Comprehensive Identification of Phosphorylation Sites in Postsynaptic Density Preparations*[S]

Jonathan C. Trinidad‡, Christian G. Specht§¶, Agnes Thalhammer§, Ralf Schoepfer§‖, and Alma L. Burlingame‡**

In the mammalian central nervous system, the structure known as the postsynaptic density (PSD) is a dense complex of proteins whose function is to detect and respond to neurotransmitter released from presynaptic axon terminals. Regulation of protein phosphorylation in this molecular machinery is critical to the activity of its components, which include neurotransmitter receptors, kinases/phosphatases, scaffolding molecules, and proteins regulating cytoskeletal structure. To characterize the phosphorylation state of proteins in PSD samples, we combined strong cation exchange (SCX) chromatography with IMAC. Initially, tryptic peptides were separated by cation exchange and analyzed by reverse phase chromatography coupled to tandem mass spectrometry, which led to the identification of phosphopeptides in most SCX fractions. Because each of these individual fractions was too complex to characterize completely in single LC-MS/MS runs, we enriched for phosphopeptides by performing IMAC on each SCX fraction, yielding at least a 3-fold increase in identified phosphopeptides relative to either approach alone (SCX or IMAC). This enabled us to identify at least one site of phosphorylation on 23% (287 of 1,264) of all proteins found to be present in the postsynaptic density preparation. In total, we identified 998 unique phosphorylated peptides, mapping to 723 unique sites of phosphorylation. At least one exact site of phosphorylation was determined on 62% (621 of 998) of all phosphopeptides, and ~80% of identified phosphorylation sites are novel. ***Molecular & Cellular Proteomics 5: 914–922, 2006.***

Protein phosphorylation has emerged as a fundamental mechanism for the modulation of protein function. It has been estimated that as many as 10,000 or about one-third of all cellular proteins may be phosphorylated (1). The postsynaptic density (PSD)[1] is a specialized protein complex for the transduction and modulation of signaling between neurons, and many of its components are regulated by phosphorylation events (2).

Several recent studies have used mass spectrometric approaches to identify between 200 and 500 proteins in PSD preparations (3–7). However, the identification of post-translational modifications, such as phosphorylation, from such complex mixtures is complicated by their relatively low stoichiometry at specific amino acid residues. For example, one of the previously mentioned studies identified 13 phosphopeptides by manual inspection of 1,830 MS/MS spectra (5).

Therefore, efforts have focused on specific enrichment methods of post-translational modifications for their improved detection from complex mixtures. Immunoaffinity separation has been used most effectively to analyze phosphorylation of tyrosine residues from a variety of cell types due to the availability of robust anti-phosphotyrosine antibodies (8–10). IMAC has been successfully used for the identification of phosphopeptides in PSD (7, 11) and synaptosome preparations (12, 13). These initial studies identified between 16 and 331 total phosphorylation sites on synaptic proteins using the esterification of carboxyl groups to reduce unspecific (non-phosphate) binding to the IMAC resin (14, 15). An alternative strategy takes advantage of the earlier elution of phosphopeptides during strong cation exchange (SCX) chromatography compared with non-phosphorylated peptides of similar compositions due to the negative charge of the phosphate group (16, 17). In addition, several groups have begun to combine various ion exchange approaches with IMAC to further enrich phosphopeptide content of LC-MS/MS fractions (18, 19). However, the relative performance of the combined ion exchange-IMAC approaches has not been directly compared.

In the present study, we combined SCX chromatography to fractionate a complex protein digest with IMAC to enrich for

From the ‡Mass Spectrometry Facility, Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143 and §Laboratory for Molecular Pharmacology, Department of Pharmacology, University College London, Gower Street, London WC1E 6BT, United Kingdom

Received, November 23, 2005, and in revised form, February 1, 2006

Published, MCP Papers in Press, February 1, 2006, DOI 10.1074/mcp.T500041-MCP200

[1] The abbreviations used are: PSD, postsynaptic density; CaMKII, calcium/calmodulin-dependent protein kinase II; p-CaMKII, phosphorylated CaMKII; CDK5, cyclin-dependent kinase 5; GluR1, ionotropic glutamate receptor subunit 1; SCX, strong cation exchange; Bicine, *N*,*N*-bis(2-hydroxyethyl)glycine.

phosphopeptides from individual SCX fractions, leading to a 3–10-fold improvement in phosphopeptide enrichment over either of the two techniques alone. Using this workflow, we obtained a comprehensive dataset of 1,264 unique proteins present in preparations of PSD fractions. We identified a total of 998 unique phosphorylated peptides mapping to 723 unique sites on 287 total phosphoproteins. This translates to ~0.58 phosphorylation sites per unique protein in our mixture (728 of 1,264), suggesting that this analysis resulted in substantial coverage of the total postsynaptic density phosphoproteome. Finally in an attempt to gain further insight into the proposed relationship between phosphorylation and *O*-linked *N*-acetylglucosaminylation, we have developed methodology to enable assignment of this modification as well (see accompanying paper, Vosseller *et al.* (24)).

MATERIALS AND METHODS

PSD Sample Preparation and Quality Control—The purification of PSD samples was performed at 4 °C as described previously (7). Protease and phosphatase inhibitors were used throughout except in the final pelleting step. Briefly adult mouse brains (2–16 months, 20 animals, 10 g wet weight) were homogenized in 40 mM Tris, pH 7.5, 0.3 M sucrose containing EDTA-free protease inhibitors (Complete, Roche Applied Science) and a mixture of phosphatase inhibitors (1 mM Na_3VO_4, 1 mM NaF, 1 mM Na_2MoO_4, 4 mM sodium tartrate, 100 nM fenvalerate, 250 nM okadaic acid) and cleared by centrifugation. The membranous fraction was separated from the cytosolic fraction at 17,000 × *g* for 20 min followed by sucrose density centrifugation (120 min at 85,000 × *g*; 0.8, 1.0, and 1.2 M sucrose step gradient). Synaptic membranes were collected at the 1.0–1.2 M interface, pelleted, homogenized in 10 mM Bicine, pH 7.5, containing 5% *N*-octyl-β-D-glucopyranoside (Calbiochem), and applied onto a second gradient (120 min at 85,000 × *g*; 1.0, 1.4, and 2.2 M sucrose). The PSD fraction was collected at the 1.4–2.2 M interface and pelleted in 10 mM Bicine, pH 7.5, in the absence of protease and phosphatase inhibitors. Yields of PSD samples were generally in the range of 0.1% of the wet weight tissue used.

Western blotting was performed as published previously (7) with the following primary antibodies: goat anti-NR2A (1:500, sc1468, Santa Cruz Biotechnology, Santa Cruz, CA); mouse anti-NR2B (1:1,000, N38120, BD Transduction Laboratories, San Jose, CA); mouse anti-NR1 (1:500, 556308 clone 54.1, BD Transduction Laboratories); mouse anti-PSD-95 (1:250, P43520, BD Transduction Laboratories); rabbit anti-GluR1 (1:1,000, AB1504, Chemicon International, Temecula, CA); goat anti-SynGap (1:1,000, sc-8572, Santa Cruz Biotechnology); mouse anti-CaMKIIα (1:1,000, MAB8699 clone 6G9, Chemicon International) and anti-phospho(Thr-286)-CaMKIIα (1:1,000, ab2724 clone 22B1, Abcam Ltd., Cambridge, UK), both kindly provided by Peter Giese; goat anti-synaptophysin (1:200, sc7568, Santa Cruz Biotechnology); and mouse anti-α-synuclein (1:1,000, S63320, BD Transduction Laboratories).

Digestion of PSD Samples—6 mg of PSD proteins were resuspended in 2 ml of 25 mM ammonium bicarbonate containing 6 M guanidine hydrochloride. The mixture was incubated for 1 h at 57 °C with 2 mM Tris(2-carboxyethyl)phosphine hydrochloride to reduce cysteine side chains; these side chains were then alkylated with 4.2 mM iodoacetamide in the dark for 45 min at 21 °C. The mixture was diluted 6-fold with 25 mM ammonium bicarbonate, and 5% (w/w) modified trypsin (Promega, Madison, WI) was added. The pH was adjusted to 8.0, and the mixture was digested for 12 h at 37 °C. The digest was desalted using a C_{18} Sep-Pak cartridge (Waters, Milford, MA).

Strong Cation Exchange Chromatography—SCX chromatography was performed using an ÄKTA Purifier (Amersham Biosciences) equipped with a Waters AP minicolumn packed in house with 5 × 115-mm polysulfoethyl A resin (Western Analytical, Lake Elsinore, CA). 6.0 mg of PSD sample digest were loaded onto the column in 30% acetonitrile, 5 mM KH_2PO_4, pH 2.7. A 90-min gradient was run from 0 to 350 mM KCl in 30% acetonitrile, 5 mM KH_2PO_4, pH 2.7. 88 fractions were collected and desalted using a MAX-RP reverse phase C_{18} cartridge (Phenomenex, Torrance, CA) and dried down using a SpeedVac concentrator (Thermo Electron, San Jose, CA). 5% of each fraction was reserved for analysis using ESI-Qq-TOF tandem mass spectrometry, and the remaining 95% was subjected to IMAC enrichment when appropriate.

Nano-LC-ESI-Qq-TOF Tandem Mass Spectrometry Analysis—Samples corresponding to 40% of each previously reserved aliquot (or 2% of each original SCX fraction) were separated using a 75-μm × 15-cm reverse phase C_{18} column (LC Packings, Sunnyvale, CA) at a flow rate of 350 nl/min, running a 3–32% acetonitrile gradient in 0.1% formic acid on an Agilent 1100 series HPLC system equipped with an autosampler (Agilent Technologies, Palo Alto, CA). Gradient cycle times were between 1.0 and 2.0 h in length. The LC eluent was coupled to a microion spray source attached to a QSTAR Pulsar mass spectrometer (Applied Biosystems/MDS Sciex, Foster City, CA). Peptides were analyzed in positive ion mode. MS spectra were acquired for 1 s. For each MS spectrum, the two most intense multiple charged peaks were selected for generation of subsequent CID mass spectra. The CID collision energy was automatically adjusted based upon peptide charge and *m/z* ratio. A dynamic exclusion window was applied that prevented the same *m/z* from being selected for 3 min after its initial acquisition.

Immobilized Metal Affinity Chromatography—IMAC purification was performed on the remainder of each SCX fraction using an ÄKTA Purifier. A polyetheretherketone column (500 μm × 8 cm) was in-house packed with POROS iminodiacetic acid resin (PerSeptive Biosystems, Framingham, MA). Flow rates for injections and between injections were 50 and 500 μl/min, respectively, with 0.1% TFA (pH 1.9) used as the mobile phase. The injection scheme was as follows: column equilibration, 1 ml; $FeCl_3$ injection, 100–200 μl of 50 mM $FeCl_3$; column equilibration, 1 ml; peptide injection, individual, desalted SCX fractions in 250 μl of 0.1% TFA; column equilibration, 500 μl; wash injection, 300 μl of 100 mM NaCl in 30% acetonitrile, 0.1% TFA, pH 1.7; column equilibration, 500 μl; and phosphate elution, 200 μl of 100 mM Na_2HPO_4 with 100 mM NaCl in 30% acetonitrile, 1% TFA, pH 2.9. The phosphate elution fractions were dried down using vacuum centrifugation, resuspended in 0.1% formic acid, and desalted with an OMIX C_{18} tip (Varian, Walnut Creek, CA). The peptides were reconstituted in 6 μl of 0.1% formic acid, and 1 μl was run using nano-LC-ESI-Qq-TOF MS/MS (see above). An additional 1-μl aliquot was run using an exclusion list of peptides fragmented in the first run.

Interpretation of MS/MS Spectra—Data were analyzed using Analyst QS software (version 1.1), and MS/MS centroid peak lists were generated using the Mascot.dll script (version 1.6b16). MS/MS centroid peaks were thresholded at 0% of the base peak. Data were initially searched against the Uniprot rodent database (47,398 entries) as of April 2005 using both Mascot (version 2.0, Matrix Science, Boston, MA) and Protein Prospector 4.15 (University of California, San Francisco, CA). Initial peptide tolerance in MS and MS/MS modes were 200 ppm and 0.2 dalton, respectively. Trypsin was designated as the protease, and up to three missed cleavages were allowed. Carbamidomethylation was searched as a fixed modification, whereas oxidation of methionine, amino-terminal protein acetylation, and phosphorylation of serine/threonine/tyrosine residues were allowed as variable modifications. The high scoring peptide matches from individual LC-MS/MS runs were then used to internally

recalibrate MS parent ion *m/z* values within that run. This resulted in an internally consistent mass accuracy with an average standard deviation of 10–15 ppm per LC-MS/MS run. Recalibrated data files were then searched with a peptide tolerance in MS mode of 50 ppm. A protein was considered to be positively identified if the following conditions were met when the data was searched against the Uniprot rodent database. If the protein was identified on the basis of one peptide, we set a conservative threshold to minimize false positive protein identifications. We set the theoretical probability of obtaining a false positive protein identification to less than one in one thousand, which meant that these peptides must have a (false positive) Mascot expectation value ≤0.001. If the protein was identified on the basis of two or more peptides, the best matched peptide for each protein must have a Mascot expectation value ≤0.01. For protein identification purposes, only peptides with a Mascot Peptide Score ≥25 were examined. In the case where a given peptide could have originated from more than one protein, those proteins were considered positively identified as long as they were identified using at least one unique peptide.

Evaluation of Phosphopeptide MS/MS Spectra—For more details on the evaluation of phosphopeptide spectra, see "Results." Seven phosphopeptides were identified matching to seven proteins for which no additional non-phosphorylated peptides were observed. These peptides had Mascot expectation values ≤0.001 when searched against the Uniprot rodent database using the previous search parameters; however, this time up to four missed cleavages were allowed. The remaining 991 phosphopeptides matched to proteins for which non-phosphopeptides were also observed. These peptides had Mascot expectation values ≤0.05 when searched against the subset of proteins identified in this study using the previous search parameters. All phosphopeptides were manually inspected to verify that the majority of high abundance peaks were y or b sequence ions, or y − H_2O/H_3PO_4 or b − H_2O/H_3PO_4 ions when appropriate. Phosphorylation site assignment was performed manually as well as automatically calculated using a Perl script described in a separate study.[2]

RESULTS

PSD fractions were isolated from mouse brain using a two-step sucrose density gradient centrifugation protocol, which includes limited solubilization by the non-ionic detergent *N*-octyl-β-D-glucopyranoside. The relative enrichment of PSD proteins throughout the purification was assessed by Western blot analysis (Fig. 1). The glutamate receptor subunits NR2A, NR2B, NR1, and GluR1 were selectively enriched in the final PSD samples together with other known postsynaptic components, PSD-95 and SynGap. In contrast, two proteins located in the presynaptic terminal, namely synaptophysin and α-synuclein, were not detected in the PSD fraction, confirming that our purification protocol excludes certain presynaptic components. CaMKIIα, known to be present in both neuronal cytoplasm as well as in the PSD was detected in fractions throughout the preparation as was its main (activated) autophosphorylated form (p-CaMKIIα Thr-286).

We developed an automated workflow that allowed us to characterize both the protein components of our purified PSD sample as well as to identify sites of phosphorylation within these proteins (Fig. 2). The PSD sample was digested with trypsin, desalted, and separated into 88 fractions using SCX chromatography (Fig. 3*A*). An aliquot of each fraction was analyzed using nano-reverse phase LC-MS/MS on a QSTAR Pulsar mass spectrometer. Interpretation of the spectra using Mascot resulted in 17,391 positive peptide identifications of which 9,236 were to non-redundant peptide sequences. Proteins were considered to be positively identified using the following criteria. If only one peptide was used for the protein identification (214 instances), that peptide must have a Mascot expectation value less than or equal to 0.001 when searched against the Uniprot rodent database. If a protein was identified using two or more peptides (1,050 instances),

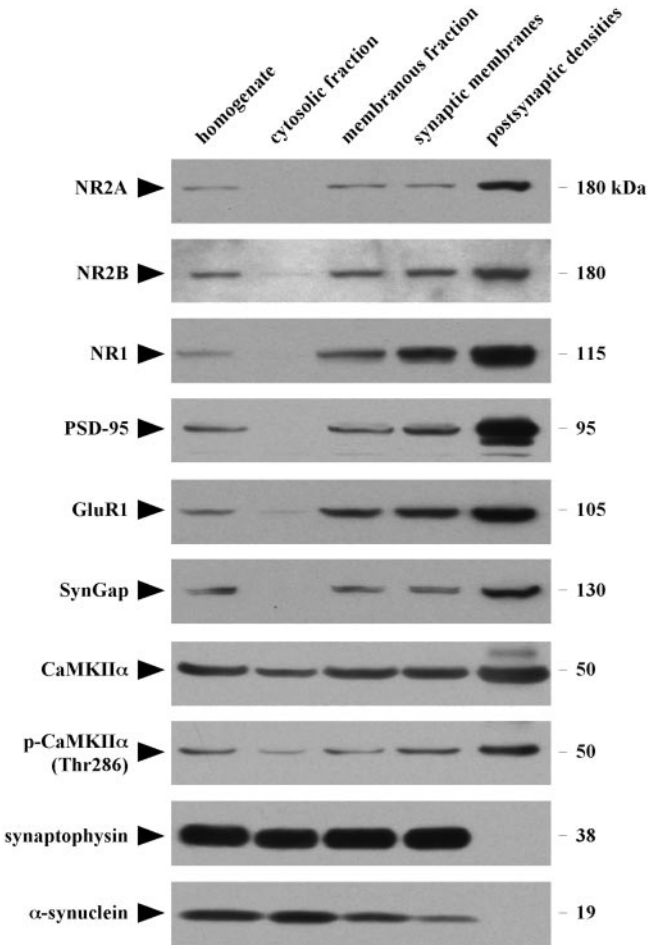

FIG. 1. **Quality control of the purified PSD samples by Western blotting.** Probing for the *N*-methyl-D-aspartate receptor subunits NR2A, NR2B, and NR1 as well as the α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor subunit GluR1 shows their enrichment in the final PSD fraction. Two additional synaptic proteins, PSD-95 and SynGap, are also found to be enriched in the PSD fraction. The presynaptic proteins synaptophysin and α-synuclein are present in the synaptic membrane fractions but lost during the second sucrose gradient. Probing for CaMKIIα subunits reveals their presence in all fractions. The detection of autophosphorylated CaMKIIα (Thr-286) confirms results obtained through mass spectrometric analysis of the PSD phosphoproteome.

[2] J. C. Trinidad, J. O. Snedecor, and A. L. Burlingame, manuscript in preparation.

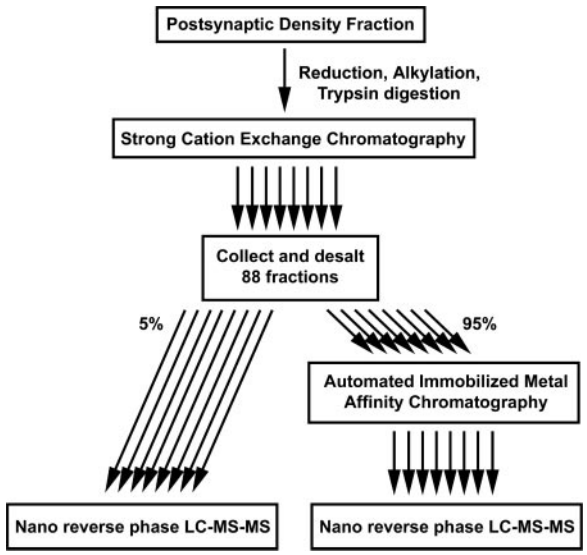

FIG. 2. **Schematic representation of the experimental workflow.** The purified PSD fraction was digested, and the tryptic peptides were fractionated using SCX chromatography. Aliquots of each fraction were analyzed by LC-MS/MS. In parallel, the remaining material of each fraction was subjected to IMAC for phosphopeptide enrichment and subsequently analyzed by LC-MS/MS.

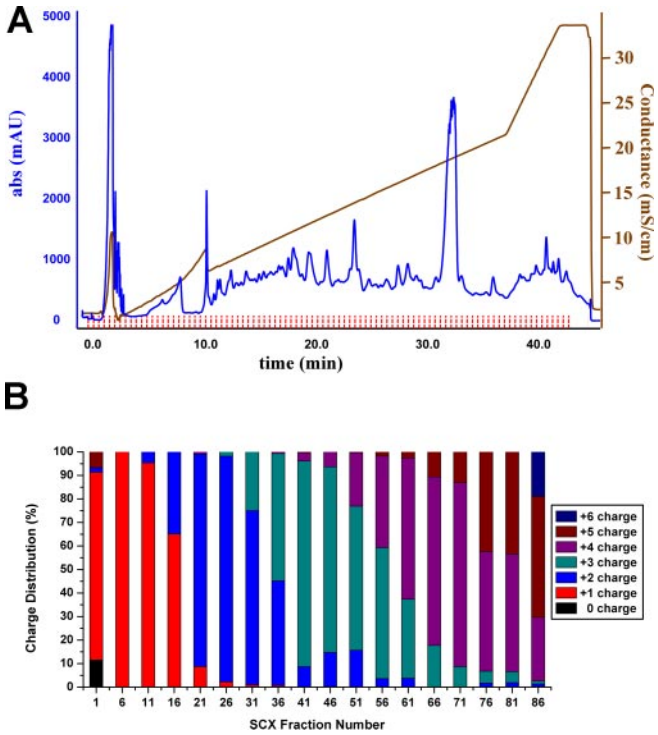

FIG. 3. **Identification of protein components of purified PSD samples.** *A*, tryptic peptides were subjected to SCX chromatography. Elution with a KCl gradient (conductance is shown in *brown*) was detected by absorbance at 215 nm (*blue trace*), and 92 fractions were collected (*red bars*). *B*, the peptide net charge varies as a function of the SCX fraction number. Unmodified tryptic peptides were given a +2 charge (protonated amino-terminal amino group and one arginine or lysine amino side chain). Additional arginine or lysine residues (due to missed tryptic cleavages) as well as His residues resulted in a unit increase in net charge. Protein amino-terminal acetylation, peptide amino-terminal pyroglutamine formation, and phosphorylation each resulted in a unit decrease in net charge. *mAU*, milliabsorbance units; *mS*, millisiemens.

the best peptide match must have a Mascot expectation value less than or equal to 0.01. By these criteria, 1,264 proteins were identified in the PSD preparation with a median sequence coverage of 17% (Supplemental Table S1).

The presence of a negatively charged phosphate group reduces the net charge of a phosphopeptide relative to a non-phosphorylated peptide of similar composition and causes phosphorylated peptides to elute earlier in SCX gradients than their non-phosphorylated counterpart. Accordingly a large number of phosphopeptides have been reported to elute in the initial phase of an SCX gradient (16, 17). To determine the extent that net charge influences SCX elution behavior in our system, we calculated the distribution of peptide net charge as a function of elution fraction across the SCX gradient (Fig. 3*B*). These data show that, although net charge clearly influences peptide elution behavior, it is not the sole determinant. For example, individual peptides with a +3 or +4 net charge in fraction 46 eluted earlier than a subset of peptides with a +2 net charge from fraction 51. A plot of SCX fraction *versus* net charge for the PSD digest gave a correlation coefficient of 0.87 (data not shown). Importantly the relatively large elution window for all peptides of a given charge was the result of a large range of individual peptide elution times rather than wide elution widths for individual peptides. The half width of elution for individual peptides was in fact much smaller than the fraction volume. A tryptic digest of α HS glycoprotein was run using identical gradient conditions, and the reduced sample complexity allowed us to measure individual peptide elution peaks by UV absorbance (data not shown). Individual peptides eluted with an average half width of 0.22 ml in contrast to the fraction size of 0.5 ml. After analyzing all 88 PSD SCX fractions, we identified a total of 311 phosphorylated peptides. Although there was a clear bias toward early elution times for phosphorylated peptides, they were in fact detected through most of the SCX chromatograph (Fig. 4, *A* and *B*). A large number (135) of phosphorylated peptides were observed in fraction 19, which corresponds to fully tryptic peptides (*i.e.* no missed cleavages) lacking histidine and containing a single site of phosphorylation (16, 17). These phosphopeptides in fraction 19 eluted two to three fractions before the bulk of peptides with a net +2 charge. Fractions 1–18 contained a large number of non-phosphorylated peptides with a +1 net charge. The vast majority of these peptides were either peptides derived from *N*-acetylated protein amino termini, peptides derived from protein carboxyl termini, or peptides with a glutamine at the amino terminus that had undergone cyclization to form pyroglutamine (19). These peptides had both a net +1 charge and a total positive charge of +1 and in general eluted earlier than

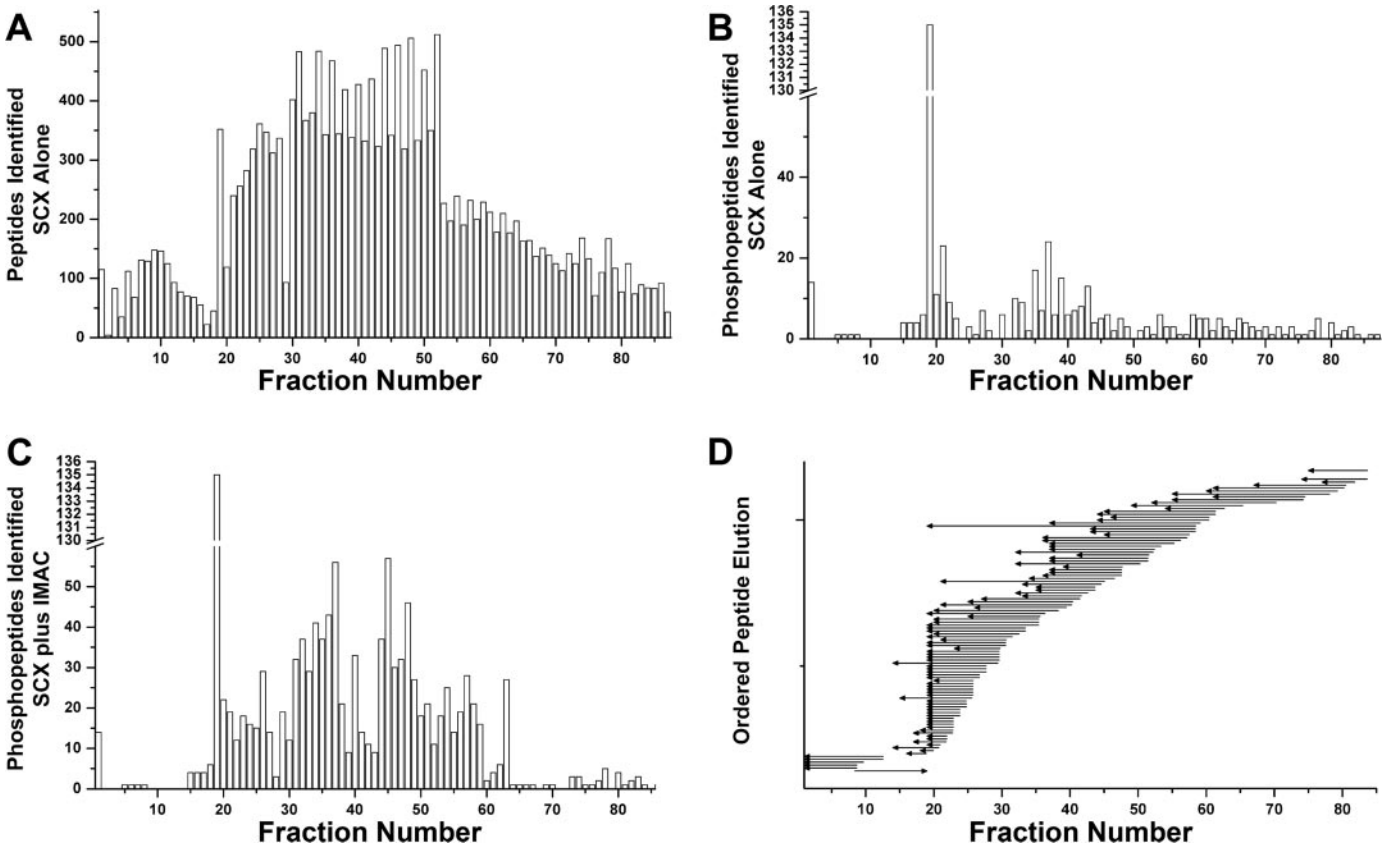

FIG. 4. **Improved identification of phosphopeptides by IMAC enrichment.** *A*, histogram showing the number of peptides identified by LC-MS/MS as a function of SCX fraction. *B*, subset of phosphorylated peptides identified in the SCX fractions from *A* without prior to IMAC enrichment. *C*, SCX fractions 24–88 were subjected to IMAC followed by LC-MS/MS. IMAC enrichment resulted in an ~3-fold increase in the total number of phosphopeptides identified. *D*, the subset of phosphopeptides whose non-phosphorylated counterparts were also observed are plotted in increasing fraction order in which the non-phosphorylated peptide was first observed. The *tails* of the *arrows* correspond to individual non-phosphopeptides observed in that specific fraction. The *heads* of the *arrows* correspond to the specific fraction in which the corresponding phosphorylated peptides were observed, and the *length* of the *shaft* therefore shows the extent to which the addition of phosphate shifted the elution behavior.

the phosphopeptides in fraction 19, which had a net +1 charge and a total positive charge of +2.

Manual inspection of MS survey scans from the majority of LC-MS/MS runs suggested that due to the complexity of individual SCX fractions MS/MS sequences were not being obtained on every peptide (and hence phosphopeptide) present above the MS/MS acquisition threshold. Particularly for fractions 24 and greater (from which a total of over 100 phosphopeptides were initially found), it was clear these fractions were being incompletely characterized. Analysis of each SCX fraction by LC-MS/MS showed that the majority of *observed* phosphopeptides eluted early in the SCX gradient (Fig. 4*B*) in agreement with previous observations by others (16, 17). However, because SCX chromatography separates in large part based upon net charge, phosphopeptides with a net +1 charge will be selectively enriched in isolation from other peptides and therefore will be detected at an increased rate. In contrast, phosphopeptides with a >+1 net charge will elute in fractions containing many high abundance non-phosphopeptides, and therefore this subpopulation of phosphopeptides will be detected at a decreased rate. It is not necessarily the case that the phosphopeptide distribution observed from the analysis of the SCX fractions accurately represents the total distribution of phosphopeptides in the mixture. Therefore, we developed an automated HPLC-based IMAC enrichment protocol to further enrich fractions 24–88 for phosphopeptides (Fig. 5*A*). A UV spectrum of a representative IMAC enrichment run is shown in Fig. 5*B*. This approach resulted in an ~3-fold increase (to a total of 998) in the total number of phosphopeptides identified (Fig. 4*C*). Importantly SCX fraction 19, which contained the most phosphopeptides of any fraction, *only* accounted for 14% of the total phosphopeptides we identified. In our analysis, we identified 105 pairs of peptides representing a singly phosphorylated peptide and its corresponding non-phosphorylated counterpart. We used these peptides to determine what effect addition of a phosphate moiety has on peptide elution during our SCX gradient. Fig. 4*D* shows these data ordered by the fraction in which the non-phosphorylated peptide was first observed. *Arrows* show the direction and magnitude of shift where the

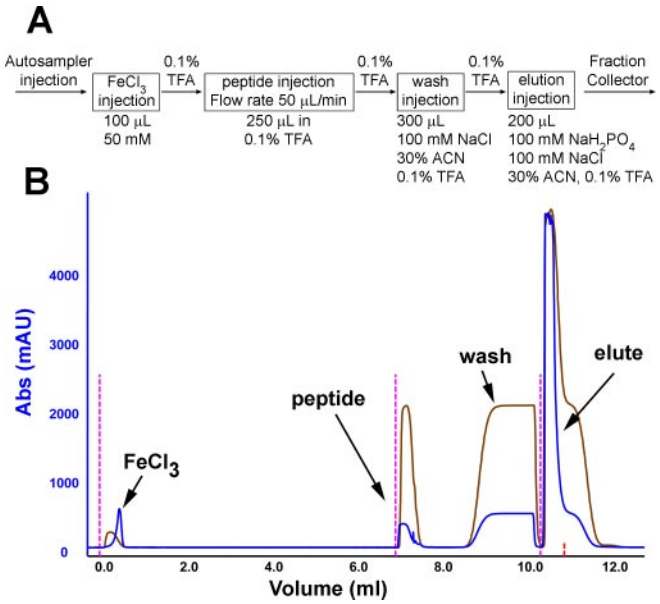

FIG. 5. *A*, schematic representation of the automated IMAC workflow. *B*, representative chromatogram of IMAC purification of phosphopeptides (215 nm absorbance is shown in *blue*; conductance is shown in *brown*). A fraction collector collected the elution fraction between 10 and 10.5 ml. The *red vertical lines* represent injection and collection time points for the various solutions. *mAU*, milliabsorbance units.

corresponding phosphopeptide was observed. The median shift was 11.4 fractions earlier in the SCX gradient or 13% earlier. Our extensive characterization of the protein composition of PSD allowed us to increase the confidence level at which individual phosphopeptides were identified. First all MS/MS spectra were searched against the Uniprot rodent database with high stringency criteria, namely a Mascot expectation value less than 0.001 for a given phosphopeptide, and 258 phosphopeptides were identified. Importantly only seven of these phosphopeptides were from proteins that had not been independently identified in the non-phosphorylated dataset, compelling evidence that the vast majority of phosphopeptides in our mixture (>97%) were derived from this set of proteins. Therefore, we examined the collection of MS/MS spectra to identify phosphopeptides using a dataset restricted to the 1,264 unique proteins determined previously. The 998 phosphopeptides reported here (Supplemental Table S2) were limited to those identified with a Mascot expectation value less than 0.05 when searched against the restricted PSD fraction protein dataset.

The Mascot expectation values used to confirm individual phosphopeptide spectra do not directly address the probability that a given serine, threonine, or tyrosine residue within the peptide was the site of phosphorylation (relative to any other serine, threonine, or tyrosine on the peptide). If a phosphopeptide has more than one potential site of modification and its spectrum is analyzed using Mascot, one can at best compare the Mascot Peptide Scores associated with phosphorylation at each of the potential sites to get a general sense of which amino acid is most likely to be the modified. Furthermore Mascot relies on neutral loss of phosphoric acid (which appears as a loss of 98 Da from the modified amino acid) in MS/MS spectra to designate the exact site of phosphorylation. This strategy is non-optimal for three reasons. First, non-phosphorylated peptide fragments can undergo loss of H_2O, yielding fragment ions with the same mass as a loss of H_3PO_4 from a peptide fragment ion containing a phosphorylated serine or threonine. Second, during CID, the phosphate group on a modified serine or threonine residue does not always undergo complete neutral loss of H_3PO_4. Fragment ions in which the phosphate group is retained can provide valuable information regarding which amino acid is the site of modification (20). Finally, instead of the neutral loss of H_3PO_4, phosphorylated serine and threonine fragment ions can undergo the neutral loss of HPO_3. This can lead to the misidentification of the site of phosphorylation if these −80 Da fragments are misinterpreted as initially non-phosphorylated sequence ions (and hence from a non-phosphorylated region of the peptide).

Using the dataset of PSD phosphorylated and non-phosphorylated peptides, we performed a statistical analysis of how frequently different ion types were observed from these two populations. Results of this analysis and the subsequent automated Perl script we developed to interpret phosphopeptide spectra will be discussed elsewhere.[2]

Analysis of the 998 unique phosphopeptides led to modification site assignment in 621 instances (Supplemental Table S2). For the 306 singly phosphorylated spectra where the site assignment was ambiguous, in 195 of these cases, the site of phosphorylation could be limited to two possible amino acids.

To estimate the percentage of novel phosphorylation sites in our dataset, we compared our results with two public databases of protein post-translational modifications, PhosphoSite (www.phosphosite.org) (21) and the Human Protein Reference Database (www.hprd.org) (22). Although not exhaustive, these databases provided a fairly comprehensive repository of phosphorylation sites reported in the literature. Compared with these databases, ~80% of the phosphorylation sites we identified appear to be novel.

DISCUSSION

We have developed a robust automated workflow to identify sites of phosphorylation from complex protein mixtures. This method was applied to proteins isolated from a PSD preparation from which we identified 1,264 unique proteins. Sites of phosphorylation were mapped on 287 of these proteins or 23% of the total. This is close to the predicted value (33%) for proteins in general (1).

The number of proteins identified in our present analysis substantially exceeds that found in previous mass spectrometric analyses of comparable PSD preparations (3–6). These earlier studies did not detect all proteins known to be present

at the PSD, and importantly, they failed to identify many of the glutamate receptor subunits identified in the present study, suggesting that our identification of PSD sample components is more extensive. Nevertheless it is unlikely that every protein we identified is a true postsynaptic component. For example, there was a minor component of contaminating proteins that are most likely nuclear in origin. This artifact of the preparation method has also been observed in previous studies. In addition, the proteins Bassoon and Piccolo, known presynaptic components, were heavily present at the peptide level. These proteins have been identified in most proteomic studies as major components of PSD preparations (3, 5), possibly a result of interactions of Bassoon and Piccolo with trans-synaptic factors. In contrast, other presynaptic components are strongly reduced in our sample, such as synaptophysin and α-synuclein (Fig. 1), and were not identified by MS analysis (Supplemental Table S1). This indicates that the chosen purification protocol leads to a defined complex containing post-, trans-, and presynaptic components. The traditional term of PSD for this fraction reflects the predominance of many postsynaptic proteins. It is important to note that excitatory synapses in the central nervous system are extremely heterogeneous in nature, and our results represent a composite view of the possible proteins and phosphorylation events present in synapses as a whole rather than the composition of a typical synapse.

We report a direct comparison of identified phosphopeptides by SCX or SCX-IMAC enrichment that demonstrates that the combination of SCX with IMAC provides superior results over SCX fractionation alone (Fig. 4*A*). A larger percentage of each SCX fraction was subjected to IMAC relative to the amount run directly by LC-MS/MS. Nevertheless the amount of each SCX fraction analyzed directly by LC-MS/MS was more than sufficient to saturate the electrospray ion source. Therefore, loading additional material for LC-MS/MS from each SCX run would not have significantly increased the total number of phosphopeptide identifications. A comparison of the results in the present work with our previously obtained data (7) suggested that using IMAC purification alone yields fewer identified phosphopeptides (a total of 83); however, this study was conducted using significantly less PSD sample. In addition, in the previous study we used methyl esterification of tryptic peptides because we had observed a high degree of non-phosphopeptide binding to the resin in the absence of esterification. However, we have found that using lower pH IMAC washing solutions to help reduce carboxylic acid-dependent retention leads to the identification of an equivalent number of phosphopeptides as in the previous study (using IMAC without additional SCX) without using esterification (data not shown). Furthermore there was less of a chance for carboxylic acid-based binding to cause a substantial contribution when SCX chromatography was used prior to IMAC because each SCX fraction was significantly reduced in complexity.

In our previous study, the methyl esterification reaction was only between 90 and 95% efficient despite extensive efforts to drive it to 100%. Such an observation has been reported by others as well (23). If the goal is only to identify phosphopeptides, then incomplete esterification is not necessarily a major issue, and indeed combining methyl esterification with IMAC and SCX may prove better than non-esterification, particularly for samples of even greater complexity, such as whole cell or organ lysates. Our eventual aim is to combine SCX-IMAC enrichment with isotope-based quantitation methods, and thus chemistries that only work at 90% will limit the quantitation dynamic range and decrease the signal to noise ratio.

Assuming 100% efficacy, an *in silico* tryptic digest of the human proteins listed at the National Center for Biotechnology Information predicts that 68% of peptides will have a net charge of +2, not counting the effect of post-translational modifications (17). Only 35% of the peptides identified in our SCX gradient possessed this +2 charge. The remaining higher charge peptides elute on average later in the SCX gradient, and their phosphorylated counterpart (if present) will not likely elute early in the gradient. Our digestion was carried out in 1 M guanidine hydrochloride, and it is possible that the guanidinium ions acted as a competitive inhibitor of the trypsin, reducing its activity and leading to a higher level of missed cleavages than if the digest had been performed using urea as the denaturant. It is therefore difficult to directly compare our method with others who may not have used guanidine during the digestion. To identify a maximal number of phosphorylated peptides from a complex mixture, it is necessary to both enrich for the phosphorylated components as well as separate these phosphopeptides into orthogonal fractions to enable a larger number of MS/MS acquisitions relative to analysis of a single phosphopeptide fraction. In this study, we used SCX primarily to fractionate our peptide mixture. Digestions performed in urea may result in more complete cleavage; however, this will result in less extensive fractionation of the peptide mixture because a greater percentage of the phosphopeptides would elute in the same fraction. Whatever extent of cleavage is achieved for a complex mixture, addition of a phosphate shifts the peptide elution to 10–15% earlier in the gradient relative to the non-phosphorylated counterpart, although the extent of this effect will be gradient-dependent. Therefore, for protein digests separated using a similar SCX gradient, if a non-phosphorylated peptide elutes approximately in the second half of the gradient, its corresponding phosphopeptide is unlikely to elute significantly early to be found in the pool of phosphopeptides with a net +1 charge.

The phosphoproteome of our PSD preparation presented here has revealed predominantly phosphoserine, less phosphothreonine, and rarely phosphotyrosine (Fig. 6*A*). These findings are in line with previous results from a variety of cell types (12, 16) and suggest that our methodological approach does not discriminate against specific amino acid residue modifications. We report sites of phosphorylation across a

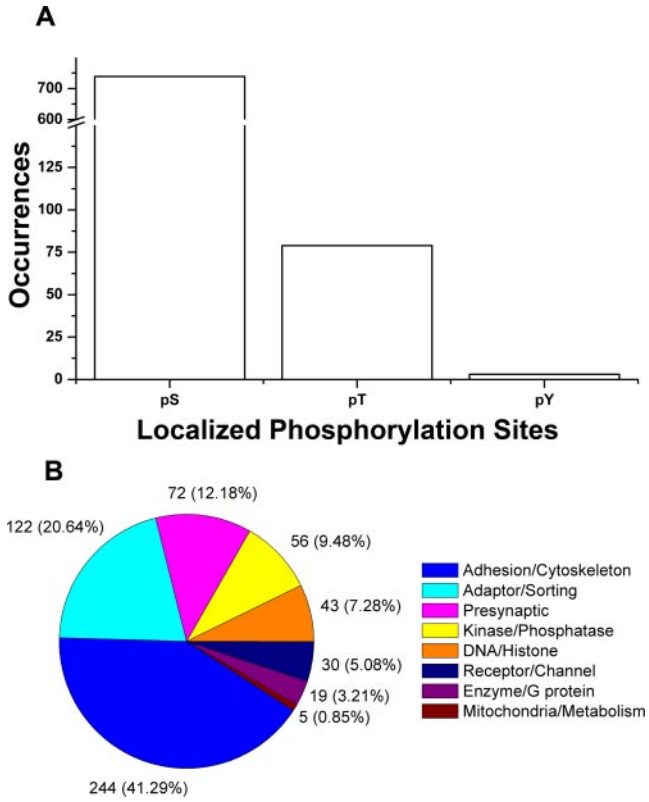

FIG. 6. *A*, the relative distribution of phosphoserines (*pS*), phosphothreonines (*pT*), and phosphotyrosines (*pY*). The majority of identified sites were phosphoserines (739) followed by phosphothreonines (79) and lastly phosphotyrosines (three). *B*, the number of unique phosphorylation sites as a function of the protein functional class. Proteins involved in adhesion/cytoskeleton represent the most prevalent class followed by proteins involved in adaptor/sorting functions.

wide range of protein classes: structural proteins, scaffolding proteins, kinases and phosphatases, membrane neurotransmitter receptors, and voltage-gated ion channels (Fig. 6*B*). From the 621 peptides for which we could identify at least one site of phosphorylation, we were able to determine 723 exact phosphorylation sites (because many of these peptides were multiply phosphorylated). When examined for obvious phosphorylation consensus motifs, ~51% of identified sites fitted the mitogen-activated protein kinase/extracellular signal-regulated kinase recognition motif (S/T)P, whereas ~31% fitted the (R/K)*X*(S/T) motif of cAMP-dependent protein kinase. In addition to the high level of phosphorylation characterized at the synapse, the companion paper to this manuscript (Vosseller *et al.* (24)) found a relatively large number of *O*-GlcNAc modifications at the synapse, underscoring the theory that post-translational modifications play a key role in regulating synaptic function.

Our study does not address a number of issues related to the PSD phosphoproteome or other phosphoproteomes in general. Importantly we cannot at the moment resolve the stoichiometry of the identified phosphorylation sites in the PSD preparation, and the turnover rates of individual phosphorylation sites are unknown. It is conceivable that a number of identified phosphorylation sites, especially sites with low stoichiometry in combination with low turnover, may lack a clear physiological function within the cell. The phosphoproteome of a cell may therefore contain a certain degree of "biochemical noise."

In summary, the combination of SCX with IMAC allowed us to obtain a comprehensive dataset of the postsynaptic density phosphoproteome with 998 unique phosphorylated peptides mapping to 723 unique sites on 287 total phosphoproteins of 1,264 unique proteins. Our results suggest substantial coverage of the PSD phosphoproteome. The obtained dataset will serve as a starting point for characterization of the physiological significance of single phosphorylation sites within the cell that will have to be addressed individually.

Acknowledgments—We thank Aenoch Lynn, June Snedecor, and Peter Baker for assistance with data processing and Peter Giese for providing antibodies against CaMKII and p-CaMKII.

* This work was funded by the Wellcome Trust (to R. S.) and by National Institutes of Health National Center for Research Resources Biomedical Research Technology Program Grants RR01614 and RR14606 (to A. L. B.). Protein Prospector software development was supported by a contribution from the Vincent J. Coates Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

S The on-line version of this article (available at http://www.mcponline.org) contains supplemental material.

¶ Present address: Pritzker Laboratory, Dept. of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA 94304-5485.

‖ To whom correspondence may be addressed. Tel.: 44-20-76797242; Fax: 44-20-76797245; E-mail: r.schoepfer@ucl.ac.uk.

** To whom correspondence may be addressed. E-mail: alb@cgl.ucsf.edu.

REFERENCES

1. Johnson, S. A., and Hunter, T. (2005) Kinomics: methods for deciphering the kinome. *Nat. Methods* **2,** 17–25
2. Yamauchi, T. (2002) Molecular constituents and phosphorylation-dependent regulation of the post-synaptic density. *Mass Spectrom. Rev.* **21,** 266–286
3. Jordan, B. A., Fernholz, B. D., Boussac, M., Xu, C., Grigorean, G., Ziff, E. B., and Neubert, T. A. (2004) Identification and verification of novel rodent postsynaptic density proteins. *Mol. Cell. Proteomics* **3,** 857–871
4. Li, K. W., Hornshaw, M. P., Van Der Schors, R. C., Watson, R., Tate, S., Casetta, B., Jimenez, C. R., Gouwenberg, Y., Gundelfinger, E. D., Smalla, K. H., and Smit, A. B. (2004) Proteomics analysis of rat brain postsynaptic density. Implications of the diverse protein functional groups for the integration of synaptic physiology. *J. Biol. Chem.* **279,** 987–1002
5. Peng, J., Kim, M. J., Cheng, D., Duong, D. M., Gygi, S. P., and Sheng, M. (2004) Semiquantitative proteomic analysis of rat forebrain postsynaptic density fractions by mass spectrometry. *J. Biol. Chem.* **279,** 21003–21011
6. Yoshimura, Y., Yamauchi, Y., Shinkawa, T., Taoka, M., Donai, H., Takahashi, N., Isobe, T., and Yamauchi, T. (2004) Molecular constituents of the postsynaptic density fraction revealed by proteomic analysis using multidimensional liquid chromatography-tandem mass spectrometry. *J. Neurochem.* **88,** 759–768
7. Trinidad, J. C., Thalhammer, A., Specht, C. G., Schoepfer, R., and Burlingame, A. L. (2005) Phosphorylation state of postsynaptic density proteins. *J. Neurochem.* **92,** 1306–1316

8. Rush, J., Moritz, A., Lee, K. A., Guo, A., Goss, V. L., Spek, E. J., Zhang, H., Zha, X. M., Polakiewicz, R. D., and Comb, M. J. (2005) Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nat. Biotechnol.* **23,** 94–101
9. Thelemann, A., Petti, F., Griffin, G., Iwata, K., Hunt, T., Settinari, T., Fenyo, D., Gibson, N., and Haley, J. D. (2005) Phosphotyrosine signaling networks in epidermal growth factor receptor overexpressing squamous carcinoma cells. *Mol. Cell. Proteomics* **4,** 356–376
10. Zhang, Y., Wolf-Yadlin, A., Ross, P. L., Pappin, D. J., Rush, J., Lauffenburger, D. A., and White, F. M. (2005) Time-resolved mass spectrometry of tyrosine phosphorylation sites in the epidermal growth factor receptor signaling network reveals dynamic modules. *Mol. Cell. Proteomics* **4,** 1240–1250
11. Jaffe, H., Vinade, L., and Dosemeci, A. (2004) Identification of novel phosphorylation sites on postsynaptic density proteins. *Biochem. Biophys. Res. Commun.* **321,** 210–218
12. Collins, M. O., Yu, L., Coba, M. P., Husi, H., Campuzano, I., Blackstock, W. P., Choudhary, J. S., and Grant, S. G. (2005) Proteomic analysis of *in vivo* phosphorylated synaptic proteins. *J. Biol. Chem.* **280,** 5972–5982
13. DeGiorgis, J. A., Jaffe, H., Moreira, J. E., Carlotti, C. G., Jr., Leite, J. P., Pant, H. C., and Dosemeci, A. (2005) Phosphoproteomic analysis of synaptosomes from human cerebral cortex. *J. Proteome Res.* **4,** 306–315
14. Ficarro, S., Chertihin, O., Westbrook, V. A., White, F., Jayes, F., Kalab, P., Marto, J. A., Shabanowitz, J., Herr, J. C., Hunt, D. F., and Visconti, P. E. (2003) Phosphoproteome analysis of capacitated human sperm. Evidence of tyrosine phosphorylation of a kinase-anchoring protein 3 and valosin-containing protein/p97 during capacitation. *J. Biol. Chem.* **278,** 11579–11589
15. Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., Burke, D. J., Ross, M. M., Shabanowitz, J., Hunt, D. F., and White, F. M. (2002) Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. *Nat. Biotechnol.* **20,** 301–305
16. Ballif, B. A., Villen, J., Beausoleil, S. A., Schwartz, D., and Gygi, S. P. (2004) Phosphoproteomic analysis of the developing mouse brain. *Mol. Cell. Proteomics* **3,** 1093–1101
17. Beausoleil, S. A., Jedrychowski, M., Schwartz, D., Elias, J. E., Villen, J., Li, J., Cohn, M. A., Cantley, L. C., and Gygi, S. P. (2004) Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 12130–12135
18. Nuhse, T. S., Stensballe, A., Jensen, O. N., and Peck, S. C. (2003) Large-scale analysis of *in vivo* phosphorylated membrane proteins by immobilized metal ion affinity chromatography and mass spectrometry. *Mol. Cell. Proteomics* **2,** 1234–1243
19. Gruhler, A., Olsen, J. V., Mohammed, S., Mortensen, P., Faergeman, N. J., Mann, M., and Jensen, O. N. (2005) Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. *Mol. Cell. Proteomics* **4,** 310–327
20. Chalkley, R. J., and Burlingame, A. L. (2003) Identification of novel sites of *O*-*N*-acetylglucosamine modification of serum response factor using quadrupole time-of-flight mass spectrometry. *Mol. Cell. Proteomics* **2,** 182–190
21. Hornbeck, P. V., Chabra, I., Kornhauser, J. M., Skrzypek, E., and Zhang, B. (2004) PhosphoSite: a bioinformatics resource dedicated to physiological protein phosphorylation. *Proteomics* **4,** 1551–1561
22. Peri, S., Navarro, J. D., Amanchy, R., Kristiansen, T. Z., Jonnalagadda, C. K., Surendranath, V., Niranjan, V., Muthusamy, B., Gandhi, T. K., Gronborg, M., Ibarrola, N., Deshpande, N., Shanker, K., Shivashankar, H. N., Rashmi, B. P., Ramya, M. A., Zhao, Z., Chandrika, K. N., Padma, N., Harsha, H. C., Yatish, A. J., Kavitha, M. P., Menezes, M., Choudhury, D. R., Suresh, S., Ghosh, N., Saravana, R., Chandran, S., Krishna, S., Joy, M., Anand, S. K., Madavan, V., Joseph, A., Wong, G. W., Schiemann, W. P., Constantinescu, S. N., Huang, L., Khosravi-Far, R., Steen, H., Tewari, M., Ghaffari, S., Blobe, G. C., Dang, C. V., Garcia, J. G., Pevsner, J., Jensen, O. N., Roepstorff, P., Deshpande, K. S., Chinnaiyan, A. M., Hamosh, A., Chakravarti, A., and Pandey, A. (2003) Development of human protein reference database as an initial platform for approaching systems biology in humans. *Genome Res.* **13,** 2363–2371
23. Larsen, M. R., Thingholm, T. E., Jensen, O. N., Roepstorff, P., and Jorgensen, T. J. (2005) Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol. Cell. Proteomics* **4,** 873–886
24. Vosseller, K., Trinidad, J. C., Chalkley, R. J., Specht, C. G., Thalhammer, A., Lynn, A. J., Snedecor, J. H., Guan, S., Medzihradszky, K. F., Maltby, D. A., Schoepfer, R., and Burlingame, A. L. (2006) *O*-Linked *N*-acetylglucosamine proteomics of postsynaptic density preparations using lectin weak affinity chromatography and mass spectrometry. *Mol. Cell. Proteomics* **5,** 923–934