, *20*, 301−305.
(27) Oda, Y.; Nagasu, T.; Chait, B. T. *Nat. Biotechnol.* **2001**, *19*, 379−382.
(28) Qian, W.-J.; Goshe, M. B.; Camp, D. G., II; Yu, L.-R.; Tang, K.; Smith, R. D. *Anal. Chem.* **2003**, *75*, 5441−5450.
(29) McLachlin, D. T.; Chait, B. T. *Anal. Chem.* **2003**, *75*, 6826−6836.
(30) Zhou, H.; Watts, J. D.; Aebersold, R. *Nat. Biotechnol.* **2001**, *19*, 375−378.
(31) Goshe, M. B.; Conrads, T. P.; Panisko, E. A.; Angell, N. H.; Veenstra, T. D.; Smith, R. D. *Anal. Chem.* **2001**, *73*, 2578−2586.
(32) Schlosser, A.; Pipkorn, R.; Bossemeyer, D.; Lehmann, W. D. *Anal. Chem.* **2001**, *73*, 170−176.
(33) Wilm, M.; Neubauer, G.; Mann, M. *Anal. Chem.* **1996**, *68*, 527−533.
(34) Carr, S. A.; Huddleston, M. J.; Annan, R. S. *Anal. Biochem.* **1996**, *239*, 180−192.
(35) Steen, H.; Kuster, B.; Fernandez, M.; Pandey, A.; Mann, M. *Anal. Chem.* **2001**, *73*, 1440−1448.
(36) Krutchinsky, A. N.; Kalkum, M.; Chait, B. T. *Anal. Chem.* **2001**, *73*, 5066−5077.
(37) Krutchinsky, A. N.; Chait, B. T. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 129−134.
(38) Kalkum, M.; Lyon, G. J.; Chait, B. T. *Proc. Natl. Acad. Sci., U.S.A.* **2003**, *100*, 2795−27800.
(39) Krutchinsky, A. N.; Zhang, W.; Chait, B. T. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 493−504.
(40) Nigg, E. A. *Trends Cell Biol.* **1993**, *3*, 296−301.

MS^n Data Analysis. MS^2 spectra were typically collected for 5−10 s using an isolation width of 4 m/z units, injection time of 500 ms, activation time of 300 ms, and relative collision energy of 30%. MS^2 spectra were analyzed using an in-house program Look4Loss written using National Instruments LabView (Austin, TX). Look4Loss analyzes Finnigan DTA format MS^2 data files for peaks arising from neutral loss of a user-defined mass. Peptides exhibiting a neutral loss of 98 Da were interrogated further by MS^3, using the $(M + H - 98)^+$ precursor ion. MS^3 spectra were typically collected for 0.5−5 min, using parameters similar to those for the MS^2 experiments. The resulting MS^3 spectra were analyzed manually.

Sample Preparation. *Synthetic Phosphopeptide Mixture.* Synthetic phosphopeptides (Anaspec, San Jose, CA) with the sequences SLRRSpSCFGGRIDRIGAQSGLGCNSFRY (monoisotopic molecular mass = 3144.5 Da), GRTGRRNpSIHDIL (1574.5 Da), KRpTIRR (909.0 Da), and RRApSPVA (835.9 Da) were reconstituted in sample solvent (50% methanol, 20% acetonitrile, and 0.1% aqueous trifluoroacetic acid (TFA)), mixed in an equimolar ratio, and diluted as appropriate. Unphosphorylated peptide standards were PPGFSPFR (903.5 Da), RPKPQQFFGLM−NH_2 (1346.7 Da), pyroELYENKPRRPYIL (1671.9 Da), VHHQKLVFFAEDVGSNK (1954.0 Da), and SYSMEHFRWGKPVGKKRRPVKVYP (2931.6 Da). These were mixed in equimolar amounts and added to the phosphopeptide mixture so that all components of the final mixture were at the same molar concentration.

For experiments with the protein digest background, bovine serum albumin (BSA) and β-casein (Sigma Chemical Co., St. Louis, MO) were digested separately with modified trypsin (Roche, Indianapolis, IN) in 100 mM ammonium bicarbonate (ABC) solution at a 20:1 substrate-to-enzyme molar ratio. Working solutions were made, each containing the two protein digests at final concentrations of 500 fmol/μL as well as the appropriate concentration of phosphopeptide mix.

2,5-Dihydroxybenzoic acid (DHB) (Sigma) was prepared as a saturated solution in the sample solvent (50% methanol, 20% acetonitrile, 0.1%TFA), and diluted 1:1 in the same solvent. Equal volumes of the diluted DHB solution and the phosphopeptide mixture or phosphopeptide/protein digest mixture were added together and vortexed thoroughly. A 2-μL spot was deposited onto the surface of our compact disk (CD) MALDI target[36] and allowed to dry in ambient air.

Ipl1. For the expression of the N-terminal fragment of Ipl1, we used a construct made in the pET-16b vector (Stratagene, La Jolla, CA) (gift from Zu-Wen Sun and David Allis). This plasmid was transformed in *Escherichia coli* BL21(DE3) (Stratagene), and cells were grown in LB medium. At an OD_{600} of 0.6−0.7, 1 mM IPTG (Sigma) was added, and the culture was transferred to room temperature for a 6-h induction. Cells were harvested by centrifugation and resuspended in 10 mL of extraction buffer (50 mM Na_2HPO_4, 300 mM NaCl, pH 8.0). Suspensions were sonicated in 30 -s bursts for 4 min with a Sonicator Ultrasonic processor (Misonix Inc, Farmingdale, NY). Lysates were clarified by centrifugation and supernatants were added to 0.5 mL Talon metal affinity resin (Clontech, Palo Alto, CA) and incubated at room temperature with agitation for 20 min. Three washes were done with 5 mL of extraction buffer for 10 min at room temperature. Two additional washes were done with 0.5 mL of extraction buffer + 100 mM imidazole (Sigma) for 10 min at room temperature. The resin was finally eluted with 0.5 mL of extraction buffer + 500 mM imidazole.

The Cdc28 yeast kinase was prepared as follows for the in vitro kinase assay using His10/NTD/Ipl1 as the substrate. Cdc28 was tagged with the IgG binding domains of protein A (PRA) from *Staphylococcus aureus* by genomic integration of DNA fragments placing the tag at the 3′ end of the gene, followed by the HIS3MX gene from *Schizosaccharomyces pombe* for selection.[41−43] The C-terminally tagged Cdc28−protein A fusion was thereby expressed under the control of its endogenous promoter. It is, therefore, expected to be expressed at a level similar to endogenous Cdc28 and to associate with its binding partners at similar stoichiometries. CDC28−PRA::HIS3 cells were grown in YPD and harvested by centrifugation. They were lysed by grinding in liquid nitrogen as described.[44,45] Protein A affinity purification was performed on the basis of a documented method.[44,46,47]

Frozen CDC28−PRA::HIS3 lysates were resuspended in 3/4 volume of buffer N (50 mM Tris−Cl, pH 7.5, 100 mM NaCl, 0.1 mM EDTA, 1 mM DTT, 10 mM NaF 0.5 M, 50 mM β-glycerophosphate, 0.1% NP-40, 4 μg/mL pepstatin A, and Protease Inhibitors Cocktail 1:200 (Sigma)). The suspension was clarified by centrifugation at 2000 rpm for 30 min (T6000D; Sorvall, Newtown, CT) and at 50 000 rpm for 1 h in a Ty50.2 rotor (Beckman Coulter, Fullerton, CA). The supernatant was incubated for 6 h with IgG-Sepharose resin (rabbit IgG, ICN; CNBr-activated Sepharose 4B, Amersham Biosciences Corp, Piscataway, NJ). The resin was washed 3 times with 10 mL of buffer N and once with 10 mL of 10 mM K-HEPES pH 7.5, 50 mM NaCl, 50% gycerol, 4 μg/mL pepstatin A, and Protease Inhibitors Cocktail 1:200 (Sigma), then resuspended in 1 volume of the same solution. The resulting resin slurry was stored at −20 °C and subsequently used in kinase assays.

The in vitro kinase assay was performed as follows. Six microliters of the Cdc28 bead slurry was incubated with 4 μL (~1 pmol/μL) His10−NTD−Ipl1, 1 μL of 100 mM ATP in 8 μL of kinase buffer containing 20 mM K-HEPES buffer, and 1 mM $MgCl_2$. The mixture was spun briefly to remove kinase beads. One-quarter of the undigested in vitro-phosphorylated protein was bound to reversed-phase C4 resin (Grace Vydac, Anaheim, CA), eluted with a solution of 50% formic acid, 33% 2-propanol, and 17% water, saturated with recrystallized α-cyano-4-hydroxycinnamic acid (4-HCCA, Sigma), directly onto a stainless steel MALDI target precoated with an ultrathin layer of 4-HCCA,[48] and analyzed by MALDI-TOF mass spectrometry using a Voyager STR instrument. The remaining kinase reaction was run on a 4−20% tris-glycine SDS−PAGE gel. The Coomassie blue-stained (GelCode Blue reagent, Pierce, Rockford, Il.) gel showed a single band (data not shown), which was excised, chopped in ~1 mm^3 pieces, and

(41) Wach, A.; Brachat, A.; Alberti-Segui, C.; Rebischung, C.; Philippsen, P. *Yeast* **1997**, *13*, 1065−1075.
(42) Rout, M. P.; Aitchison, J. D.; Suprapto, A.; Hjertaas, K.; Zhao, Y.; Chait, B. T. *J. Cell Biol.* **2000**, *148*, 635−651.
(43) Cross, F. R.; Archambault, V.; Miller, M.; Klovstad, M. *Mol. Biol. Cell* **2002**, *13*, 52−70.
(44) Archambault, V.; Li, C. X.; Tackett, A. J.; Wasch, R.; Chait, B. T.; Rout, M. P.; Cross, F. R. *Mol. Biol. Cell* **2003**, *14*, 4592−4604.
(45) Schultz, M. C. *Methods* **1999**, *17*, 161−172.
(46) Rout, M. P.; Blobel, G.; Aitchison, J. D. *Cell* **1997**, *89*, 715−725.
(47) Aitchison, J. D.; Blobel, G.; Rout, M. P. *Science* **1996**, *274*, 624−627.
(48) Cadene, M.; Chait, B. T. *Anal. Chem.* **2000**, *72*, 5655−5658.

subjected to destaining in 50% acetonitrile and 50 mM ABC, followed by wash steps with 100% acetonitrile, 50 mM ABC, and again with acetonitrile. Twenty nanograms of modified trypsin (Roche) in 2 μL of 50 mM ABC was applied to the gel pieces, followed by 10 μL of 50 mM ABC, and the pieces were incubated at 37 °C overnight. A 1:1 mixture of R2 and R3 Poros (Applied Biosystems) resin was resuspended in 19 volumes of an aqueous solution containing 5% formic acid and 0.2% TFA, and 12.5 μL of this slurry was added to the in-gel digest. After incubation at 4 °C for 8 h to extract the peptides, the resin was collected on a C18 ZipTip (Millipore, Billerica, MA) (preconditioned by washing 2× with 0.1% TFA, 4× with sample solution, and again 4× with 0.1%TFA) and washed 2× with 0.1% TFA. Sample solution saturated with DHB was diluted 1:2 with sample solution, and 2.5 μL of this solution was applied to the resin. Peptides were eluted directly onto a polycarbonate compact disk MALDI target. MS analysis was performed using the Sciex Centaur mass spectrometer,[39] and MS^n HMS-MS experiments were performed using the MALDI-ion trap instrument.[36]

Two phosphopeptides from Ipl1 that we observed by HMS-MS (ie. IPpSPVREK and ISHpSPQQRNPNSKIPpSPVREK) were also synthesized in the solid phase (Rockefeller University Proteomics Resource Center) to allow us to verify and examine in detail their observed fragmentation patterns. MALDI samples (200, 20, and 2 fmol) were prepared in 1-μL spots of 1/3 saturated DHB.

Orc6. ORC6−PRA::HIS3 cell preparation, growth, lysis, and PrA affinity purification were performed essentially using the same procedure as for CDC28−PRA::HIS3. Two batches of Orc6-PrA were prepared. For each batch, 10 g of ORC6−PRA::HIS3 frozen lysate was resuspended in 50 mL of extraction buffer (containing 20 mM K-HEPES pH 7.4, 110 mM AcOK, 0.1% Tween-20, 1% Triton X-100, 1 mg/mL heparin, 2 mM $MgCl_2$, 1 mM DTT, 0.2 mg/mL PMSF, 4 μg/mL pepstatin A, Protease Inhibitors Cocktail 1:200, and either 300 mM NaCl or no NaCl), homogenized with a Polytron (PT 10/35; Brinkmann, Westbury, NY) and incubated at 4 °C for 1 h with gentle agitation. The soluble fraction was isolated by centrifugation at 3000 rpm (T6000D; Sorvall) for 10 min, and another centrifugation at 30 000 rpm for 30 min in a Ty50.2 rotor (Beckman). The supernatant was incubated overnight with 80 μL of IgG-Sepharose resin (as above). The resin was collected by quick centrifugation and washed once for 5 min with 10 mL of extraction buffer and 3 times with 1 mL of extraction buffer. The resin was further washed twice with 400 μL of 100 mM NH_4OAc and 1 mM $MgCl_2$. The resin was eluted twice with 300 μL of 0.5 M NH_4OH and 0.5 mM EDTA. The eluates were pooled and lyophilized in a SpeedVac (ThermoSavant, Holbrook, NY).

The resulting dried proteins were resuspended in SDS−PAGE sample buffer and separated on a Novex 4−20% Tris−glycine polyacrylamide gel (Invitrogen, Carlsbad, CA). Proteins in the gel were visualized by Coomassie staining with GelCode Blue. Staining patterns for the two preparations were essentially identical (data not shown). Tryptic digestion and sample preparation for MALDI analysis was performed essentially as described,[44] and two bands containing Orc6-PrA were identified from MS and MS^2 spectra using the protein identification programs ProFound[49,50] and Sonar (Genomic Solutions, Ann Arbor, MI), respectively. Previously, it has been indicated that the more slowly migrating of the two Orc6 SDS−PAGE bands is hyperphosphorylated;[51] therefore, material from the upper band of our gel (Figure 4A) was used for the HMS-MS phosphorylation studies. Data reported below is an aggregate for experiments on two gel lanes.

Period. A domain of the *D. melanogaster* PERIOD protein containing residues 531−700 (Per 531), phosphorylated in vitro by rat casein kinase I in a solution of 150 mM NaCl, 50 mM Tris, and 1 mM EDTA (concentration ~200 μg/mL) was obtained from Saul Kivimäe and Michael Young (Rockefeller University). This solution was diluted 10-fold in 0.1% TFA, and 2.5 μL was deposited directly onto the CD MALDI target. This spot was allowed to dry at 37 °C. To denature proteins on the polycarbonate CD surface, 2.5 μL of 1% TFA was applied to the dried spot and allowed to evaporate at 37 °C. To remove soluble contaminants, 2.5 μL of ice-cold 0.1% TFA was applied to the dried spot and quickly removed under vacuum. Trypsin was added to the dried spot in 2.5 μL 50 mM ABC so that the final molar ratio of enzyme to Per 531 was 1:20. The CD was placed in its plastic case, and the spot was surrounded by multiple 100-μL droplets of HPLC grade water such that when the case was shut, the 100-μL droplets formed columns from the CD surface to the inside of the case cover. These columns of water helped to keep the spot hydrated during the course of the trypsin digestion. After incubating the CD at 37 °C for 3 h, the spot had completely dried out and was rehydrated by adding 2.5 μL of 50 mM ABC. The reaction was incubated for another 3 h at 37 °C, at which point the spot was dried out again. At this point, the surrounding columns of water were then removed by vacuum aspiration, and 2.5 μL of DHB solution prepared as above was added to the spot, mixed with the dried tryptic digest by pipetting up and down several times, and allowed to dry in ambient air.

RESULTS AND DISCUSSION

HMS-MS Strategy. In the HMS-MS strategy, we hypothesize that any possible site of phosphorylation within a given protein may, in fact, be phosphorylated. We calculate the m/z of all the corresponding singly protonated phosphopeptide ions $(M + H)^+$ that could be generated by the enzyme employed for proteolysis and test these values for the presence of phosphoserine/threonine residues by MALDI-MS^2 (using the occurrence of an $(M + H - 98)^+$ peak to indicate the presence of a putative phosphopeptide). Subsequent MS^3 analysis of the $(M + H - 98)^+$ peaks either confirms or rejects the hypotheses.

The hypotheses concerning the sites of phosphorylation can be based on prior knowledge (e.g., from known consensus motifs, homologues, or previous mutational studies) or can simply be generated as the entire set of peptides that contain one or more Ser/Thr residues. Because the effective sensitivity of the MALDI-ion trap mass spectrometer is greater in MS^2 mode than in single-stage MS mode,[36,37] observation of the −98-Da signature from peptides provides a sensitive test for the presence of the hypothetical phosphopeptides.[14,52] Efficient sample utilization in the MALDI-ion trap mass spectrometer allows not only hypothesis-driven MS^2 but also extensive MS^3 interrogation of the (M + H

(49) Zhang, W.; Chait, B. T. *Anal. Chem.* **2000**, *72*, 2482−2489.

(50) http://prowl.rockefeller.edu.

(51) Nguyen, V. Q.; Co, C.; Li, J. J. *Nature* **2001**, *411*, 1068−1073.

(52) Qin, J.; Chait, B. T. *Anal. Chem.* **1997**, *69*, 4002−4009.

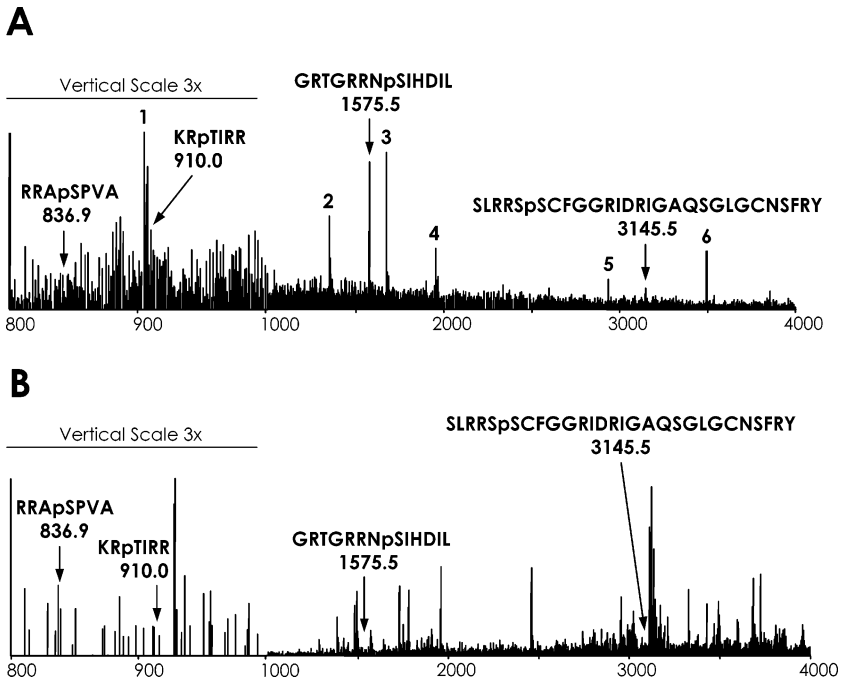


Figure 1. Single-stage mass spectra of four phosphorylated peptides in the presence of low- and high-complexity background peptide mixtures. (A) MALDI-ion trap mass spectrum of 20 fmol each of four synthetic phosphopeptides (sequences labeled) mixed with 20 fmol each of six standard peptides (labeled 1−6, sequences given in the Experimental Section). (B) MALDI-ion trap mass spectrum of 20 fmol each of the same four synthetic phosphopeptides mixed with a tryptic digest of 500 fmol each of bovine serum albumin and bovine β-casein. None of the phosphopeptides are observed in (B). In both spectra, the region marked vertical scale 3× is also expanded 7.5× on the horizontal scale.

$- 98)^+$ ion products. This further stage of analysis can prove decisive in confirming the identity of the peptide in cases in which the $(M + H - 98)^+$ signature peak in the MS^2 is weak or inconclusive, when different phosphorylated peptides coincidentally have the same nominal mass, and when a peptide is multiply phosphorylated.

HMS-MS of Synthetic Phosphopeptides: Sensitivity Limits. For HMS-MS to yield useful results, the mass spectrometer must perform informative multiple-stage mass spectrometry on low-abundance species in the presence of a complex background of peptides (which may be of relatively high abundance). To assess the ability of our MALDI-ion trap mass spectrometer to detect phosphopeptides by HMS-MS, known amounts of synthetic phosphopeptides were titrated into a complex mixture containing a molar excess of background peptides obtained from a tryptic digest of 500 fmol each of BSA and β-casein. Successive dilutions of an equimolar mixture of four phosphopeptides containing, respectively, 200, 20, and 2 fmol of each were mixed with the background digest and analyzed by MS, MS^2 and MS^3. For comparison, a similar titration was analyzed using an equimolar mixture of the phosphopeptides and the unphosphorylated standard peptides.

Figure 1A shows the spectrum of a mixture containing 20 fmol each of the four phosphopeptides and the six unphosphorylated standard peptides. Although all of the unphosphorylated peptides yielded relatively intense singly protonated peptide ion peaks, just two of the four phosphopeptides were readily discernible in the single-stage mass spectrum. Only at the 200-fmol level were we able to discern all four phosphopeptides in the single-stage mass spectra (data not shown). This problem of detecting phosphopeptides in the presence of unphosphorylated peptides is even more pronounced when the background peptides are present in molar excess. Thus, when 500 fmol of a digest of BSA and β-casein were added to the 20-fmol mixture of phosphopeptides, no trace of the phosphopeptides was discerned in the single-stage mass spectrum (Figure 1B). This latter case mimics the situation that is often encountered during the characterization of substoichiometric protein phosphorylation. Similar difficulties in detecting phosphopeptides in mixtures have been reported by others.[4,14,15]

Given that it can be problematic to detect phosphopeptides in single-stage MS, we explored the use of MS^2 for this purpose. It has been shown previously that singly charged phosphopeptides preferentially lose 98 Da during resonant excitation in the ion trap mass spectrometer[52−54] and that this specific signature[14] can be used to detect the presence of phosphopeptides. We found that MS^2 revealed this preferential loss of 98 Da for all phosphopeptides that were present in the mixture (Figure 2), even when only 2 fmol of the phosphopeptide was present. MS^3 experiments on the resulting $(M + H - 98)^+$ species provided confirmation of the identities of the phosphopeptides (Figure 2). As we decreased the phosphopeptide concentration or increased the complexity and concentration of the background peptides, the −98 Da signatures became less intense and, hence, less certain. In such cases, the MS^3 experiment on the $(M + H - 98)^+$ ion was especially useful for confirming the identity of the phosphopeptides, although ion statistics were ultimately limiting in such experiments (Figure 2). We conclude from these titrations that HMS-MS should be useful for phosphopeptides amounts as low as 2−20 fmol.

HMS-MS of In Vitro Phosphorylated Ipl1. Ipl1 is the *S. cerevisiae* homologue of the mammalian Aurora B kinase, a protein implicated in the regulation of chromosome segregation.[55,56] In

(53) Jonscher, K. R.; Yates, J. R., 3rd. *Biol. Chem.* **1997**, *272*, 1735−1741.
(54) Annan, R. S.; Carr, S. A. *Anal. Chem.* **1996**, *68*, 3413−3421.
(55) Chan, C. S.; Botstein, D. *Genetics* **1993**, *135*, 677−691.

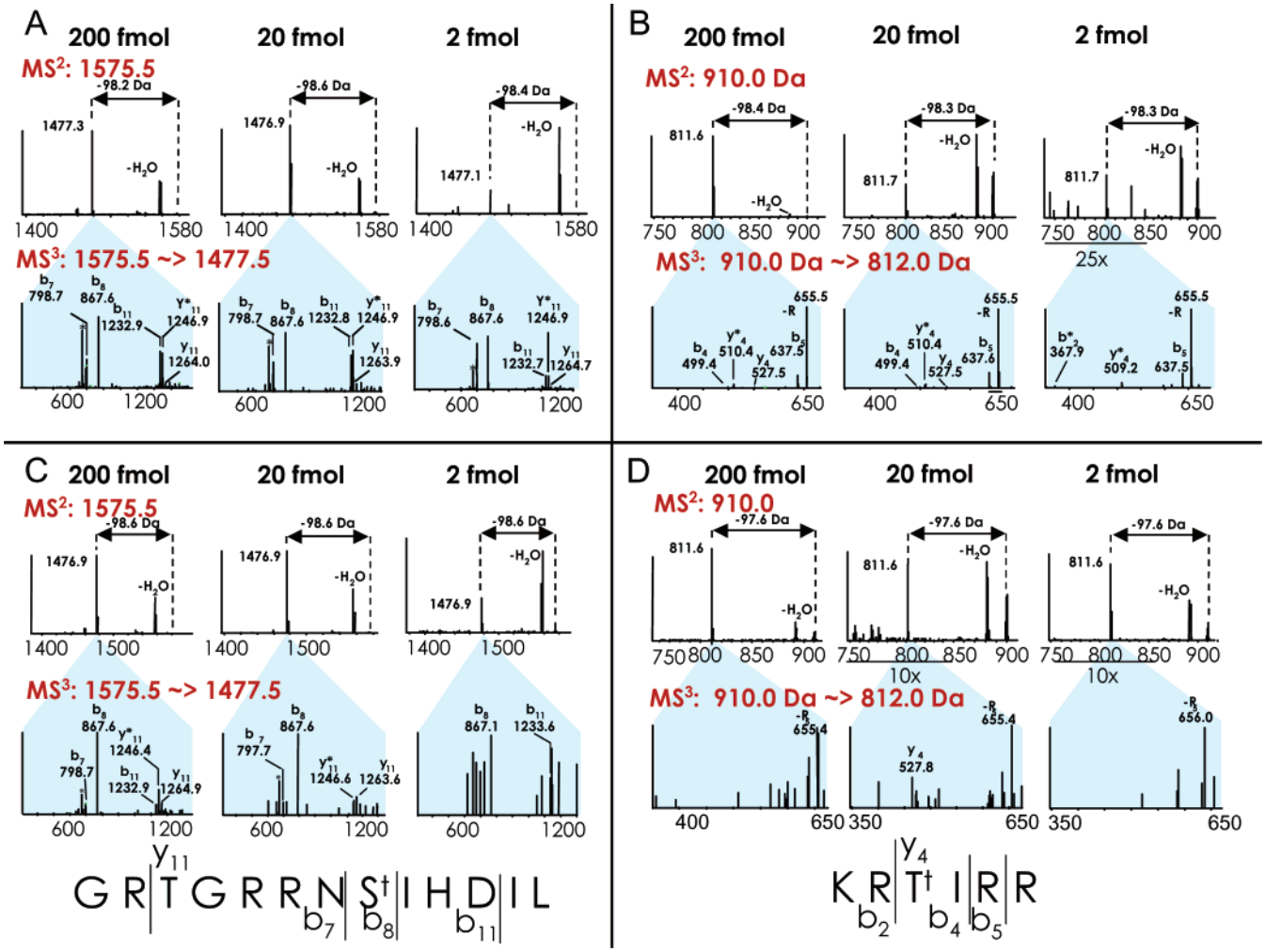
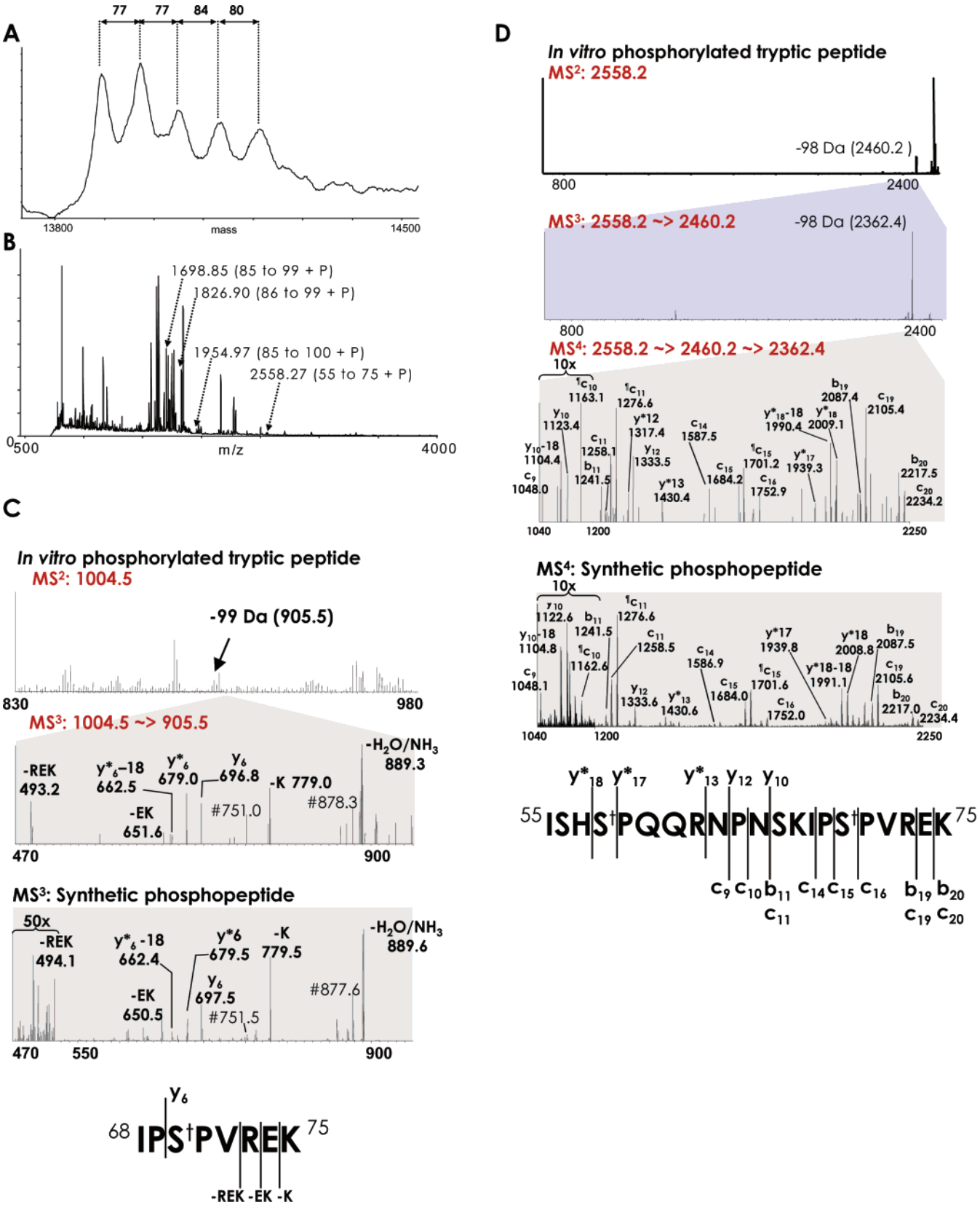


Figure 2. HMS-MS detection limits. (A) MS^2 and MS^3 analysis of a titration of the peptide GRTGRRNpSIHDIL from 200 to 2 fmol, in the presence of six synthetic peptides. (B) MS^2 and MS^3 analysis of a titration of the peptide KRpTIRR from 200 to 2 fmol in the presence of six synthetic peptides. In both (A) and (B), 2 fmol of phosphopeptide is detected using MS^2 and MS^3. (C) MS^2 and MS^3 analysis of a titration of the peptide GRTGRRNpSIHDIL, from 200 to 2 fmol in the presence of 500 fmol of proteolytic protein digest. (D) MS^2 and MS^3 analysis of a titration of the peptide KRpTIRR, from 200 to 2 fmol in the presence of 500 fmol of proteolytic protein digest. In cases (C) and (D), MS^2 can detect phosphopeptides down to 2 fmol, and the sensitivity limit for MS^3 lies between 2 and 20 fmol. The phosphorylated residues in the peptide sequences are marked with a superscript dagger (†). Peaks labeled with an asterisk (*) in panels A and C could not be readily explained.

vivo phosphorylation of Ipl1 by Cdc28 cyclin-dependent kinase (Cdk) complexes has been suggested but not conclusively demonstrated.[57−59] Here, we used HMS-MS to test whether Ipl1 was an in vitro substrate of Cdc28 and, if so, which sites were phosphorylated. Cdks are proline-directed Ser/Thr kinases; i.e., they phosphorylate Ser/Thr residues followed C-terminally by a Pro residue.[40] The N-terminal domain of Ipl1 (NTD-Ipl1, encompassing the first 100 residues of the protein and containing 4 SP motifs), was used in this study.

MALDI linear-TOF analysis of intact in vitro-phosphorylated NTD-Ipl1 yielded five peaks, spaced ~80 Da apart, suggesting that a fraction of the protein had been phosphorylated at least four times (Figure 3A). The resulting protein mixture migrated as a single band on SDS−PAGE (data not shown). The protein mixture within this band was digested with trypsin, and the products were analyzed by QqTOF-MS. The resulting single-stage mass spectrum yielded four discernible peaks corresponding to potential phosphopeptides from NTD-Ipl1, encompassing three of the four Ser/Pro sites (Figure 3B). Assuming that NTD-Ipl1 is exclusively phosphorylated at Cdk consensus motifs, this single-stage spectrum can at best explain only three of the four observed phosphorylations. We therefore subjected the digest to phosphopeptide mapping by HMS-MS in an attempt to detect and positively identify phosphorylated peptides whose abundance or response rendered them invisible in the single-stage mass spectrum. For this purpose, phosphorylation at any or all proline directed sites was hypothesized, a total of 57 peptides allowing for up to four missed cleavages by trypsin (Table S1). MS^2 spectra were acquired at m/z values corresponding to the calculated values of the hypothetical phosphopeptides. Signal corresponding to the neutral loss of 98 Da was detected in 27 of the 57 acquired MS^2 spectra. Ten of these 27 peptides yielded MS^3 spectra with structural information consistent with the hypothesized phospho-

(56) Ke, Y. W.; Dou, Z.; Zhang, J.; Yao, X. B. *Cell Res.* **2003**, *13*, 69−81.
(57) Haydon, C. E.; Eyers, P. A.; Aveline-Wolf, L. D.; Resing, K. A.; Maller, J. L.; Ahn, N. G. *Mol. Cell Proteomics* **2003**, *2*, 1055−1067.
(58) Francisco, L.; Chan, C. S. *Cell Mol. Biol. Res.* **1994**, *40*, 207−213.
(59) Ubersax, J. A.; Woodbury, E. L.; Quang, P. N.; Paraz, M.; Blethrow, J. D.; Shah, K.; Shokat, K. M.; Morgan, D. O. *Nature* **2003**, *425*, 859−864.


A
77
77
84
80
13800
mass
14500
B
1698.85 (85 to 99 + P)
1826.90 (86 to 99 + P)
1954.97 (85 to 100 + P)
2558.27 (55 to 75 + P)
0
500
m/z
4000
C
In vitro phosphorylated tryptic peptide
MS2: 1004.5
-99 Da (905.5)
830
980
MS3: 1004.5 ~> 905.5
-REK
493.2
y*6–18
662.5
y*6
679.0
y6
696.8
-K 779.0
-H2O/NH3
889.3
-EK
651.6
#751.0
#878.3
470
900
MS3: Synthetic phosphopeptide
50x
-REK
494.1
y*6 -18
662.4
y*6
679.5
-K
779.5
-H2O/NH3
889.6
-EK
650.5
y6
697.5
#751.5
#877.6
470
550
900
y6
68 IPS†PVREK 75
-REK -EK -K
D
In vitro phosphorylated tryptic peptide
MS2: 2558.2
-98 Da (2460.2)
800
2400
MS3: 2558.2 ~> 2460.2
-98 Da (2362.4)
800
2400
MS4: 2558.2 ~> 2460.2 ~> 2362.4
10x
MS4: Synthetic phosphopeptide
10x
y*18 y*17 y*13 y12 y10
55 ISHS†PQQRNPNSKIPS†PVREK 75
c9 c10 b11 c14 c15 c16 b19 b20
c11 c19 c20

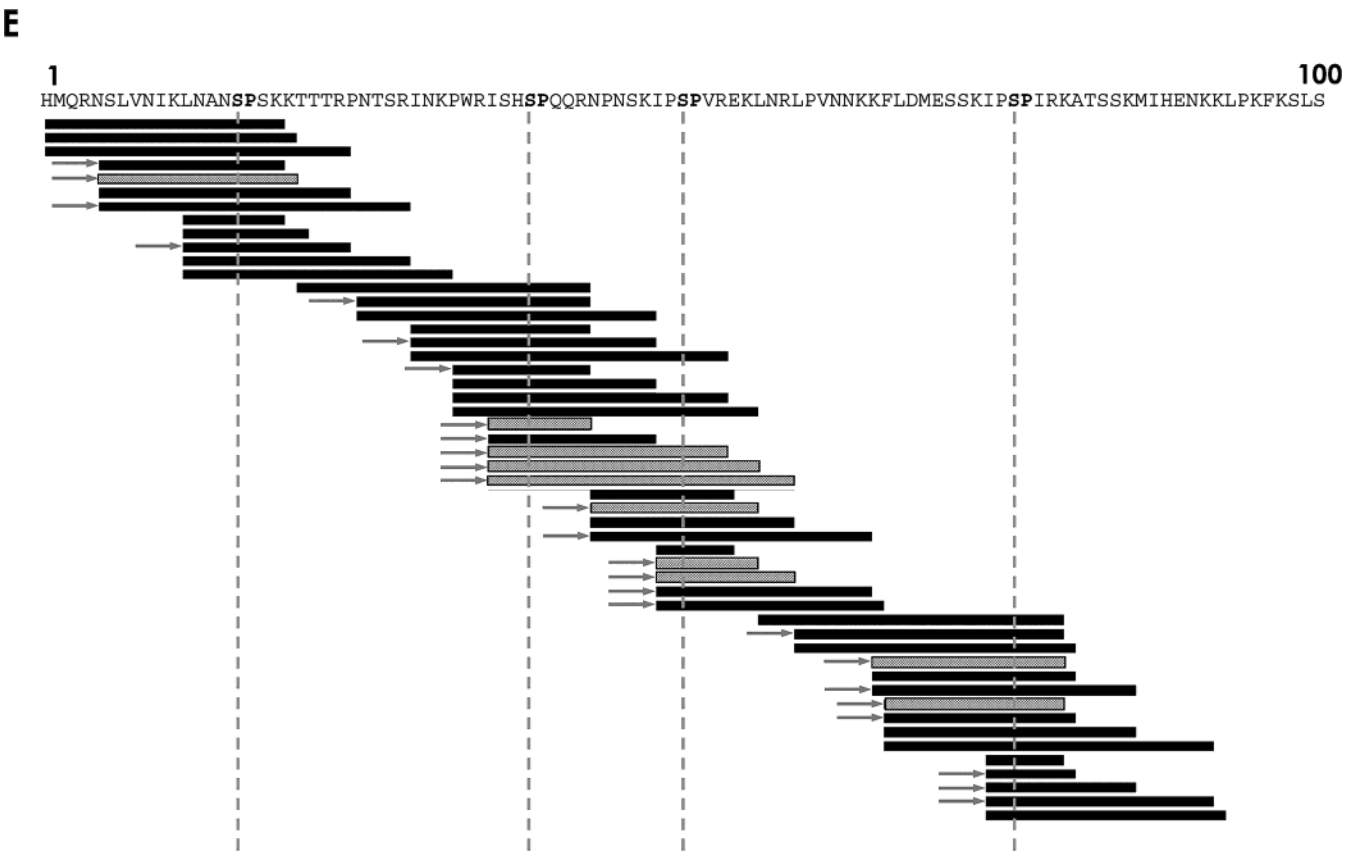


Figure 3. In vitro phosphorylation of NTD-Ipl1. (A) MALDI-linear-TOF mass spectrum of in vitro Cdk phosphorylated Ipl1 shows five distinct peaks spaced ~80 Da apart. (B) MALDI-QqTOF spectrum of tryptic digest of in vitro phosphorylated Ipl1 reveals four potential phosphopeptides, corresponding to only three of the four Cdk consensus sites. (C) HMS-MS detects phosphorylation of peptide 68−75, a tryptic phosphopeptide not seen in the single-stage spectrum (top panels). The synthetic phosphopeptide IPpSPVREK undergoes neutral loss of 98 Da in MS^2 (data not shown). The MS^3 spectrum of the peptide from the in vitro phosphorylated sample is essentially identical to the corresponding MS^3 spectrum of the synthetic peptide (bottom panel.) Peaks labeled with a pound sign (#) are not readily identifiable but nevertheless appear in the spectra of both the in vitro phosphorylated and synthetic peptides (D) MS^2, MS^3, and MS^4 are used to confirm the sequence of the doubly phosphorylated peptide 55−75 (top three panels). The corresponding synthetic phosphopeptide undergoes neutral loss of 98 Da in MS^2 and MS^3 (data not shown) and yields an MS^4 spectrum (bottom panel) that is essentially identical to that from the in vitro phosphorylated sample. Peaks labeled with a superscript ¶ appear to arise from a species that has lost water from a residue C-terminal to Ser 70. (E) Summary of HMS-MS data for NTD-Ipl1. The sequence is shown, with bars corresponding to the potential phosphopeptides that were subjected to MS^2. Arrows represent potential phosphopeptides for which −98 Da fragments were observed in MS^2. Light-colored bars represent phosphopeptides confirmed by MS^3. A total of 10 phosphopeptides were confirmed, corresponding to all four Cdk consensus sites.

peptide. HMS-MS detected all four phosphopeptides observed in the single-stage spectrum, and six additional phosphopeptides that were not detected in the single-stage MS spectrum.

The MS^2 and MS^3 spectra of one of these peptides (residues 68−75, IPpSPVREK) are shown in the top two panels of Figure 3C. The preferred y_6 fragmentation, C-terminal to a Pro residue in the second position, is characteristic of MALDI-ion trap fragmentation.[60] Noting a series of unusual fragment ions in the MS^3 spectrum (which correspond formally to the loss of C-terminal residues by hydrolysis), we obtained a synthetic version of the peptide IPpSPVREK and analyzed it by MS^2 and MS^3. The MS^2 spectrum was dominated by a peak 98 Da less than the parent (data not shown). As seen by comparing the bottom two panels of Figure 3C, the major peaks from the MS^3 spectrum of the in vitro phosphorylated tryptic peptide all overlap with the major peaks of the corresponding MS^3 spectrum from the synthetic peptide, including the series of "hydrolysis-loss" fragment peaks. In addition, the MS^3 spectrum of the purified, synthetic material contained several identifiable low-intensity peaks that were not readily observed in the corresponding spectrum from the in vitro phosphorylated material, as might be expected from the better ion statistics of the former.[60]

Although inspection of the MS^3 spectrum from the in vitro phosphorylated sample did not allow us to conclusively differentiate between "c ions" and the "hydrolysis-loss" fragments (a difference of 1 mass unit), the better statistics and mass accuracy (≤200 ppm) obtained from MALDI-ion trap MS^3 studies using 200 fmol synthetic peptide clearly demonstrate that these fragments arise from "hydrolysis-loss" products. Such "hydrolysis-loss" products in MALDI-ion trap MS^3 spectra of phosphorylated peptides have been observed previously (ref 29, Supporting Information).[29] Both "c ions" and "hydrolysis-loss" products have been observed in the MALDI-QqTOF MS^2 spectra of phosphorylated peptides by Standing and co-workers.[17] Intramolecular rearrangements of this sort have previously been observed in fast atom bombardment MS^2 experiments.[61] Similar fragmentation of

(60) Zhang, W.; Krutchinsky, A. N.; Chait, B. T. *J. Am. Soc. Mass Spectrom*. **2003**, *14*, 1012−1021.

the C-terminal residue from tryptic and partially tryptic peptides has also been observed during ion trap MS^2 fragmentation.[62,63]

Although the most straightforward explanation for the MS^3 fragmentation of the $(M + H - 98)^+$ species involves neutral loss of both phosphate and water from the phosphorylated residue, this assumption may not always be true.[17,20] The data in Figure 3C is consistent with a model in which the peptide loses 98 in MS^2 through two pathways. The first involves loss of phosphate and water from Ser in the third position, whereas the second involves loss of phosphate accompanied by loss of water from a C-terminal residue.(Figure S1). Further studies of the mechanism(s) of fragmentation are in progress.

Seven of the MS^3-confirmed phosphopeptides were singly phosphorylated, including one containing a site that could not be inferred from the single-stage mass spectrum. Three other phosphopeptides were observed to undergo a second loss of 98 Da in the MS^3 spectra, confirming that they were doubly phosphorylated. One of these doubly phosphorylated peptides (55−75, ISHpSPQQRNPNSKIPpSPVREK) yielded a sufficient ion count to confirm the sequence in MS^4 mode (Figure 3D, top three panels). MS^2 and MS^3 spectra of a synthetic version of this peptide also showed strong loss of 98 Da (data not shown). The major peaks from the MS^4 spectrum of the in vitro phosphorylated tryptic peptide also all overlapped with the major peaks from the corresponding MS^4 spectrum from the synthetic peptide (Figure 3D, bottom two panels). The MS^4 spectrum of the synthetic phosphopeptide again suggests at least partial loss of water from a nonphosphorylated residue and exhibits fragments that correspond formally to "c ions". The exact nature of these ions and the mechanism of their formation are current topics of investigation. To further investigate sensitivity limits, the two synthetic Ipl1-derived phosphopeptides were subjected to titrations similar to those described above for the standard phosphopeptides. These two Ipl1-derived phosphopeptides were detected by MS^2 and MS^3 down to 2 fmol, in both the absence and the presence of 500 fmol protein digest (data not shown).

As summarized in Figure 3E, hypothesis-driven MS^2 and MS^3 demonstrated that all four of the potential Cdk phosphorylation sites were indeed phosphorylated.

HMS-MS of In Vivo Phosphorylation of Orc6: Phosphopeptide Mapping of an In Vivo Phosphorylated Protein Expressed at Endogenous Levels. The origin recognition complex (ORC) is an essential heterohexameric protein assembly localized to origins of replication on DNA[51,64] which serve as a platform for assembly of the prereplicative complex.[65] This complex is necessary to ensure proper timing of DNA replication. Three components of ORC—Orc1, Orc2, and Orc6—contain at least one full Cdk consensus motif (i.e., [S/T]PX[K/R],[40] where X is any residue). Mutation of the Ser/Thr residues at the sites in Orc2 and Orc6 sensitizes the cells to inappropriate reinitiation of DNA replication.[65] Although the phosphorylation of Orc6, which has six minimal Cdk motifs (i.e., [S/T]P[40]), all located between residues 106−171, has been followed by change in migration of the protein band in SDS−PAGE,[51] the actual sites of Orc6 phosphorylation have not yet been directly observed. We applied phosphopeptide mapping by HMS-MS to Orc6 (expressed at endogenous levels) after protein A affinity purification from asynchronous yeast cells (Figure 4A). Orc6 copurified with the five other members of the Orc complex (Orc1−Orc5). Consistent with earlier work, Orc6 was identified in two bands with the slower-running band presumably containing hyperphosphorylated forms of Orc6. This slower running Orc6 band was excised for HMS-MS analysis. We hypothesized phosphorylation on all peptides containing the minimal Cdk motifs, allowing for up to three missed cleavages. A total of 57 m/z values were interrogated by MS^2. Twenty-three of these (corresponding to 17 tryptic amino acid sequences) exhibited a detectable loss of 98 Da, of which five yielded MS^3 spectra, allowing positive identification of the phosphopeptide. These five confirmed phosphopeptides contained four of the six candidate minimal Cdk sites (Ser 106, Thr 114, Ser 116, and Ser 174) in Orc6 (Figure 4B, Table S2).

An example of the process whereby we confirmed the phosphopeptides in Orc6 is given in Figure 4B for the hypothetical doubly phosphorylated peptide spanning residues 110−122 (at m/z 1746.6). MS^2 yielded a spectrum in which the neutral loss of 98 was the eighth most intense peak (excluding the residual precursor ion peak and the peak corresponding to the loss of water/ammonia from the precursor). Thus, most data-dependent scanning methods would have missed this signal. However, when hypothesis-driven MS^3 was performed on the m/z (1746.8−98) peak, the base peak of the nearly noise-free resulting spectrum corresponded to a further neutral loss of 98, indicating the loss of a second phosphate. Although the ion count was insufficient to perform MS^4 on this species, Orc6 contains no other tryptic peptides with possible multiple phosphorylation sites that would result in a mass close to 1746.8. Therefore, we conclude that the peptide 110−122 is indeed doubly phosphorylated at its only two Ser/Thr residues, Thr 114 and Ser 116. Both of these are proline-directed sites at which phosphorylation has not previously been directly observed.

HMS-MS of Phosphorylated PERIOD Protein: HMS-MS without the Constraint of a Consensus Motif. *D. melanogaster* protein PERIOD (PER) is a 1224-amino acid protein that plays a central role in the circadian rhythm cycle by regulating transcription of clock-controlled genes.[66,67] Several regions in *Drosophila* PER are highly enriched with serine and threonine residues as well as with acidic amino acids. One of these regions lies in the middle of the protein, between amino acid residues 536 and 640, with almost 30% of its residues being either Ser or Thr. This domain has affinity in vitro for several protein kinases, including casein kinase I (CKI), probably because of its acidic amino acid content.[68,69] We used HMS-MS to map the phosphorylation sites of a truncated construct of PER (Per 531, containing residues 531−700) phosphorylated in vitro with rat CKI. Because CKI does not have a strict primary sequence requirement for phosphorylation (although it shows a preference for nearby N-terminal acidic residues or phosphorylated Ser or Thr residues[68,69]), we hypoth-

(61) Ballard, K. D.; Gaskell, S. J. *J. Am. Chem. Soc.* **1992**, *114*, 64−71.
(62) Qin, J.; Fenyo, D.; Zhao, Y.; Hall, W. H.; Chao, D. M.; Wilson, C. J.; Young, R. A.; Chait, B. T. *Anal. Chem.* **1997**, *69*, 3995−4001.
(63) Qin, J.; Chait, B. T. *Int. J. Mass Spectrom.* **1999**, *190/191*, 313−320.
(64) Diffley, J. F.; Cocker, J. H.; Dowell, S. J.; Rowley, A. *Cell* **1994**, *78*, 303−316.
(65) Bell, S. P.; Dutta, A. *Annu. Rev. Biochem.* **2002**, *71*, 333−374.
(66) Young, M. W.; Kay, S. A. *Nat. Rev. Genet.* **2001**, *2*, 702−715.
(67) Young, M. W. *Trends Biochem Sci* **2000**, *25*, 601−606.
(68) Flotow, H.; Roach, P. J. *J. Biol. Chem.* **1991**, *266*, 3724−3727.
(69) Flotow, H.; Graves, P. R.; Wang, A.; Fiol, C. J.; Roeske, R. W.; Roach, P. J. *J. Biol. Chem.* **1990**, *265*, 14264−14269.

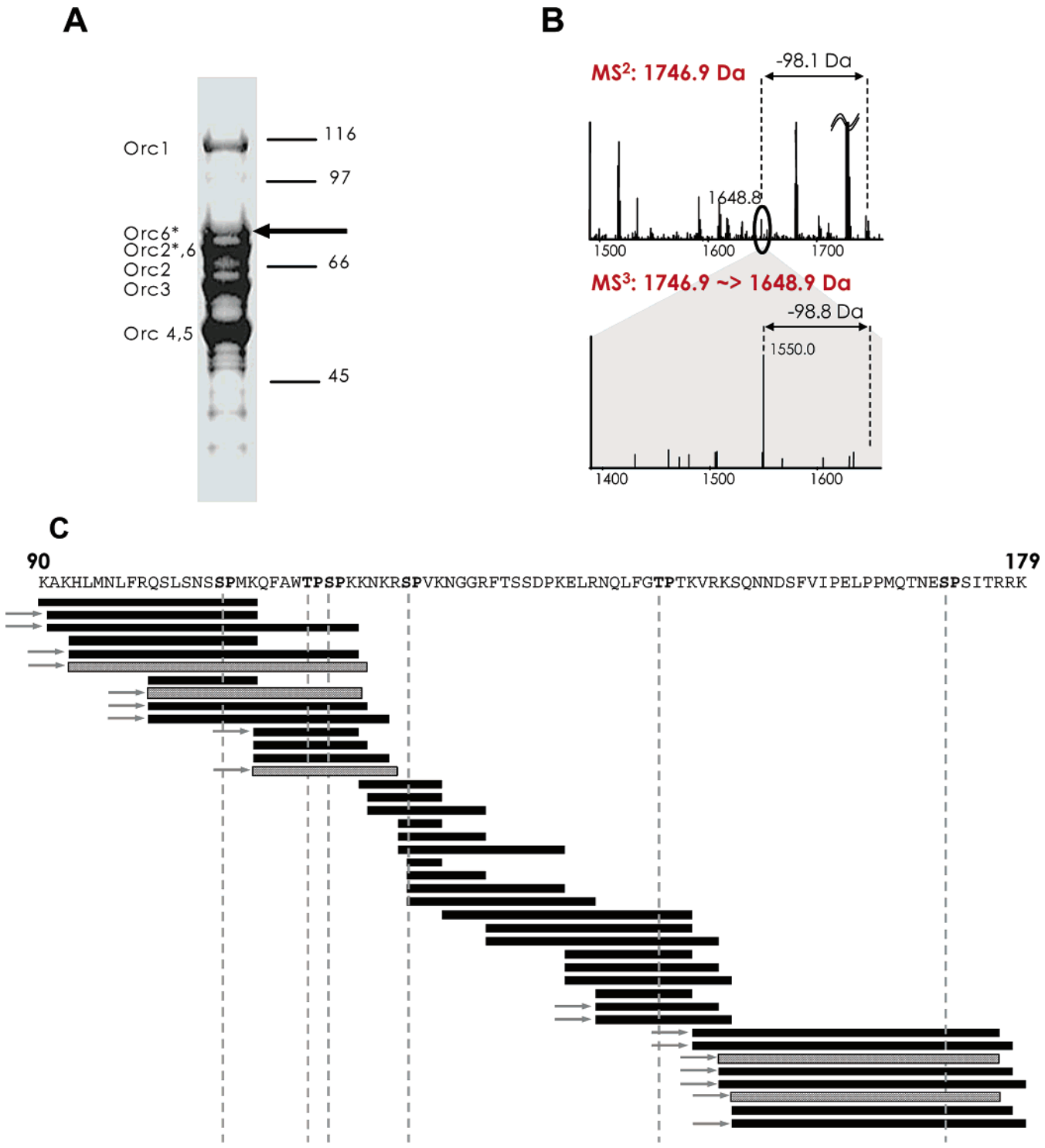


Figure 4. HMS-MS of in vivo phosphorylated Orc6 at Cdk consensus motifs. In vivo phosphorylation of Orc6. (A) SDS−PAGE (visualized by Coomassie blue staining) of proteins purified from an Orc6−PrA tagged yeast strain, by IgG affinity chromatography. All six members of the Orc complex are observed. The arrow indicates the band that was excised for HMS-MS studies. (B) HMS-MS analysis of the doubly phosphorylated peptide 110−122. Loss of 98 Da is observed successively in MS^2 and MS^3. (C) Summary of HMS-MS experiments. The sequence of the region of Orc6 containing the Cdk consensus motifs is shown, with bars corresponding to the potential phosphopeptides tested by MS^2. Arrows represent potential phosphopeptides for which −98-Da fragments were observed in MS^2. Light-colored bars represent phosphopeptides confirmed by MS^3.

esized possible phosphorylation at every Ser and Thr residue from Thr537 through Thr633. Thus, we hypothesized 38 phosphorylated amino acid sequences that resulted from the theoretical trypsin digest with up to four missed cleavages. Because many of these theoretical tryptic peptides contain multiple Ser/Thr residues, it was necessary to consider the possibility of multiple phosphorylation, yielding a total of 120 distinct hypothetical phosphopeptides masses with m/z accessible to our MALDI-ion trap mass spectrometer (m/z < 4000). Twenty-five of these yielded spectra exhibiting neutral loss of 98 Da and were further interrogated by

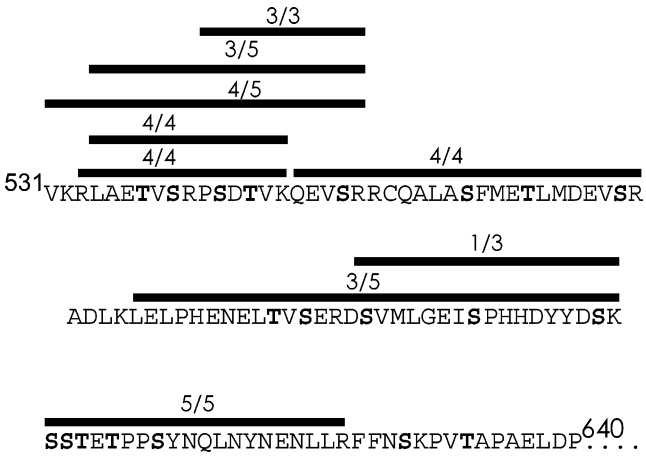


Figure 5. Summary of HMS-MS results for in vitro phosphorylated PERIOD 531 protein. The region of interest (PERIOD protein residues 531−640, corresponding to residues 10−111 in the truncated construct we analyzed) is shown, with Ser and Thr residues in bold. The bars represent sequences observed to be phosphorylated by HMS-MS, and the numbers above the bar indicate the highest number of phosphorylations observed over the total possible number of phosphorylations in the given sequence. At least 16 of the 20 Ser/Thr residues were observed to be phosphorylated.

MS^3. In some cases, when a −98 product was confirmed in MS^3, we also performed extended collection time MS^3 on other potential phosphoforms of the same sequence (to reduce the chance of missing low-abundance phosphoforms). In all, 36 different m/z values were interrogated by MS^3, and 23 gave fragmentation data consistent with their corresponding hypothesized peptides (Table S3). The analysis indicates that at least 16 of 20 possible residues in the region of interest are phosphorylated. (Figure 5). Each of the Ser/Thr residues in the region that we determined to be phosphorylated (except for Ser 558) has, within three residues to its N-terminal side, either an acidic or a possibly phosphorylated Ser/Thr residue. This finding is consistent with the rather promiscuous CKI primary sequence preference. Thus, the nearly comprehensive phosphorylation of this region of PER indicates that under the in vitro conditions used, at least trace phosphorylation occurred at nearly every possible site analyzed. The artificial combination of the rat kinase with the *Drosophila* substrate, combined with the use of an extended reaction time and an excess of enzyme, likely relaxes the specificity of the enzyme (personal communication, Saul Kivimäe) with respect to the true in vivo specificity, resulting in the observed nonspecific in vitro phosphorylation pattern. The phosphorylation of this PER construct demonstrates that HMS-MS is a useful method for analysis of phosphorylation, even in the absence of an assumed consensus motif and can approach comprehensive phosphopeptide determination.

CONCLUSION

In this work, we have demonstrated the ability of HMS-MS to detect femtomole amounts of Ser/Thr phosphorylated peptides using vacuum MALDI-ion trap mass spectrometry. Few false positives are expected, because the approach is essentially a coincidence experiment, requiring the observation of both the neutral loss of 98 Da and MS^3 fragmentation data that confirm the hypothesized phosphopeptide sequence. This high sensitivity and selectivity for the detection of phosphorylated peptides arises from the characteristics of our in-house-assembled vacuum MALDI-ion trap mass spectrometer. Singly phosphorylated phosphopeptides undergo preferential neutral loss of 98 Da during resonant excitation in the ion trap, providing a clear signature for the presence of the phospho-Ser/Thr moiety. The ability to readily perform MS^3 on the resulting $(M + H - 98)^+$ ion species (or MS^4 on the $(M + H - 98-98)^+$ ion species) allows us to confirm the identity of the putative phosphopeptides with high confidence, even when the signature is weak or where the MS^2 spectrum is complicated by the presence of multiple peptides.

Inasmuch as the MALDI HMS-MS experiment relies on the neutral loss of H_3PO_4 as an assay for phosphorylation, it is similar to a conventional neutral-loss linked scan experiment. However, because the HMS-MS experiment interrogates only the masses of hypothesized peptides instead of the entire mass range, losses associated with scanning the mass spectrometer are reduced accordingly. In addition, the use of MALDI has the important characteristic that the sample can be interrogated until it is completely depleted from the sample probe. Because the present MALDI-ion trap mass spectrometer has relatively high efficiency, it is usually feasible to test more than 100 different hypotheses from a single sample. Thus, the situation is different from ESI of peptides eluting from an LC column, where the time for analysis is limited to the elution period. In addition, weakly responding peptides can be interrogated for extended periods of time in order to obtain better statistics, and because solid MALDI sample depositions are stable, multiple rounds of data collection can be performed over a period of days or even weeks.

Although MALDI HMS-MS proves an effective means for detecting phosphorylated peptides, it is limited in its ability to pinpoint specific sites of phosphorylation when more than one potential site is present on the phosphopeptide. This limitation is a result of the highly preferred loss of the elements of H_3PO_4 from singly protonated phosphopeptides upon resonant excitation in the ion trap mass spectrometer and because the loss of H_2O moieties does not always occur from the same amino acid residue as does the loss of the HPO_3 moiety. In such cases, other means (e.g., ECD[70,71] or CID[20,72] of multiply protonated ions formed by ESI) are required to pinpoint the site of phosphorylation. In any event, knowledge of the m/z and amino acid sequence of the target phosphopeptide greatly facilitates such studies.

The described applications of phosphorylation site mapping by HMS-MS demonstrate the applicability of the method to both in vitro and in vivo phosphorylated proteins and to proteins digested either in gel, in solution, or directly on the surface of a polycarbonate MALDI probe. Additional applications to the elucidation of in vivo phosphorylation of Cdk substrates in *S. cerevisiae* can be found in ref 73.[73] Because the technique can be used to systematically survey the entire length of proteins in a

(70) Stensballe, A.; Jensen, O. N.; Olsen, J. V.; Haselmann, K. F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1793−1800.
(71) Shi, S. D.; Hemling, M. E.; Carr, S. A.; Horn, D. M.; Lindh, I.; McLafferty, F. W. *Anal. Chem.* **2001**, *73*, 19−22.
(72) Zarling, A. L.; Ficarro, S. B.; White, F. M.; Shabanowitz, J.; Hunt, D. F.; Engelhard, V. H. *J. Exp. Med.* **2000**, *192*, 1755−1762.
(73) Archambault, V.; Chang, E. J.; Drapkin, B. J.; Cross, F. R.; Chait, B. T.; Rout, M. P. *Mol. Cell* 2004, in press.

sensitive and specific manner for Ser/Thr phosphorylation, it represents a significant step toward our goal of comprehensive phosphorylation mapping.

ACKNOWLEDGMENT

E.J.C. is a Burroughs-Welcome fellow. DTM was supported by a postdoctoral fellowship from the Canadian Institutes of Health Research. Thanks to Markus Kalkum for the computer program Look4Loss, to Zu-wen Sun and David Allis (Rockefeller University) for plasmid expressing His10-NTD-Ipl1, to the Rockefeller University Proteomics Resource Center for Ipl1-derived synthetic phosphopeptides, and to Saul Kivimäe and Michael Young (Rockefeller University) for phosphorylated PERIOD protein. This work was funded by NIH grants RR00862 and CA89810 to B.T.C.. This work was presented in part in the Proceedings of the 51st ASMS Conferences on Mass Spectrometry and Allied Topics, June 8 -12, 2003, Montreal, Canada.

SUPPORTING INFORMATION AVAILABLE

Supporting tables S1, S2, and S3 contain summarized HMS-MS data for Ipl1, Orc6, and PER, respectively. Supporting figure S1 contains a detailed interpretation of the MS^3 spectrum of the synthetic peptide IPpSPVREK and a proposed model for interpreting the observed fragmentation pattern. This material is available free of charge via the Internet at http://pubs.acs.org.

Received for review March 6, 2004. Accepted May 6, 2004.

AC049637H