Proteomic Analysis of \textit{in Vivo} Phosphorylated Synaptic Proteins*§

Received for publication, September 30, 2004, and in revised form, November 16, 2004
Published, JBC Papers in Press, November 30, 2004, DOI 10.1074/jbc.M411220200

Mark O. Collins‡§, Lu Yu§, Marcelo P. Coba§, Holger Husi‡, Iain Campuzano†, Walter P. Blackstock†, Jyoti S. Choudhary§, and Seth G. N. Grant‡§**

From the ‡Division of Neuroscience, University of Edinburgh, Edinburgh EH9 3JZ, United Kingdom, the §Wellcome Trust Sanger Institute, Hinxton CB10 1SA, United Kingdom, ¶Waters Corp., Manchester M22 5PP, United Kingdom, and the ¶Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield S10 2TN, United Kingdom

The molecular architecture of the synaptic junction has been studied intensely for many years, yielding information on its composition and function based on studies of individual receptors or small groups of proteins (1). In recent years, with the advent of proteomic technologies, a coherent map of the mammalian synapse proteome has been emerging. Mass spectrometry (MS)\textsuperscript{1}-based analysis of the postsynaptic density (PSD) has established for the first time a detailed list of its molecular components (2–4). Systematic analysis of functional multiprotein complexes embedded in the PSD (5) has also added to our knowledge of the overall organization of the postsynaptic proteome.

Central to the functioning of signaling complexes and indeed the most basic signaling pathways is the process of reversible phosphorylation. The propagation of an appropriate synaptic response to receptor stimulation is highly regulated by phosphorylation cascades. This is exemplified by the process of synaptic plasticity, a process whereby glutamate receptor activation results in divergence signaling cascades, which ultimately lead to activation of transcription factors and modulation of gene expression. Phosphorylation is also employed to modulate protein function and stability and to mediate phosphorylation-dependent protein-protein interactions (e.g. Src homology 2 binding of phosphatidylinositol 3-kinase to NR2B) (6), conferring a higher order level of regulation in such protein complexes.

Historically, synapse phosphorylation and its importance in regulating neuronal signal transduction and brain function has been studied at the level of single molecules (7, 8), but new proteomic strategies lend themselves to the global characterization of the signaling properties of the synapse proteome. Phosphopeptides can be purified from complex protein mixtures using immobilized metal affinity chromatography (IMAC) (9) and identified using MS (10). Recent phosphoproteomic studies have utilized various peptide IMAC approaches, sometimes with methyl esterification, to enrich and improve specificity for phosphopeptides prior to MS. This approach has been successfully used to study phosphorylation in yeast (10), \textit{Arabidopsis} (11), and cell lines (12). However, the application of these approaches to complex mammalian subcellular organelles such as the synapse has yet to be established.

MS, although a powerful tool for analysis of protein phosphorylation, has several on-going technical challenges. In particular, these result from heterogeneity arising from dynamic site occupancy, multiple phosphorylation sites in the same low abundance peptide, and inherently poor fragmentation of phosphopeptides. Continued improvements in three critical areas, unbiased sample enrichment methods, high sensitivity MS, and data analysis methods, are required to fully harness the potential of MS. The field is in a phase of rapid development, and various strategies for protein or modification centric analyses are being explored (13, 14).

Here, we describe the use of a combination of cellular fractionated repeat analysis; NMDA, \textit{N}-methyl-d-aspartate; NRC, NMDA receptor complex; IP\(_3\), inositol triphosphate; PSD, postsynaptic density; LC, liquid chromatography; ESI, electrospray ionization; MOPS, 4-morpholinepropanesulfonic acid; PKA, protein kinase A; PKC, protein kinase C; SH3, Src homology 3.
tion procedures with large scale IMAC phosphoprotein and phosphopeptide enrichment protocols and complementary MS analytical strategies to characterize mouse forebrain synaptosomes. This has resulted in the unambiguous identification of 289 sites of phosphorylation in 79 synaptic proteins involved in important pre- and postsynaptic multiprotein complexes and signaling pathways. Large scale in vitro phosphorylation screening on peptide arrays for 95 sites with seven kinases identified 28 phosphorylated sites and a total of 52 phosphorylation events. The simultaneous identification of large numbers of sites of phosphorylation and identification of responsible kinases, as exemplified by this study, is a powerful approach to expand the current knowledge of cell signaling in a particular system.

EXPERIMENTAL PROCEDURES

Isolation of Synaptosomes—Synaptosomes were prepared as described by Carlin et al. (1980) (15) with minor modifications. In brief, mouse forebrains were homogenized in a cold buffer containing 50 mM Tris acetate, pH 7.4, 10% (w/w) sucrose, 5 mM EDTA, 1 mM phenylmethylsulfonyl fluoride, 2 mg/ml aprotinin, and 2 mg/ml leupeptin. The sample was then centrifuged for 20 min at 800 x g, and the resulting supernatant was centrifuged again for 30 min at 16,000 x g. The pellet was then resuspended in 5 ml/g of original weight in a buffer containing 5 mM Tris acetate, pH 8.1, 1 mM phenylmethylsulfonyl fluoride, 2 mg/ml aprotinin, and 2 mg/ml leupeptin, quickly homogenized, and incubated for 45 min on ice. After a further homogenization step and the addition of sucrose to 34% (w/w), the sample was overlaid with solutions containing 28.5% (w/w) sucrose in 50 mM Tris acetate, pH 7.4, and 10% (w/w) sucrose in 50 mM Tris acetate, pH 7.4, and centrifuged for 2 h at 60,000 x g at 4 °C. The protein-containing band, which formed between the 34 and 28.5% sucrose gradients, was collected and diluted with 50 mM Tris acetate, pH 7.4, to 10% sucrose and then centrifuged for 30 min at 48,000 x g. The supernatant was then resuspended in 50 mM Tris acetate, pH 7.4, and homogenized gently to form the synaptosomal preparation.

Protein IMAC of Urea-soluble Synaptosomal Fraction—Fast-flow chelating Sepharose with iminodiacetic acid (IDA) (Amersham Biosciences) or nitrotriacetic acid (Qiagen) chelating groups were charged with GaCl3 or FeCl3, Synaptosomal proteins (12.5 mg) were solubilized in 6 ml urea, and the supernatant was removed and incubated with 2 ml of the metal charged resin with mixing for 1 h at room temperature. The unbound protein was washed with buffer A (6 ml urea, 50 mM Tris acetate) to base line, and the phosphoproteins were specifically eluted with buffer B (6 ml urea, 50 mM Tris acetate, 100 mM EDTA, 100 mM EGTA). The fractions were collected, concentrated, and washed with buffer B in a Vivaspin 6 PES membrane spin column (Vivascience). 13.4 mg of protein was diluted with buffer C (1M urea, 0.125M thiourea, 5% CHAPS) and incubated with 2 ml chelating Sepharose with iminodiacetic acid (IDA) (Amersham Biosciences) for 60 min at 4°C. The chelating resin with metal proteins was separated on a 12% SDS-polyacrylamide gel and sequentially stained with Pro-Q diamond (phosphoprotein) and SYPRO Ruby (total protein) stains (Molecular Probes, Inc., Eugene, OR). Peptidomimetic Stock phosphoprotein molecular weight standards (Molecular Probes) served as a molecular weight marker and internal control. Images were captured with a Typhoon scanner (Amersham Biosciences) and overlaid using TotaLab software (Non-linear Dynamics).

Table I

| Solvent | 0 min | 180 min | 240 min | 242 min | 247 min |
|---------|-------|---------|---------|---------|---------|
| A (%)   | 94    | 62      | 60      | 60      | 60      |
| B (%)   | 6     | 70      | 50      | 50      | 50      |

On-line Nano-LC-MS/MS—A nanoflow high pressure liquid chromatography system, Ultimate™ (LC Packings) or CapLC (Waters), was coupled to a Q-ToF 1, Q-ToF Ultima (Waters/Micromass), or 4000 QTRAP (Applied Biosystems). Tryptic peptides from the phosphoprotein digest were loaded in 0.1% aqueous formic acid and desalted on PepMap C18 trapping cartridge (LC Packings). BetaMax Neutral (Thermo Hypersil-Keystone) was used to trap the IMAC-enriched phosphopeptides in 0.5% aqueous formic acid. Peptides on the trap were back-flushed to and separated on the analytical column (PepMap C18, 75-μm inner diameter × 15 cm; LC Packings). The gradients are shown in Tables I and II.

In the LC-MS/MS analysis of the phosphoprotein digest, the Q-ToF Ultima was operated in automated data-dependent acquisition mode. Each cycle had a 1-s MS survey (m/z 400–1500), and up to three of the highest intensity multiply charged ions (+2 and +3) were selected for MS/MS (m/z 50–2000) each for 5 s. The collision energy in MS/MS was varied according to the m/z and the charge state of the precursor ion. Due to the high complexity of the sample, there were two LC-MS/MS runs for the same LC gradient. After the first standard run (Run 1), survey data were examined, and multiply charged ions above the intensity threshold that had not been subjected to MS/MS in Run 1 were incorporated into the inclusion list for the second run (Run 2). We term this latter LC-MS/MS approach, incorporating a primary analysis followed by a repeat experiment based on an inclusion list, targeted repeat analysis (TRA).

Analysis of IMAC-enriched phosphopeptides was performed on the Q-ToF using similar acquisition parameters to those for the Q-ToF Ultima. Two different precursor-scanning approaches were adopted on the 4000 QTRAP to selectively analyze phosphopeptides. In the first approach, the instrument was used in negative ion mode (+ESI) to scan for the precursors of m/z 79 (PO4) with automatic switching to positive ion mode (+ESI) for MS/MS (m/z 50–2000) for the detected precursors. In the second approach, we applied TRA strategy. First, one-third of the sample was analyzed in negative mode for detecting precursors of m/z 79. Data were examined to generate the inclusion list for the second run in positive ion mode (+ESI) using the remaining two-thirds of the sample in information-dependent acquisition mode. Raw data were processed to give a peak list file and submitted to a local Mascot version 2.0 (Matrix Science) server for iterative searching on a custom, nonidentical, combined human and mouse IPI data base (EBI). Assignment of phosphorylation sites was verified manually with the aid of PEAK Studio (Bioinformatics Solutions) software.

Peptide Array Phosphorylation Assays—Jenri Phosphosite detector™ peptide arrays (Jerini Peptide Technologies, Gmbh) were used to determine which of the MS-identified phosphorylation sites (95 selected sites) could be phosphorylated by seven kinases in an in vitro assay. 15-amino acid-long peptides that encompassed the selected sites were synthesized on cellulose membranes in a parallel manner using SPOT technology (16) deposited to glass slides and were covalently immobi- lized to the glass slide surface. Each peptide was present in triplicate on the chip, and seven full-length proteins that are capable of being phos- phorylated were also included. Negative control peptides for each phospho- pylation site were included, replacing serine or threonine with alanine or valine, respectively, and positive control peptide sequences for each kinase were present.

Peptide arrays were sealed with Gene-Frame™ incubation chambers (Agene house, Surrey, UK), and the chambers were filled with 330 μl of kinase buffer (20 mM MOPS, pH 7.2, 25 mM β-glycerol phosphate, 5 mM, EGTA, 1 mM sodium orthovanadate, 1 mM diithiothreitol, 100 μM

phosphoprotein samples were separated on a 12% SDS-polyacryl- amide gel and sequentially stained with Pro-Q diamond (phosphoprotein) and SYPRO Ruby (total protein) stains (Molecular Probes, Inc., Eugene, OR). Peptide arrays were sealed with Gene-Frame™ incubation chambers (Agene house, Surrey, UK), and the chambers were filled with 330 μl of kinase buffer (20 mM MOPS, pH 7.2, 25 mM β-glycerol phosphate, 5 mM, EGTA, 1 mM sodium orthovanadate, 1 mM diithiothreitol, 100 μM
ATP, 15 mM MgCl2, and 10 μM [γ-32P]ATP. Recombinant active kinases (Upstate Biotechnology, Inc., Lake Placid, NY) (3 μg of PKA catalytic subunit (with 2 μM eAMP), 2 μg of Akt1 (α, PH, S473D), 2 μg of Erk1, 3 μg of p38α, 3 μg of CKII, 3 μg of Cdk5/p35, and 1.5 μg of PKC (α, β, γ)) were included in the appropriate kinase assays. In the case of PKC, a modified kinase buffer was used (10 mM MOPS, pH 7.2, 12.5 mM β-glycerol phosphate, 2.5 mM EGTA, 0.5 mM sodium orthovanadate, 0.5 mM dithiothreitol, 0.1 mM CaCl2, 0.1 mg/ml phosphatidylserine, 25 mM dithiothreitol, 0.5 mM CaCl2, 0.1 mg/ml phosphatidylserine, 25 μM dithiothreitol).

After a 45-min incubation at 32 °C, the peptide microarrays were washed six times, alternating between 0.15 M phosphoric acid and distilled water. γ-32P incorporation in the immobilized peptide spots was detected on a Typhoon 8600 PhosphorImager (350 phosphopeptides containing 331 sites of phosphorylation).

Protein IMAC Protocol—Unlike peptide IMAC that is widely used for phosphoproteomic analysis, to our knowledge protein IMAC approaches have not been reported. In order to develop a protein IMAC protocol, we tested two resins: Sepharose-IDA (a tridentate ligand) and agarose-nitriotriacetic acid (a quadridentate ligand) and two metal ions, FeCl3 and GaCl3 (Fig. 1B), to isolate phosphoproteins from urea-soluble preparations. The IDA resin with both metal ions showed marked enrichment indicated by specific phosphoprotein staining on SDS-polyacrylamide gels. The Ga3+ resin showed more effective depletion of phosphoprotein from the unbound fraction together with stronger phosphostaining of the purified sample (Fig. 1C) and therefore was used in all subsequent protein IMAC experiments.

Analysis of tryptic peptides from the protein IMAC enrichment by LC-MS/MS identified 152 proteins and 19 phosphopeptides. A further 28 different phosphopeptides from these samples were identified using the precursor-scanning approach, applied in the routine as well as the targeted manner as described under “Experimental Procedures.” Clearly, the TRA strategy is effective in both standard LC-MS/MS and precursor scanning modes for extending both protein and phosphorylation mapping from complex peptide mixtures.

186 proteins have been identified in the protein IMAC sample. 105 of these have been confirmed as phosphoproteins in the literature and by phosphorylation sites identified in this study (Fig. 2). This suggests that whereas a small proportion may represent contaminating proteins, the majority are probably phosphoproteins. The identified components represent a diverse range of protein classes (Fig. 2B and Supplemental Table I), and low abundance protein classes, such as kinases, phosphatases, and small G-proteins and modulators, are well represented even in the presence of very abundant cytoskeletal proteins.

The protein IMAC approach presents important benefits for phosphorylation analysis. It offers the opportunity to use basic protein identification to identify many candidate phosphoproteins, which are of low abundance and would not have been in the dynamic range required for direct phosphorylation analysis. The majority of phosphopeptides characterized in MS analyses of the phosphoprotein mixture cover single phosphorylation events (Fig. 3) and thus are complementary to measurements made from the peptide IMAC approach that better represents highly phosphorylated peptides. A further important benefit is that it provides scope for two stages of sample enrichment, at the protein level and subsequently at the peptide level.

**Double IMAC Protocol**—LC-MS/MS analysis of phosphopeptide from double IMAC yielded identification and characterization of 176 phosphopeptides (Fig. 2A and Supplemental Table II) derived from 41 phosphoproteins. The results from the double IMAC protocol significantly extended those from the protein IMAC protocol on several levels. First, of the 25 validated
phosphoproteins found in the protein IMAC experiments, we found an additional 115 phosphopeptides on 16 proteins, indicating greater depth of analysis. We also identified a further 14 phosphopeptides corresponding to seven proteins detected in the protein IMAC sample. Second, 19 phosphoproteins containing 37 phosphopeptides that were not characterized in the protein IMAC were observed in the double IMAC protocol. Third, the increased depth of phosphorylation coverage through overlapping peptides reflects variability in site occupation that can occur on a protein and a level of heterogeneity associated with this type of modification. This is well illustrated in Bassoon, for which we found 15 phosphorylation sites (10 peptides) in the protein IMAC protocol and 16 new sites (16 peptides) in the double IMAC protocol (Supplemental Table III). It is evident that combination of protein and double IMAC strategies enables identification and characterization of phosphoproteins across a wider range of protein abundance and phosphorylation states.

**Synaptic Membrane Phosphoproteomic Analysis**—Integral membrane proteins are usually difficult to analyze, since detergent-based extraction methods are not readily compatible with subsequent purification strategies and LC-MS/MS analysis. However, for a phosphoproteomic analysis, cytoplasmic domains of integral membrane proteins are sufficient, since they contain the phosphorylation sites involved in intracellular signaling. Therefore, we digested 6 M urea-insoluble fractions of synaptosomes; the supernatant was desalted and esterified as described above and was subjected to a single peptide IMAC and LC-MS/MS analysis (Fig. 1A). This approach allowed identification and characterization of 60 phosphopeptides from 31 proteins (Supplemental Table II), of which 12 are predicted to be integral membrane proteins and many others, such as PSD-95 and Adapter-related protein complex-2 α-1, are known to be membrane-associated. 10 of the identified membrane proteins were not previously known to be phosphoproteins. Of these 31 phosphoproteins, 18 were not detected in either protein or double IMAC protocols (Fig. 3A).

**Complementarity of Analytical Approaches**—Inspection of the summarized data (Figs. 2A and 3) shows that the combination of subcellular fractionation and IMAC protocols with MS analyses are complementary, both in terms of the protein identification and phosphorylation site characterization. Overall, we sequenced 350 phosphopeptides covering 653 phosphorylation events. These correspond to a total of 331 phosphorylation sites, of which 289 were localized unambiguously (Fig. 1A). The phosphorylation sites were found in 79 synaptic proteins, and we identified a further 149 candidate phosphoproteins. Interestingly, the large numbers of phosphorylation events characterized by the double IMAC approach mainly correspond to greater coverage of a small set of proteins. The phosphopeptide-based methods, in contrast, cover a proportionately larger number of proteins but with fewer representative sites.

Further complementarity is observed in the frequency of eluted protein from each purification was separated on a 12% SDS-polyacrylamide gel and stained and imaged sequentially. Phosphorylated protein appears pink to black, whereas unphosphorylated protein appears green. Lane 1 shows Peppermint Stick molecular mass markers in which proteins of molecular mass 23 and 45 kDa are phosphorylated, and those of 55 and 116 kDa are not phosphorylated, thus serving as an internal control for selective staining of phosphorylated proteins. A 25-kDa band (see arrow) was shown to consist of ribosomal proteins, which were retained on the IMAC resin due to ion exchange effects and therefore removed from the data set. C, relative signal intensities of the Pro-Q diamond phosphoprotein stained image used in the overlay in B. Phosphoprotein signals for lanes 2–5 in B are shown and plotted versus molecular weight. IDA-Ga has the highest overall signal intensity and thus is the most enriched in phosphoprotein.
phosphorylation of peptides detected using the different protocols (Fig. 3). Peptide IMAC approaches have been reported to enrich for multi-phosphorylated peptides (10, 11), and we have also observed this. In contrast, protein IMAC shows a shift in specificity favoring detection of monophosphorylated peptides. The combination of protein and peptide IMAC results in more

**FIG. 2.** A, Venn diagram of phosphoproteins detected in the protein, double, and peptide IMAC protocols. Data from peptide IMAC experiments on whole synaptosomal digests and urea-insoluble synaptosomal digests were combined to simplify the comparison. Values in brackets indicate the number of phosphoproteins in each set for which evidence of their phosphorylation was found in the literature. *B*, class distribution of phosphoproteins identified. Phosphoproteins were cross-referenced with six MS-based studies of the postsynaptic density and with the NRC. Proteins in which we found phosphorylation sites (experimental) and known phosphoproteins (literature) are indicated.

**FIG. 3.** Distribution of the number of phosphates present per phosphopeptide identified. The values are representative of all phosphopeptides detected by each approach and therefore contain some peptides covering overlapping sequences in proteins but with differing numbers of phosphates attached.
representative coverage, with very similar preferences for mono- and diphasphorylated peptides together with improved detection of higher phosphorylated states. Another approach, using strong cation exchange and exploiting the net charge on phosphopeptides at low pH, has a bias toward singly phosphorylated peptides (17). This approach yielded peptides with an average of 1.15 phosphates/peptide, compared with our 1.85 phosphates/peptide. It is likely that our combination of protein and peptide protocols is more representative of the distribution of phosphorylation sites in vivo.

Interestingly, the peptide-based approach on the membrane fraction shows more multiply phosphorylated peptides (up to six sites) than the other methods. Moreover, the average number of phosphorylation sites per peptide detected in the direct peptide IMAC (2.53) on the membrane fraction is higher than that of whole synaptosomes (2.05) and is greater that that observed for the protein IMAC (1.44) and double IMAC (1.75) analysis of the urea-soluble fraction (Fig. 3). The enrichment of multiply phosphorylated peptides in the urea-insoluble fraction may reflect clusters of phosphorylation sites, often found on the intracellular domains of membrane proteins.

Of the total 331 sites defined, 281 correspond to phosphoserine phosphorylation events, and 43 and 8 can be attributed to phosphothreonine and phosphotyrosine, respectively. Tyrosine phosphorylation is usually associated with a higher gain in signals because it is less abundant and more tightly regulated. We have observed more tyrosine phosphorylation sites than would be expected from previous reports in other large scale studies (11–13). Since most of the phosphotyrosines we characterized are on multiphasphorylated peptides that were identified by the protein or double IMAC approach, this increased coverage can be attributed to increased phosphoprotein enrichment, together with the improvements in analysis of multiphasphorylated peptides.

**Literature Mining and Bioinformatic Assessment**

Literature mining and bioinformatic analysis of the 228 proteins that we have identified in this study clearly indicate an enrichment of phosphoproteins (Fig. 2 and Supplementary Table I). Systematic PubMed searching on the identified components revealed that 110 have been described in the literature as phosphoproteins (Supplemental Table I and Fig. 2). Of the 79 synaptic phosphoproteins characterized here, only 41 were previously reported as phosphorylated, the present work representing a 2-fold increase to the published literature. We also report an additional 114 putative phosphoproteins, which we have characterized based on protein identifications in the protein IMAC samples. 69 (46%) of these components have been reported as phosphorylated in the literature, thereby strongly indicating that further mapping of these proteins will produce confirmation of their phosphorylation (Fig. 2A) and further validation of this approach. Within our data set, we have also identified phosphorylation sites in 14 novel or uncharacterized proteins, adding functional annotation in an unbiased manner not possible in single protein focused studies.

To further explore the putative phosphoproteins as well as phosphopeptides with ambiguous site assignment, we utilized phosphorylation prediction software that provides information on sites and cognate kinases. We found that prediction programs were particularly useful for assigning specific sites to phosphopeptides with multiple potential phosphorylation sites. For example, the peptide 44IG(S)(T)(T)NPFLDIPHDPNAAV-YK64 from mKIAA0942 protein was detected in the protein IMAC analysis and contained one p(S/T) phosphorylation site. Scansite predicted that the most likely site to be phosphorylated in this sequence (by PKCδ) was Ser46, and this site was subsequently confirmed when it was detected again in the double IMAC analysis. We have observed accurate predictions in many similar cases, which support the use of in silico localization of the remaining 43 ambiguous phosphorylation sites (Supplemental Table III).

Based on bioinformatic prediction and, in a few cases, data from the literature, we identified the kinases that were most likely to phosphorylate 289 defined sites (Supplementary Table III). All sites could be accounted for by a minimum set of 23 kinases (Fig. 4), and 14 of these are shown in the literature to phosphorylate proteins in our data set. The predicted kinases emerged as two groups: those that phosphorylate multiple sites on proteins (phosphate/substrate ratio >1 (nine kinases)) (left of the dashed line; Fig. 4) and those that phosphorylate a single site per protein (phosphate/substrate ratio = 1 (14 kinases)). Nine kinases could each phosphorylate 10 or more sites on multiple substrates and together could account for 258 of the observed sites, and eight kinases could phosphorylate 10 or more distinct synaptic substrates. In agreement with another report, we noted that the majority of sites cluster outside characterized protein domains (18) (Supplemental Table III). 25 phosphorylation sites corresponded to residues predicted to mediate phosphorylation-dependent protein-protein interactions such as 14-3-3 and Src homology 2 domains.

**In Vitro Phosphorylation Assays**

Peptide array technology was employed to evaluate kinase specificity and multiplicity inferred from bioinformatic analysis of the synaptic phosphorylation data set. 190 peptides were synthesized that encompassed 95 phosphorylation sites from the described MS analyses. Control peptides were used to confirm enzymatic activity for each kinase, and analytical spot signals were only considered positive when they were positive after background subtraction and when serine/threonine substitutions on control peptides eliminated or significantly reduced the resultant signal. Seven kinases (PKA, Akt, PKC (α, β, and γ), CK2, p38 mitogen-activated protein kinase, extracellular signal-regulated kinase 1, and Cdk5) were used to screen for positive signals across 190 peptides, arrayed in triplicate. Robust positive signals for 28 unique phosphorylation sites were detected, resulting in 52 kinase substrate site identifications (Supplementary Fig. 1A) with an observed kinase/substrate site ratio of 1.9. To evaluate the agreement of this in vitro data with the in silico kinase predictions by Scansite and NetPhosK, predictions were made with both programs for the seven kinases used on the 28 positive sites in the peptide array experiments. Scansite and NetPhosK predicted 58 and 63% of sites, respectively (Supplemental Fig. 1B). The overlap of positive predictions by Scansite and NetPhosK for this set of phosphorylation sites was 26%, strongly indicating that neither algorithm is sufficient and that both must be used to obtain good coverage of a data set.

The number of kinases and binding sites with known consensus sequences constrains global predictions; however, in combination with data sets of defined in vivo phosphorylation sites and in vitro phosphorylation data, they point to the fact that multiple synaptic substrates may exist for each kinase, and a complex regulatory network may arise.

**DISCUSSION**

The Synapse Phosphoproteome—The aim of this study was to establish an in vivo map of the synaptic phosphoproteome. We therefore examined the representation of phosphorylation data across synaptic proteins and important synaptic subcomponents and organelles. Our data set contained postsynaptic components, including the postsynaptic density (134 proteins) and NMDA receptor-associated proteins (44 proteins) (Fig. 2) and...
presynaptic components, of which the most striking examples were Piccolo-Bassoon transport vesicles.

Presynaptic Multiprotein Complexes—A number of protein complexes have been identified at the presynaptic active zone, a region where synaptic vesicles dock, fuse, and release their neurotransmitters into the synaptic cleft (19). It is well known that phosphorylation is critical to the regulation of calcium-regulated synaptic vesicle exocytosis (20), and we have identified a number of phosphoproteins known to play a role in this process (Fig. 5). Piccolo-Bassoon transport vesicles have been shown to be putative precursor vesicles in the active zone (21) and contain a number of proteins functionally coupled to synaptic vesicle exocytosis. Bassoon, a presynaptic protein with homology to Piccolo, is intimately involved in vesicle release (21) and, together with Piccolo and Rim-1/2, constitutes half of the protein content of Piccolo-Bassoon transport vesicles. We have detected 30 phosphorylation sites on Bassoon and a further 16 sites on Piccolo.

We have identified a number of phosphorylation sites on Rim-1, a Rab3 effector involved in the regulation of synaptic vesicle fusion that is associated with Piccolo-Bassoon transport vesicles (21). PKA phosphorylation of an isoform of Rim1 regulates presynaptic long term potentiation (22), highlighting the regulatory importance of such presynaptic molecules. In addition to the main constituents of Piccolo-Bassoon transport vesicles, we mapped phosphorylation sites on Munc-18. Phosphorylation of Munc-18 by PKC has been shown to disassemble Munc-18-syntaxin 1A complexes and may have a regulatory role in presynaptic plasticity (23). We have identified a number of phosphorylation sites in components of other presynaptic multiprotein complexes such as cysteine string protein, SNAP25-interacting protein, and the Rab11-interacting protein, Rip11 (Fig. 5).

Postsynaptic Multiprotein Complexes and Pathways—The neurotransmitter glutamate activates synaptic plasticity primarily via the ionotropic NMDA receptor and metabotropic (mGluR) receptors. This leads to Ca2+/H11001 elevation in the dendritic spine and signal transduction to α-amino-3-hydroxy-5-methylisoxazole-4-propionate receptors and other effector mechanisms. Proteomic analysis of proteins associated with the NMDA receptor revealed that a multiprotein complex of over 180 proteins is embedded in the postsynaptic density (5, 24). We have identified 115 phosphopeptides derived from 17 NMDA receptor complex (NRC) proteins, most of which have been also identified as PSD proteins and an additional 142 phosphopeptides from another 29 PSD proteins (Figs. 2 and 4). Phosphopeptides derived from three PDZ domain-containing proteins (Shank1, PSD-93/Chapsyn110, and PSD-95) were detected in this study. PDZ domain-containing proteins are important components of the NRC and the postsynaptic density and form protein scaffolds that support and regulate many essential synaptic signaling processes (25). We found a phosphorylation site in PSD-93 (Ser414), which was also found in MS analyses of purified NRC complexes. Since this site was detected in the NRC with no specific enrichment for phosphoproteins, it must be highly phosphorylated and enriched in this complex, in agreement with another study of PSD-93 phosphorylation in the NRC (26).

Three phosphorylation sites were mapped in PSD-95 within a 20-amino acid stretch, two Ser(P) just before its SH3 domain and Tyr(P)432 located at the start of its SH3 domain in the β1 sheet. Structural analysis of the SH3 and guanylate kinase domains of PSD-95 (27, 28) showed that there is an intermo-

\[ \text{H. Husi, J. S. Choudhary, W. P. Blackstock, and S. G. N. Grant, unpublished results.} \]
molecular interaction between the SH3 and guanylate kinase domains involving the formation of a β-sheet, including residues N- and C-terminal to the SH3 domain. Tyr432 is located in this hydrophobic core mediating this interaction and is itself interacting with His461. Introduction of a negative charge by phosphorylation of Tyr432 could disturb this interaction and may have a regulatory role in the formation of an SH3 and guanylate kinase intermolecular interaction and also affect the interaction of other proteins, such as AKAP79, that bind to this region (29). It has been proposed that this intermolecular interaction is a conserved feature of the membrane-associated guanylate kinase superfamily based on mutational data available for different membrane-associated guanylate kinase proteins (28). These phosphorylated sequences in PSD-95 are also present in two paralogous proteins, SAP97 and SAP102, representing a possible conservation of regulatory functions.

There is a growing body of evidence that scaffold proteins are regulated by phosphorylation (30), which is consistent with the number of identified scaffold phosphoproteins identified in this study. In fact, we have identified two phosphorylation sites on PKA-RIIα, which map to a defined 15-amino acid sequence, which has been shown to interact with MAP2 (31). This interaction is believed to influence cytoskeletal localization of the PKA holoenzyme (31), and phosphorylation of the binding site could modulate this interaction and ultimately the localization of PKA.

The usefulness of a differential urea extraction strategy is shown by the identification of phosphorylation sites in a number of integral membrane or membrane-associated proteins. We have found sites in cell adhesion molecules such as neural cell adhesion molecule and neurofascin, potassium channels, and a number of other known membrane proteins. However, the value of such a targeted extraction procedure is exemplified by the identification of components of a plasma membrane to intracellular membrane signaling complex associated with mGluR5 (32) (Fig. 5). We have identified phosphorylation sites in mGluR5 (group I metabotropic glutamate receptor) and IRBIT (membrane-associated inositol trisphosphate (IP3) receptor-binding protein) in the urea-insoluble synaptosomal fraction and phospholipase C-β in the urea-soluble fraction. Activation of mGluR5 results in the hydrolysis of membrane phosphatidylinositol bisphosphate to diacylglycerol, which activates PKC, and IP3, which in turn activates the IP3 receptor to release intracellular calcium (33). IRBIT binds to the IP3 receptor via its N terminus and is released from the IP3 receptor upon IP3 binding. The N-terminal region of IRBIT responsible for this interaction contains three novel phosphorylation sites, and phosphorylation in this region has been postulated to regulate this interaction (34). mGluR5, G-proteins, and phospholipase C-β form a multiprotein complex with the IP3 receptor and IRBIT via the scaffolding properties of Homer (35). This complex seems to facilitate a signaling pathway from mGluR5 to modulation of intracellular calcium, which is known to regulate a myriad of intercellular signaling activities.

mGluR5 has a critical role in NMDA receptor-dependent forms of synaptic plasticity and excitotoxicity. Co-activation of mGluR5 and NMDA receptors is required for potentiation of excitatory synaptic transmission in hippocampal neurons (36), and this process requires IP3 receptor-mediated mobilization of intracellular calcium and activation of PKC for maintenance of potentiation. As well as being functionally coupled, there is evidence to suggest that these receptors are physically associated; mGluR5 is a component of the NMDA receptor complex (5), NR2A (an NMDA receptor subunit) is present in the mGluR5 receptor complex (32), and 20 other proteins are common to these complexes. Functional analysis of novel phosphorylation sites on mGluR5 and other components of the associated pathway to calcium release and PKC activation may provide novel insights into the functional interaction of these two important glutamate receptors.
Peptide Array-based Kinase Screen—Oriented peptide libraries were first used in the study of protein phosphorylation to map target specificity of kinases. This approach has contributed much of the kinase consensus information currently known and constitutes the basis of a widely used phosphorylation prediction algorithm (37). It is thought that many phosphorylation sites tend to occur in accessible and flexible regions in three-dimensional protein structures, and in agreement with another study (18), the majority of phosphorylation sites (65%) we identified are predicted to be outside structural domains. This would indicate that phosphorylation of linear peptide sequences in vitro should be similar to phosphorylation of the intact protein for the majority of sites. Data derived from immobilized peptide array experiments are consistent with known kinase consensus sequences (38, 39) and therefore represent a useful tool for studying phosphorylation.

Peptide array screening of a limited set of phosphorylation sites with seven kinases resulted in the identification of 52 kinase phosphorylation site sets. Kinase data for two phosphorylation sites screened on peptide arrays were available in the literature. Microtubule-associated protein Tau was shown to be phosphorylated at Ser493 by Cdk5 (40). This site was identified and was hyperphosphorylated only in p25-overexpressing mice (increased activation of Cdk5) and along with other hyperphosphorylated residues contributed to neurodegeneration and formation of neurofibrillary tangles (40). We identified this site in vivo in wild type mice and have shown that it can be phosphorylated by Cdk5 on a peptide array. Also, Myelin basic protein is phosphorylated at Ser144 by PKA and PKC (41), in agreement with our peptide array observations for this site. The reported validation and use of peptide arrays in the literature and confirmation of cognate kinases for these two sites supports the value of using peptide array technology as a screening tool for in vivo identified phosphorylation sites. We recognize that these in vitro data are not sufficient on their own to definitively prove that a kinase may phosphorylate a given site in vivo; however, because these phosphorylation sites were identified from in vivo preparations, it is reasonable to use this peptide array technology as a first approach to screen for possible substrates.

Large scale screening of kinase substrates directly in a protein complex is currently not feasible due to technical limitations. However, an alternative strategy would be to first assign kinases to phosphorylation sites using peptide array screening,
which is very scalable, and then for example, phospho-specific antibodies or specific MS-based approaches could be used to assess the relevance of a substrate phosphorylation in a particular complex or organelle. Since phosphoproteomic approaches are beginning to yield unprecedented numbers of phosphorylation site identifications, peptide array-based screening of cognate kinases may prove to be very useful for selecting and prioritizing phosphorylation sites or phosphoproteins for further functional investigation.

**Kinase Association—**An important issue in assigning kinases to specific phosphorylation sites on substrates is one of kinase proximity. The potential of a substrate to be phosphorylated by a kinase depends on the presence of the appropriate consensus sequence as well as the presence of the kinase in the microenvironment of the substrate. Kinases can be targeted to their substrates via adapter proteins (e.g., a membrane-associated guanylate kinase-AKAP79 complex recruits PKA to the postsynaptic membrane to phosphorylate the α-amino-3-hydroxy-5-methylisoxazole-4-propionate receptor subunit GluR1 (29)). Furthermore, it is probable that a given phosphorylation site can be phosphorylated by multiple kinases in an *in vitro* assay, but either the site is specific for a given kinase in a given complex or, in some cases, the substrate would not usually be in proximity to that kinase in physiological circumstances.

We addressed this issue of proximity by focusing on protein-protein interactions in the NRC. A network of interacting NRC proteins was constructed, and kinase-substrate data from MS and peptide array experiments were superimposed onto this network (Fig. 6A). Six kinases were shown to phosphorylate nine NRC components on the peptide array. It should be noted that three of these kinases (AKT1, p38 mitogen-activated protein kinase, and Cdk5) were not found in the NRC; perhaps these kinases are transiently associated with the NRC (thus beyond detection limits), or they phosphorylate NRC components in a different physiological context. Certain substrates such as SAPAP1/2 act as phosphorylation hubs, being phosphorylated on various sites by all six kinases, whereas others are phosphorylated by a single kinase. It is apparent that some kinases such as PKA act on sets of interacting proteins (Fig. 6D). PKA is linked to four interacting substrates (PSD-95, PSD-93, and SAPAP1 and -2) via AKAP79, a kinase anchor protein that recruits PKA to the postsynaptic membrane in proximity to its physiological substrates. The layering of kinase-substrate data onto protein-protein interaction networks such as this can result in the identification of interacting substrates, which are more likely to be physiologically important and therefore should be prioritized for further functional annotation. A prerequisite for this kind of approach is a well defined and functionally relevant interactome such as the NRC, and layering dynamic aspects of phosphorylation such as phosphorylation time courses, in response to stimuli, will certainly advance the dissection of complex signaling pathways.

**Synaptic Plasticity and Disease—**In addition to identifying and characterizing phosphoprotein components in important synaptic multiprotein complexes, we have found phosphorylation sites in a number of proteins, which are directly implicated in synaptic plasticity and disease (Supplemental Table IV). It is generally accepted that phosphorylation is important in the regulation of synaptic plasticity, a process that is believed to be involved in learning and memory (42), drug addiction (43), and pain (44). Functional studies have shown that seven phosphoproteins we identified are involved in synaptic plasticity, and perturbations of five in rodents show impairments in learning and memory. Also, seven of the top 11 kinases predicted to phosphorylate our total phosphoprotein data set are known to be intimately involved in synaptic plasticity. We have identified eight phosphoproteins that have been linked to schizophreria and other mental disorders. Finally, we have detected at least seven known phosphorylation sites on microtubule-associated protein tau (and one novel site), many of which have been shown to be involved in neurodegeneration and the development of neurofibrillary pathology (40).

We have described large scale analysis of phosphorylation of synaptic proteins using multiple complementary approaches at the levels of protein extraction, phosphoprotein and phosphopeptide enrichment, analysis by MS, *in vitro* phosphorylation assays, and network analysis. This integrated large scale approach has several advantages over more traditional methods, where a single protein or site is characterized, and in particular it can be used to derive a picture of the global organization of the synapse phosphoproteome. The establishment of phosphorylation maps will provide the basis for numerous functional studies, which will aid the understanding of complex signaling pathways at the synapse. In addition to enhancing this map of the synapse phosphoproteome with further studies, there is an important need to develop novel functional assays to monitor the regulation of multiple phosphorylation events in a physiological context.

**Acknowledgment—**Animals were treated in accordance with the UK Animal Scientific Procedures Act (1986) and National Institutes of Health guidelines.

**REFERENCES**

1. Sheng, M. (2001) *Proc. Natl. Acad. Sci. U. S. A.* 98, 7058–7061
2. Yoshimura, Y., Yamauchi, Y., Shinkawa, T., Taoka, M., Donai, H., Takahashi, N., Isobe, T., and Yamauchi, T. (2004) *J. Neurochem.* 88, 759–768
3. Feng, J., Ji, M. J., Cheng, D., Duong, D. M., Gygi, S. P., and Sheng, M. (2004) *J. Biol. Chem.* 279, 21003–21011
4. Li, K. W., Hornshaw, M. P., Van Der Schors, R. C., Watson, R., Tate, S., Castetta, B., Jimenez, C. R., Gouwenberg, Y., Gundelfinger, E. D., Smalla, K. H., and Smith, A. B. (2004) *J. Biol. Chem.* 279, 987–1002
5. Husi, H., Ward, M. A., Choudhary, J. S., Blackstock, W. P., and Grant, S. G. (2000) *Nat. Neurosci.* 3, 661–669
6. Hitakume, C., Umemori, H., Mishima, M., and Yamamoto, T. (1999) *Genes Cells* 4, 657–666
7. Greengard, P., Valtorta, F., Czernecki, A. J., and Benfenati, F. (1993) *Science* 259, 780–785
8. Sweatt, J. D., and Kandel, E. R. (1989) *Nature* 339, 51–54
9. Andersson, L., and Porath, J. (1986) *Anal. Biochem.* 154, 250–254
10. Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., Burke, D. J., Ross, M. M., Shahbazian, J., Hunt, D. F., and White, F. M. (2002) *Nat. Biotechnol.* 20, 301–305
11. Nuhse, T. S., Stensballe, A., Jensen, O. N., and Peck, S. (2003) *Cell Proteomics* 11, 1234–1243
12. Shu, H., Chen, S., Bi, Q., Mumb, Y., and Brecken, D. L. (2004) *Cell Proteomics* 3, 279–286
13. Beausoleil, S. A., Jedrychowski, M., Schwartz, D., Elias, J. E., Villen, J., Li, J., Cohn, M. A., Csanley, L. C., and Gygi, S. P. (2004) *Proc. Natl. Acad. Sci. U. S. A.* 101, 12130–12135
14. Chang, E. J., Archambault, V., Mclellan, D. T., Krutchinsky, A. N., and Chant, B. T. (2004) *Anat. Cell Biol.* 78, 4472–4483
15. Carlin, R. K., Grab, D. J., Cohen, R. S., and Siekerivitz, P. (1980) *J. Cell Biol.* 86, 831–845
16. Wenslau, H., Volkmer-Engert, R., Schmidt, M., Schulz, M., Schneider-Mergener, J., and Reineke, U. (2000) *Biopolymers* 58, 189–206
17. Ballif, B. A., Villen, J., Beausoleil, S. A., Schwartz, D., and Gygi, S. P. (2004) *Mol. Cell Proteomics* 11, 1093–1101
18. Nuhse, T. S., Stensballe, A., Jensen, O. N., and Peck, S. C. (2004) *Plant Cell 4*, 2294–2405
19. Sudhof, T. C. (2004) *Annu. Rev. Neurosci.* 27, 509–547
20. Turner, K. M., Burgoyne, R. D., and Morgan, A. (1999) *Trends Neurosci.* 22, 459–464
21. Shapira, M., Zhu, R. G., Dresher, T., Bresler, T., Torres, V. I., Gundelfinger, E. D., Ziv, N. E., and Garner, C. C. (2003) *Neuron* 38, 237–252
22. Leonard, G., Schoch, S., Karrick, P. S., Larkin, C. J., Sudhoff, T. C., and Linden, D. J. (2003) *Cell 115*, 45–90
23. Craig, T. J., Evans, G. J., and Morgan, A. (2003) *J. Neurochem.* 86, 1450–1457
24. Grant, S. G. (2003) *BioEssays* 25, 1229–1235
25. Nourry, C., Grant, S. G., and Berg, J. P. (2003) *Science*’s STKE http://stke.sciencemag.org/cgi/content/full/OC_sigtrans/2003/798/7
26. Nada, S., Shimna, T., Yanai, H., Husi, H., Grant, S. G., Okada, M., and Akymama, T. (2003) *J. Biol. Chem.* 278, 47610–47621
27. McGee, A. W., Dakoji, S. R., Olsen, O., Breitt, D. S., Lim, W. A., and Preboda, K. E. (2001) *Mol. Cell 8*, 1291–1301
28. Tavares, G. A., Panepucci, E. H., and Brunger, A. T. (2001) *Mol. Cell* 8, 1315–1325
29. Colldege, M., Dean, R. A., Scott, G. K., Langebein, L. K., Huganir, R. L., and Scott, J. D. (2000) *Neuron* 27, 107–119
30. Giordani, F., Mauzer, D., Fiorentini, C., Bellone, C., Misella, C., Chatabene, F.,
and Luca, M. D. (2003) J. Biol. Chem. 278, 44745–44752

31. Scott, J. D., Stofko, R. E., McDonald, J. R., Comer, J. D., Vitalis, E. A., and Mangili, J. A. (1999) J. Biol. Chem. 274, 21561–21566

32. Farr, C. D., Gafken, P. R., Norbeck, A. D., Doneau, C. E., Stapels, M. D., Barofsky, D. F., Minami, M., and Saugstad, J. A. (2004) J. Neurochem. 91, 438–450

33. Kawabata, S., Kohara, A., Tsutsui, R., Itahana, H., Hayashihara, S., Yamaguchi, T., and Okada, M. (1998) J. Biol. Chem. 273, 17381–17385

34. Ando, H., Mizutani, A., Matsura, T., and Mikoshiba, K. (2003) J. Biol. Chem. 278, 10602–10612

35. Tu, J. C., Xian, B., Yuan, J. P., Lanahan, A. A., Leffert, K., Li, M., Linden, D. J., and Worley, P. F. (1998) Neuron 21, 717–726

36. Kotecha, S. A., Jackson, M. F., Al-Mahrouki, A., Roder, J. C., Orser, B. A., and MacDonald, J. F. (2003) J. Biol. Chem. 278, 27742–27749

37. Obenauer, J. C., Cantley, L. C., and Yaffe, M. B. (2003) Nucleic Acids Res. 31, 3635–3641

38. Lizcano, J. M., Deak, M., Morrice, N., Kieloch, A., Hastie, C. J., Dong, L., Schutkowski, M., Reimer, U., and Alessi, D. R. (2002) J. Biol. Chem. 277, 27839–27849

39. Rychlewski, L., Kechis, M., Deng, L., Schutkowski, M., and Reimer, U. (2004) J. Mol. Biol. 336, 307–311

40. Cruz, J. C., Tseng, H. C., Goldman, J. A., Shih, H., and Teai, L. H. (2003) Neuron 40, 471–483

41. Kishimoto, A., Nishiyama, K., Nakashibi, H., Uratsuji, Y., Nomura, H., Takeyama, Y., and Nishizuka, Y. (1985) J. Biol. Chem. 260, 12492–12499

42. Tokuda, M., and Hatase, O. (1998) Mol. Neurobiol. 17, 137–156

43. Yao, W. D., Gainetdinov, R. R., Arbuckle, M. I., Sotnikova, T. D., Cyr, M., Beaulieu, J. M., Torres, G. E., Grant, S. G., and Caron, M. G. (2004) Neuron 41, 625–638

44. Garry, E. M., Moss, A., Delaney, A., O’Neill, F., Blakemore, J., Bowen, J., Husi, H., Mitchell, R., Grant, S. G., and Fleetwood-Walker, S. M. (2003) Curr. Biol. 13, 321–328
Proteomic Analysis of in Vivo Phosphorylated Synaptic Proteins
Mark O. Collins, Lu Yu, Marcelo P. Coba, Holger Husi, Iain Campuzano, Walter P. Blackstock, Jyoti S. Choudhary and Seth G. N. Grant

J. Biol. Chem. 2005, 280:5972-5982.
doi: 10.1074/jbc.M411220200 originally published online November 30, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M411220200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2005/01/14/M411220200.DC1

This article cites 43 references, 15 of which can be accessed free at
http://www.jbc.org/content/280/7/5972.full.html#ref-list-1