ANALYSIS OF THE MOLECULAR COMPONENTS AND PHOSPHORYLATION OF MOUSE BRAIN PROTEOMES

Mark Oliver Collins

A Thesis Submitted For The Degree Of Doctor Of Philosophy

The College of Medicine & Veterinary Medicine
The University of Edinburgh

2006

DECLARATION

I declare that the work presented in this thesis is my own, except where noted in the text and as detailed below.

Lu Yu and Jyoti Choudhary (The Wellcome Trust Sanger Institute) performed all LC-MS/MS analyses described in this thesis

Marcelo Coba (The Wellcome Trust Sanger Institute) performed the in vitro kinase assays described in Chapter 3.

Holger Husi (University of Edinburgh) performed MASC and PSD purifications described in Chapter 5

Luis Valour (The Wellcome Trust Sanger Institute) generated the gene expression data for developing primary neurons described in Chapter 5.

Lawrence Humphreys and Paul Charlesworth (The Wellcome Trust Sanger Institute) performed electrophysiological recordings of developing primary neurons described in Chapter 5.

Mark O. Collins

ACKNOWLEDGEMENTS

Firstly, I would like to thank Seth Grant for his guidance, support and enthusiasm for this project. In addition, I would like to thank Holger Husi who educated me in the field of proteomics; he imparted his boundless knowledge of protein biochemistry and bioinformatics, even when it was not always requested! I am indebted to Lu Yu and Jyoti Choudhary, my collaborators, for their interest, commitment and hard work on this project and for their patience in explaining the finer points of the world of mass spectrometry to me.

Thank you to Luis Valor, Marcelo Coba and Julia Brandon, with whom I had the pleasure to work with in the latter part of my PhD. A big thanks of course, to those of you that have helped keep me sane over the last few years, all on the fourth floor in 1 George Square, especially Lynsey Forsyth and Cathy Vickers. Also, thanks to my non-scientist friends in Edinburgh that had to put up with listening to science chat in BM, you loved it really! In addition, I would like to thank Lianne, Espe, Fred & Co for looking after me during my write-up and for keeping me entertained.

A special thank you to Therese, who has been there for me in so many ways over the last three years, it would have been much more difficult and a lot more boring without you.

Finally, I would like to thank Mum, Dad and family for being great and for supporting me throughout my many years as a student.

ABBREVIATIONS

AD	Alzheimer's disease
AKAP	A Kinase Anchor Protein
AMPA	α-amino-3-hydroxy-5-methylisoxazole-4-propionic acid
APP	Amyloid precursor protein
ATP	Adenosine Triphosphate
BM	BetaMax
CaM	Calmodulin
CaMKII	Ca^{2+} /Calmodulin Protein Kinase II
CID	Collision-induced dissociation
CK2	Casein Kinase 2
cPSD	Consensus PSD
DAG	Diacylglycerol
DDA	Data-dependent acquisition
EJC	Exon junction complex
FA	Formic acid
FT-ICR	Fourier transform-ion cyclotron resonance
GK	Guanylate kinase
GO	Gene ontology
GRAVY	Grand Average of Hydropathy
HD	Huntington's disease
IDA	Iminodiacetic acid
IDA	Information Dependent Acquisition
IMAC	Immobilised metal-affinity chromatography
IP3	Inositol 1,4,5-trisphosphate
LC/MS-MS	Liquid chromatography-tandem mass sepctrometry
LTD	Long-term depression
LTP	Long-term potentiation
MAGUK	Membrane-associated guanylate kinase
MALDI-MS	Matrix-assisted laser desorption/ionisation mass spectrometry
MASC	MAGUK-associated signalling complexes
MEA	Multi-electrode array
MS	Mass spectrometry
MudPIT	Multidimensional protein identification technology
NMDA	N-methyl-D aspartate
NRC	NMDA receptor complex
NTA	Nitrilotriacetic acid
PHF's	Paired helical filaments
PIP2	phosphatidyl-inositol bisphosphate
PKA	cAMP-dependent protein kinase A
PLC	Thin layer chromatography
PP2B	Protein phosphatase 2B
PPID	Protein-protein interaction database
PSD	Postsynaptic density
PSP	Postsynaptic proteome
PVDF	Polyvinylidene fluoride
PVTs	Piccolo-Bassoon transport vesicles
Q-Tof	Quadrupole-time-of-flight
RP	Reversed-phase
RT	Room temperature

SILAC	Stable Isotope Labelling with Amino acids in Cell culture
TRA	Targeted-repeated analysis

ABSTRACT

The mammalian synapse is essential for transmission of electrical impulses from presynaptic terminals to postsynaptic sites. The translation of this information into a long-term physical modification of the synapse is believed to be central to the process of learning and memory. This physical modification may be manifested by changes in the composition and post-translational modification of synaptic proteins, which together alter the signalling properties of a synapse. Reversible protein phosphorylation mediated by kinases, phosphatases and regulatory molecules, is an essential mechanism of signal transduction in living cells. Although phosphorylation is the most intensively studied of the several hundred known post-translational modifications on proteins, until recently the rate of identification of phosphorylation sites has remained low. Historically, synapse phosphorylation and its importance in regulating neuronal signal transduction and brain function has been studied at the level of single molecules, but new proteomic strategies lend themselves to the global characterisation of the signalling properties of the synapse proteome.

We have developed upon existing immobilised metal-affinity chromatography (IMAC) techniques for capturing phosphopeptides, to selectively purify phosphoproteins from complex mixtures. Using this novel approach, combining both protein and peptide IMAC and mass spectrometry (MS) data acquisition strategies, comprehensive mapping of both the mouse forebrain cytosolic and synaptic phosphoproteomes was achieved. In total, 1138 phosphorylation events were detected, corresponding to 652 phosphorylation sites, of which 618 are unique from the two phosphoproteomes. 331 unique phosphorylation sites were mapped on synaptic phosphoproteins and 92% of these phosphorylation sites are novel, indicating a previously underestimated complexity in synaptic signalling. Bioinformatic and in vitro phosphorylation assays of peptide arrays suggest a small number of kinases phosphorylate many proteins and each substrate is phosphorylated by many kinases. 321 unique phosphorylation sites were mapped to cytosolic phosphoproteins and bioinformatic analysis revealed that this dataset was rich in intrinsic protein disorder, providing insight into the possible functions of phosphorylation in the regulation of these proteins.

In recent years, MS-based analyses of the postsynaptic density (PSD) and receptor complexes, have established for the first time a detailed list of its molecular components. In order to provide a coherent map of the molecular components of the synapse proteome, a bioinformatic approach was used to combine six PSD and three postsynaptic receptor complex datasets into a single database of synaptic proteins. This process of data integration allowed comparisons of analytical approaches used and revealed the most effective biochemical and MS-based methodologies. This data was used as a framework, on which multiple data sources were integrated and allowed the derivation of proteome-wide molecular network maps at the level of gene regulation, protein interaction and protein phosphorylation. These maps were merged with functional or phenotype data from individual molecule studies to derive new models of synapse function. This work significantly increases our existing knowledge of synapse protein phosphorylation and together with our analysis of the postsynaptic proteome; we observe evidence of multiple levels of organisation These data indicate the existence of a tightly regulated and highly complex signalling network for translating electrical information into biochemical changes.

PUBLICATIONS FROM THIS THESIS

This thesis contains the following publications that are included in the appendix.

CHAPTER 1 (In part)
Appendix 9 page 340

Collins, MO, Husi, H, and Grant, SGN. (2004) *Proteomics in the Neurosciences*. In Proteomics: Biomedical and Pharmaceutical Applications. Edited by Hondermarck, H. (Kluwer Academic Press).

CHAPTER 2 (In full)
Appendix 10 page 361

Collins MO, Yu L, Husi H, Blackstock WP, Choudhary JS, Grant SG. (2005) *Robust enrichment of phosphorylated species in complex mixtures by sequential protein and peptide metal-affinity chromatography and analysis by tandem mass spectrometry*. Science STKE. 2005(298):pl6.

CHAPTER 3 (In part)
Appendix 11 page 385

Collins MO, Yu L, Coba MP, Husi H, Campuzano I, Blackstock WP, Choudhary JS, Grant SG. (2005) *Proteomic analysis of in vivo phosphorylated synaptic proteins*. Journal of Biological Chemistry. 280(7):5972-82. (Editors choice Science. (2005) The Big Picture of Synaptic Phosphorylation. Vol 307, Issue 5714.)

CHAPTER 5 (In part)
Appendix 12 page 396

Collins, MO, Husi , H, Yu, L, Brandon, JM, Anderson, CNG, Blackstock, WP, Choudhary, JS, Grant, SGN. (2005) *Molecular characterization and comparison of the components and multi-protein complexes in the postsynaptic proteome. Journal of Neurochemistry*. (In press, 2006)

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION AND AIMS

1. Introduction

1.1 Introduction and Aims

The field of proteomics aims to provide a comprehensive view of the characteristics and function of every protein present within a given biological system. The genome, which encodes the entire complement of genes in an organism, has been seen as the fundamentally complex component of cells. However it has emerged in recent years that its complexity is overshadowed by that of the transcriptome, which describes the full complement of mRNA transcripts transcribed from the genome of the cell. This in turn gives rise to yet another level of complexity see in that of the cellular proteome. This proteome is derived from the variety of alternative splicing events that occur at the level of transcription and the use of alternative start and stop sites and frame-shifting which occurs at the level of translation. This complexity is compounded by the 300 or more different protein modifications, which have been, reported (Aebersold and Goodlett, 2001) of which phosphorylation and glycosylation are prototypic (Figure 1.1).

Comprehensive proteome analysis also seeks to elucidate protein sub-cellular distribution and protein-protein interactions, which are crucial in determining the integration of large numbers of proteins into interaction networks, and functional molecular complexes that ultimately perform a myriad of cellular functions (Figure 1.1). Finally, the proteome is highly dynamic and temporal and thus describes the complement of proteins expressed in a cell at a single time point. The synapse is a sub-cellular specialisation that converts patterns of electrical activity into cellular memory. This processing of electrical information is mediated by the protein components of the synapse and the phosphorylation of these components is very important in this process. This specialised proteome can be isolated and characterised by mass spectrometry in terms of its components and phosphorylation. Establishment of such a molecular map permits insight into the organisation and functioning of the synaptic proteome.

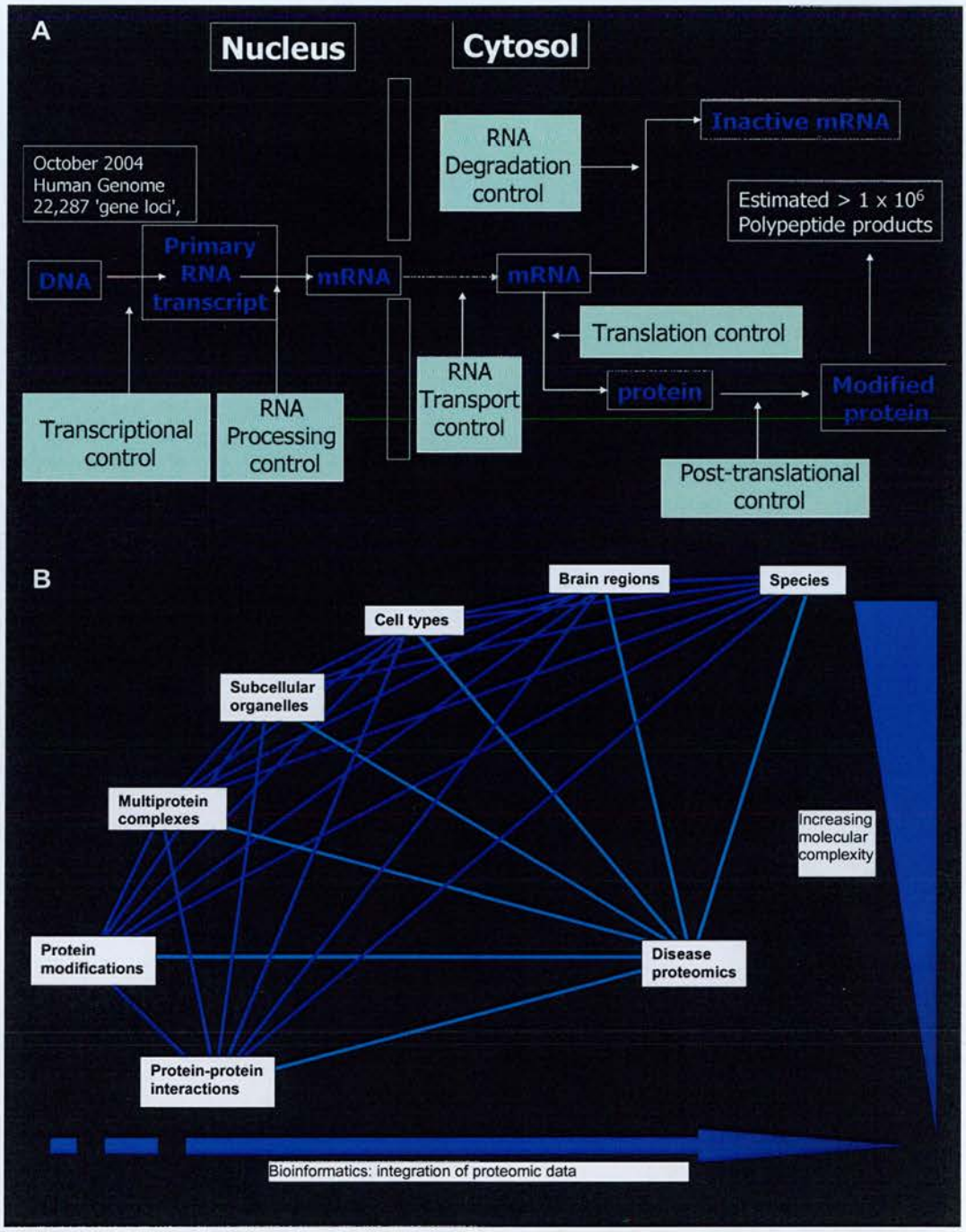


Figure 1.1 - Overview of levels of proteome complexity. (A) There are at least six points where the complexity of the information encoded in the genome is increased by transcriptional, translational and post-translational control. At the post-translational level, more than 300 different types of modifications, including protein phosphorylation exist. These processes would easily transform the 22, 000 gene loci in the human genome into an estimated 1 million different polypeptide forms, greatly increasing genome to proteome complexity. (B) Once a polypeptide has undergone splicing, post-translational modifications etc, there are further levels of proteome complexity in how proteins exist in a cell. Proteomes from different species, tissues or cell types differ not only in their composition but in their protein interactions, protein complexes, and organelles. Complex diseases are usually caused by changes in more than one level in this complexity and therefore disease proteomics should take this into account. An essential point about this complexity is the need for adequate bioinformatic approaches to integrate such proteome data.

1.2 Proteomics in neuroscience

There are many approaches used to study large-numbers of proteins in the brain, from comparative analysis of protein expression to mapping molecular complexes. A number of examples of such approaches will be reviewed to highlight the diversity of proteomic applications to neuroscience.

1.2.1 Two-dimensional electrophoresis

The biochemical approach of two-dimensional electrophoresis which has become the classical proteomic approach to whole proteome analysis has the capacity to display a large number of proteins expressed the studied system under given physiological conditions. Construction of global expression maps for defined proteomes is the most widely used application of proteomics and when used in combination with mass spectrometry (MS) techniques can be a powerful approach. There have been a number of studies focused on global neuroproteomics from whole brain analysis to the analysis of synaptic components. Two-dimensional maps have been constructed for whole human (Langen et al., 1999) mouse (Gauss et al., 1999) and rat brains (Fountoulakis et al., 1999) to analyse the entire brain proteome. These studies yielded more than 200 discrete proteins with significant overlap between datasets. The mouse brain proteome study is the most comprehensively characterised to date by 2D electrophoresis with 8767 resolved spots and over 300 proteins identified by a combination of proteomic and genetic methods, (Figure 1.2).

A number of brain region proteomes have been studied to investigate differences in protein expression. Preliminary analysis of the mouse cerebellar proteome has identified 30 proteins (Beranova-Giorgianni et al., 2002) and analysis of the porcine cerebellum led to identification of 56 spots (Friso and Wikstrom, 1999). A developmental proteomic study of the rat cerebellum resulted in resolution of over 3000 spots and identification of 67 of these (Taoka et al., 2000). Most proteins

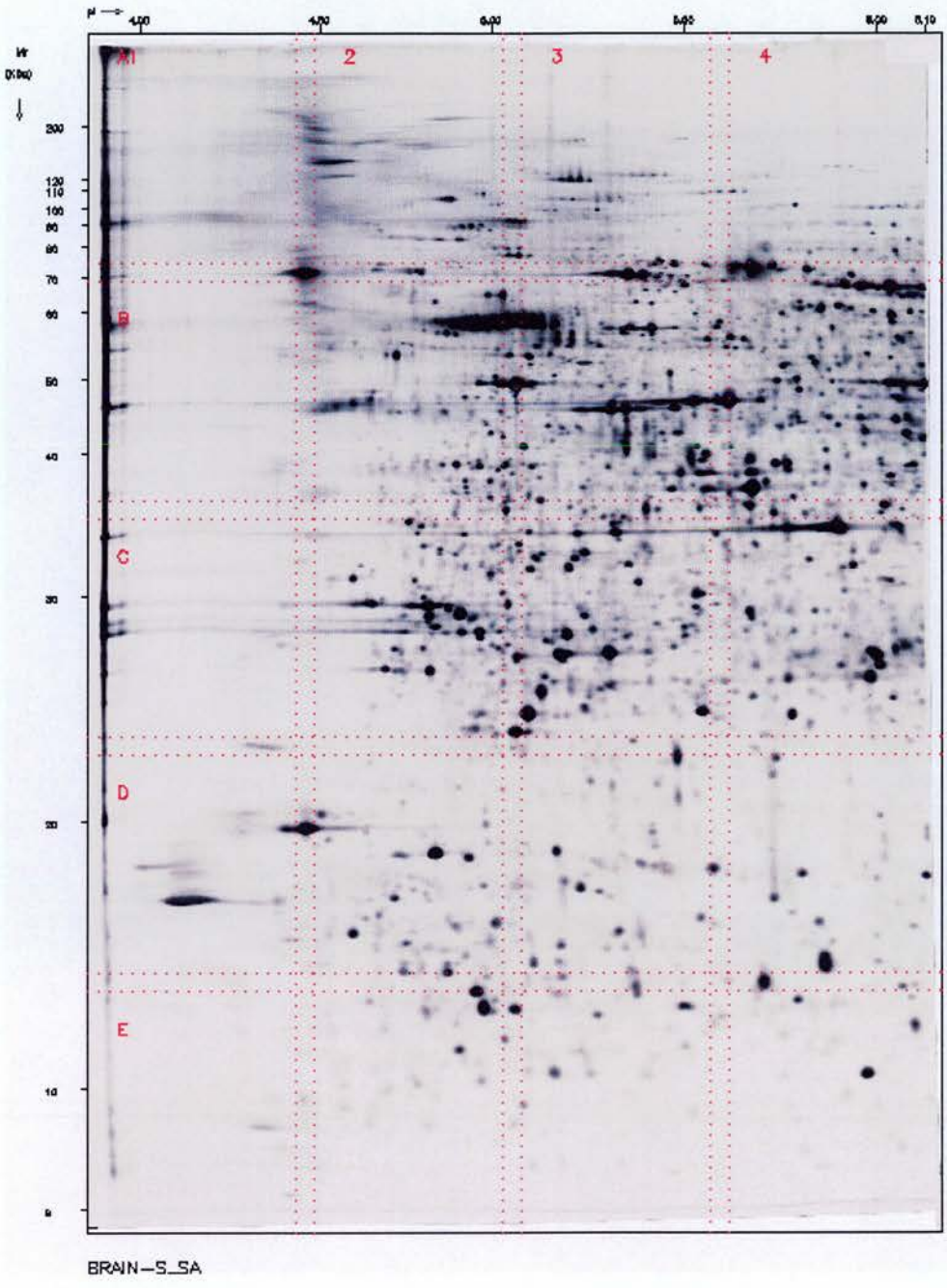


Figure 1.2 - Protein standard map of the mouse brain supernatant fraction obtained by large-gel 2-D electrophoresis
(image from Gauss, C., Kalkum, M., Lowe, M., Lehrach, H. & Klose, J. Analysis of the mouse proteome. (I) Brain proteins: separation by two- dimensional electrophoresis and identification by mass spectrometry and genetic variation. Electrophoresis 20, 575-600. (1999))

showed an increase in abundance as the cerebellum matured, however, 42 spots appeared to be exclusively expressed in the immature cerebellum. Some of the latter were identified by MS and included proteins with defined roles in nervous system development. Although, the approach of using 2D gels in conjunction with MS techniques is a classical proteomic approach, it is evident that it falls short as a sole platform for global proteomics. The range of biochemically different proteins, co-migration of proteins and difficulty of resolution of low abundance molecules has hampered the use of this technology to comprehensively analyse highly complex mixtures of proteins present in brain preparations. Problems associated with 2D gel resolution of membrane proteins have been overcome to some extent (Santoni et al., 2000), however gel-free proteomic solutions such as multidimensional protein identification technology (MudPIT) (Wu et al., 2003) allow greatly increased coverage of the membrane proteome. However, 2D-electrophoresis will remain a useful tool in proteomics, due to sophisticated image analysis, software and fluorescent dyes, which now permit more sensitive and accurate comparative proteome analysis.

1.2.2 Subcellular proteomics

The observation of cells under an electron microscope for the first time in the 1940's revealed a previously unexpected level of complexity (Palade, 1966). Extensive compartmentalisation was observed and was found to be achieved by intracellular membranes. This realisation prompted intensive investigation of sub-cellular organelles and in the following decades led to the development of a number of methodologies to physically isolate cellular organelles. The identification of the protein composition of sub-cellular organelles has been a significant challenge for many years and is only now becoming routine with sophisticated mass spectrometry-based platforms (Table 1.1). Most sub-cellular organelles are isolated using density gradient centrifugation, in which a gradient of sucrose solutions is used to partition different organelles according to their density. A limitation of this

approach is the difficulty in assessment of the degree of purity of the isolated fractions. Usually, the presence and absence of sub-cellular markers is used to indicate the level of contamination of co-sendimenting cellular debris.

Organelle	Number of proteins	Reference
Phagosome	525	(Houde et al., 2003), (Garin et al., 2001), (Gagnon et al., 2002), (Dermine et al., 2001)
Nucleolus	489	(Andersen et al., 2005)
Spliceosome	292	(Neubauer et al., 1998), (Rappsilber et al., 2002)
Exosome	295	(Pisitkun et al., 2004)
Lysosome	215	(Bagshaw et al., 2005)
Centrosome	64	(Andersen et al., 2003)
Mitochondria	847	(Cotter et al., 2004)
Clathrin-coated vesicles	209	(Blondeau et al., 2004)
Lipid raft	241	(Foster et al., 2003)
Golgi	136	(Bell et al., 2001), (Wu et al., 2000), (Wu et al., 2000)
Nuclear pore	174	(Rout et al., 2000)
Neutrophil granules	286	(Lominadze et al., 2005)

Table 1.1 – Summary of organellar proteomic studies

Recently, with the advent of quantitative mass spectrometry-based approaches, it is possible to asses the relative amounts of each protein present in a fraction compared to the starting material of the isolation. This approach was successfully applied to the identification of proteins in the centrosome (Andersen et al., 2003). In this study, they quantitatively profiled hundreds of proteins across several centrifugation fractions and identified proteins which were quantitatively enriched in centrosomes (Andersen et al., 2003).

Traditionally, the protein content of purified organelles are separated by either 1D or 2D SDS-PAGE and protein containing bands or spots are excised, digested and subjected to analysis by mass spectrometry. 1D SDS-PAGE has the advantage of being particularly well suited for separation of membrane proteins, while 2D SDS-PAGE provides better separation of protein species. However, both of these

approaches require band or spot excision, which is very labour intensive, not very reproducible and causes significant sample loss due to incomplete peptide extraction from the gel slices/spots. These difficulties have been overcome, in part, by gel-free separations of peptides such as MudPIT (multidimensional protein identification technology) pioneered by John Yates, in which the protein sample is denatured, digested in solution and the resultant peptides separated in the first dimension by ion-exchange chromatography and in the second dimension, by reverse-phase chromatography (Washburn et al., 2001). The first application of this methodology was the profiling of the yeast proteome and this facilitated identification of 1484 proteins (Washburn et al., 2001), far exceeding the numbers of protein identifications from traditional methods.

Complete characterisation of the protein inhabitants of every organelle in a cell type would provide valuable information on protein localisation, provide annotation for the large numbers of proteins of unknown function and would provide a starting point for large-scale functional investigation of the biology of each organelle. Such data would be invaluable for neuroscience research and application of this approach to synapses, which are specific to neurons, would certainly provide an unprecedented parts list and platform for the study of synaptic biology.

1.2.3 Multiprotein complex proteomics

1.2.3.1 NMDA receptor-adhesion protein signalling complex

Protein-protein interactions are the basis on which cellular structure and function are built and interaction partners are an immediate lead into biological function that can be exploited for therapeutic purposes. Proteomics has been shown to make a crucial contribution to the study of protein-protein interactions (Neubauer et al., 1997). Proteomic approaches to tackle multiprotein complexes usually involve purification of the entire complex by a variety of affinity methods, and protein

identification by western blotting or mass spectrometry based approaches. The first proteomic analysis of multiprotein complexes relevant to the brain was the purification and identification of the molecular constituents of the NMDA receptor-adhesion protein signalling complexes (NRC) (Husi et al., 2000), (Husi and Grant, 2001).

1.2.3.1.1 Methodology

Similar to most neurotransmitter receptors, the NMDAR has a low abundance in neurons compared with the overall protein content. Therefore, it is important to generate highly enriched samples containing the target molecule. This can be achieved by a variety of methods, generally involving affinity chromatography steps. These can range from ligand affinity (drugs, substrates or co-factors) to the more widely used approach of immunoprecipitation, involving specific antibodies to the molecule of interest. A major drawback in using antibody-based approaches for proteomic analysis is the introduction of the IgG itself, which interferes with the analysis by overshadowing proteins of interest that co-migrate during the gel-separation. This same problem arises with protein-affinity based methods, in which a recombinant protein carries one or several affinity-tags (e.g. poly-histidines), or is expressed as a GST-fusion protein. This problem can be eliminated using either the tandem affinity purification (TAP) method (Rigaut et al., 1999) or with the use of non-protein based methods, which exploit substrate or ligand affinities, such as peptide-based isolation procedures (Husi and Grant, 2001). All of these ligands, groups or proteins are then coupled to a resin either irreversibly (usually in the case of small molecules), or through a secondary affinity linkage specific for the bait-molecule. Neuronal extracts are then incubated with the resins bearing the bait and bound proteins separated from each other as well as from the matrix by conventional gel analysis.

Identification of isolated proteins can be readily achieved using western blotting or mass spectrometry. Western blotting (also known as immunoblotting) utilises specific antibodies to detect a protein that has been transferred from a gel onto a membrane. A major advantage of western blotting is its sensitivity and ability to detect amounts of protein beyond the range of current sequencing-based or mass spectrometry methodologies. It is obvious that this approach is biased by the assumption that a given molecule might be present, and cannot be used to identify unknown or unsuspected proteins. However, there is a large amount of neurobiological evidence concerning the putative involvement of molecules that influence many target proteins and such proteins pose ideal targets for western blotting approaches. Additionally, changes in protein levels, as a result of modulation or modification in the receptor's environment, can easily be visualised using specific antibodies against the proteins under investigation. A disadvantage of this method is its labour intensive nature and thus difficulty in performing on a large scale.

The main approaches to protein identification are peptide-mass fingerprinting using matrix-assisted laser desorption/ionisation (MALDI) (Henzel et al., 1993), (Shevchenko et al., 1996) and electrospray ionisation using a tandem mass spectrometer (Shevchenko et al., 1997) or a combination of both (Shevchenko et al., 2000). Recent developments led to new methodologies using direct liquid chromatography/tandem mass spectrometry (LC/MS-MS) techniques, which allow fast and reliable mapping of very low amounts of peptides in an almost completely automated fashion (Shevchenko et al., 2000). Furthermore, using this method, it is also possible to determine if a protein or peptide carries any post-translational modifications, such as glycosylation or phosphorylation (Pandey et al., 2000).

The sample is usually separated on a gel and the bands of interest excised and treated with trypsin in order to generate fragments of proteins that can subsequently be extracted from the gel slice. There are several reasons for analysing

peptides rather than proteins. The peptide's mass can be measured at much greater precision than the mass of intact proteins. Peptides give rise to a library of molecular masses that are directly derived from the proteins, which can be identified using computational techniques.

Using the approaches described above the NRC was shown in 2000, to comprise 77 proteins organised into receptor, adaptor, signalling, cytoskeletal and novel proteins, of which 30 are implicated from binding studies and another 19 participate in NMDAR signalling (Figure 1.3) (Husi et al., 2000).

NMDAR and metabotropic glutamate receptor subtypes were linked to cadherins and L1 cell-adhesion molecules in complexes lacking AMPA receptors. These neurotransmitter-adhesion receptor complexes were bound to kinases, phosphatases, GTPase-activating proteins and Ras with effectors including MAPK pathway components. A striking feature of the composition is that there appears to be 'modules' or sets of signalling proteins that are known to comprise key components of signal transduction pathways that can be distinctly regulated. For example, all of the molecules necessary to induce the phosphorylation of mitogen-activated protein kinase (MAPK) following NMDAR stimulation are present in the NRC (including CaMKII, SynGAP, Ras, MEK and ERK) (Figure 1.5). There modules may allow the NMDAR and mGluR to integrate signals within the complex and then couple to down-stream cellular effector mechanisms, such as trafficking of AMPA receptors and cytoskeletal changes that mediate structural and physiological plasticity. Furthermore, at least 18 NRC constituents are regulated by synaptic activity indicating that the composition of the complex is dynamic. In the hippocampus, these activity-dependent genes are known to undergo specific temporal changes following the induction of plasticity.

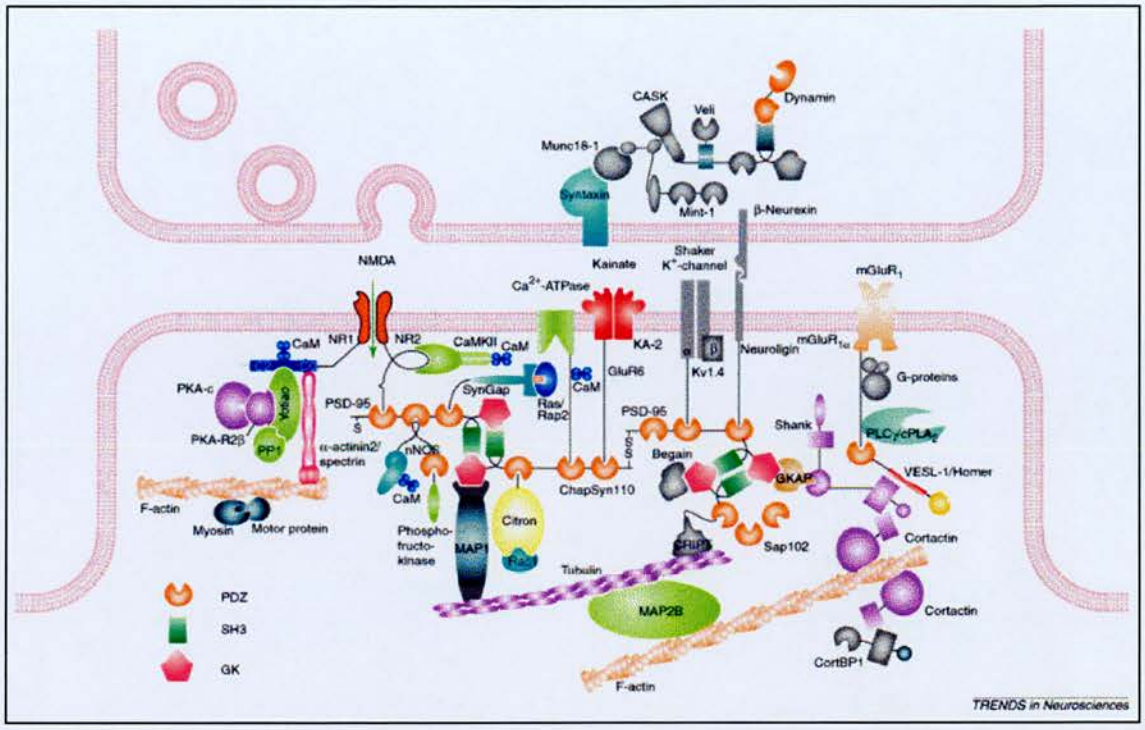


Figure 1.3 – Schematic representation of the NMDA receptor complex (NRC). Proteins depicted here (except for those in grey, which were found in the literature) were found in the NRC by proteomic screens. Protein-protein interactions were mined from the literature and can be found in PPID (www.PPID.org).

Adapted from "Proteomics of the nervous system" Husi H, Grant SG. Trends Neurosci. 2001 24(5):259-66.

1.2.3.2 Other membrane receptor complexes

Since characterisation of the NRC, two other receptor complexes have been reported. P2X receptors are ATP-gated ion channels in the plasma membrane, and activation of the P2X7 receptor also leads to rapid cytoskeletal re-arrangements such as membrane blebbing. 11 proteins in human embryonic kidney cells that interact with the rat P2X7 receptor were identified by affinity purification followed by mass spectroscopy and immunoblotting (Kim et al., 2001). Also using a proteomic approach based on peptide affinity chromatography followed by mass spectrometry and immunoblotting, another study identified 15 proteins that interact with the C-terminal tail of the 5-hydroxytryptamine 2C (5-HT (2C)) receptor, a GPCR (Becamel et al., 2002). These proteins include several synaptic multidomain proteins containing one or several PDZ domains (PSD95 and the proteins of the tripartite complex Veli3-CASK-Mint1), proteins of the actin/spectrin cytoskeleton and signalling proteins. Co-immunoprecipitation experiments showed that 5-HT (2C) receptors interact with PSD-95 and the Veli3-CASK- Mint1 complex in vivo. Electron microscopy also indicated a synaptic enrichment of Veli3 and 5-HT (2C) receptors and their co-localisation in microvilli of choroidal cells. These results indicate that the 5-HT (2C) receptor is associated with protein networks that are important for its synaptic localisation and its coupling to the signalling machinery (Becamel et al., 2002).

It is now emerging that multiprotein signalling complexes or "signalling machines", of which the NRC is prototypic, are responsible for orchestrating the complex signalling events at ionotrophic glutamate receptors (Husi et al., 2000), ATP receptors (Kim et al., 2001) and G-protein coupled receptors (Becamel et al., 2002). They seem to share a common organisation of receptor, protein scaffold, (in which the signalling molecules are localised) and membrane to cytoskeletal interactions. Therefore, study of individual receptor multiprotein complexes such as the NRC will most likely provide a mechanistic basis that will be very useful and applicable

for the systematic identification and study of membrane receptor complexes associated with the estimated 100 other receptor types believed to exist in the brain.

1.3 One hundred years of protein phosphorylation

Phosphorylation was first detected in the protein Vitellin, in 1906 (Alsberg, 1906) and phosphoserine was detected in Vitellin nearly thirty years later (Levene, 1933). It took another 20 years to discover that this phenomenon was actively catalysed by enzymes (Burnett and Kennedy, 1954). It was in the 1950's when phosphorylation became of interest to the fields of metabolism and hormone action, in which phosphorylase A was intensively studied. The effects of epinephrine and glucagon on phosphorylase activity were investigated and it was found that these hormones stimulate cAMP (3', 5'-cyclic-adenosine monophosphate), which regulates phosphorylase activity (Rall and Sutherland, 1958). This regulation was subsequently shown to be mediated by incorporation of phosphate into phosphorylase A, converting the enzyme to phosphorylase B, the active form (Fischer and Krebs, 1955). The necessity for cAMP for phosphorylase activation promped speculation that there was a cAMP-dependent kinase that was responsible for phosphorylase A activation and it was the first time a signalling cascade was proposed. In 1968, this cAMP-dependent kinase (PKA) was purified (Walsh et al., 1968) and subsequently, intensive study of the first protein kinase revealed that is was a tetramer composed of regulatory and catalytic subunits (Taylor et al., 1990).

The significance of phosphorylation was realised through the 1970's and the early 1980's. Important studies by Paul Greengard in the field of synaptic phosphorylation paved the way for the countless studies on synaptic phosphorylation ever since. These studies were the first to report phosphorylation of synaptic membrane proteins in response to cAMP stimulation (Ueda et al., 1973), (Ueda and Greengard, 1977). The first calmodulin-dependent kinase was identified

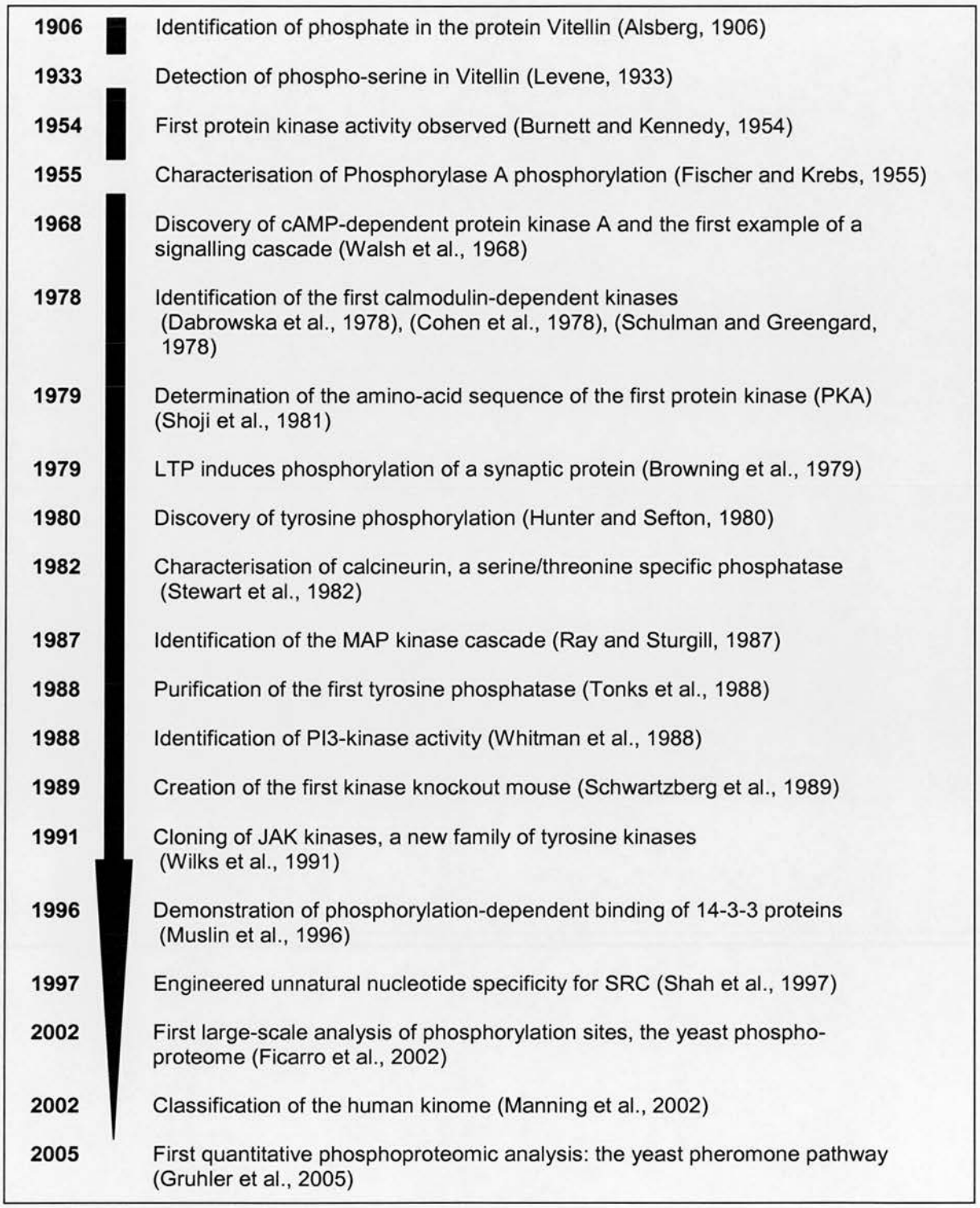


Figure 1.4 – One hundred years of protein phosphorylation

in late 1970's which included CaMKI, CaMKII (Schulman and Greengard, 1978), phosphorylase kinase (Cohen et al., 1978) and myosin light-chain kinase (Dabrowska et al., 1978). In 1979, the amino-acid sequence of the first kinase (PKA) was determined (Shoji et al., 1981). In the same year, an exciting report that showed that LTP (Long-term potentiation) induced phosphorylation of a synaptic protein (Browning et al., 1979), a key finding which provided a link between electrophysiological phenomena with protein phosphorylation.

The next major finding occurred in 1980, when Tony Hunter discovered tyrosine phosphorylation (Hunter and Sefton, 1980), which is now recognised as a very important and specific type of phosphorylation. A number of important serine/threonine protein kinases were characterised in the 1980's providing insight into how the process of reversible phosphorylation is controlled (Ingebritsen and Cohen, 1983) (See Ingebritsen and Cohen, 1983 for review). The late eighties and nineties were characterised by the signalling cascade, with much attraction focused on the MAP kinase cascade. In addition, this period saw the identification of the first protein tyrosine phosphatase (Tonks et al., 1988), PI3-kinase (Whitman et al., 1988), and JAK kinases (Wilks et al., 1991). One of the major landmarks of the late 1980's was the creation of the first protein kinase knockout in mouse (Schwartzberg et al., 1989), this was closely followed by a Fyn knockout, the first kinase knockout mouse relevant to neuroscience. Analysis of this Fyn tyrosine-kinase mutant revealed defects in LTP, spatial learning and hippocampal development and thus, established for the first time, the importance of a specific kinase in synapse function (Grant et al., 1992).

The 1990's brought many studies focused on specific kinases and phosphorylation sites and also the demonstration of phospho-dependent binding of adapter proteins such as 14-3-3 proteins (Muslin et al., 1996). In 1997, the sophisticated use of genetic engineering allowed manipulation of the residues in Src tyrosine kinase responsible for nucleotide specificity (Shah et al., 1997). These residues were altered to allow the

use of unnatural nucleotides which the kinase would use to specifically label its substrates. The new millennium brought phosphorylation research into proteomics and the new field of phosphoproteomics begun. This new approach to study phosphorylation was exemplified by the work of Ficarro et al. in 2002. This reported identification of hundreds of phosphopeptides and phosphorylation sites from S. *cerevisiae* by using a combination of phosphopeptide enrichment and tandem mass spectrometry (Ficarro et al., 2002) and set the standard for phosphoproteomic experiments to follow. In the same year, the portion of the genome encoding protein kinases was analysed and for the first time a complete set of kinase sequences were assembled and classified (Manning et al., 2002). The result, the human kinome, which contains 518 protein kinases, is a useful data resource and profiling of kinomes is becoming relevant to therapeutics (Vieth et al., 2005).

The assembly of the human kinome as well as the possibility of identification of hundreds of phosphorylation sites started to reveal the complexity of the phosphoproteome and prompted researchers to ask how to understand and address this complexity. This has been addressed in part by using quantitative mass spectrometry to monitor protein phosphorylation in response to specific conditions and stimuli, e.g. the yeast pheromone pathway (Gruhler et al., 2005). In addition, attempts to integrate phosphoproteomic data have been made, using systems biology type approaches such as network analysis to derive understanding of complex signalling pathways. This has been applied to derive network and pathway maps of EGF (epidermal growth factor) receptor signalling (Oda et al., 2005), one of the best characterised signalling systems to date.

1.4 Phosphorylation in neuronal signalling

In the 1894, Ramón y Cajal proposed a theory that changes in synaptic strength (synaptic plasticity) maybe involved in memory formation (Ramón y Cajal, 1894) and this basic mechanism is now the most widely accepted theory of memory

storage on a cellular level. The phenomena of long-term potentiation (LTP) and long-term depression (LTD), the enduring increase or decrease in the amplitude of excitatory-postsynaptic potentials as a result of high or low frequency stimulation of afferent pathways, are considered to be the electrophysiological manifestations of synaptic plasticity.

Since the discovery of cAMP-dependent phosphorylation of proteins in the synaptic membrane in the early 1970's by Paul Greengard and co-workers (Ueda et al., 1973), (Ueda and Greengard, 1977), the field of synaptic phosphorylation has become an essential part of the study of synaptic physiology. It was realised in the late 1970's that repetitive stimulation of the of the Schafer collateral-commissural system of the rat hippocampus induced LTP and resulted in phosphorylation of a 40 kDa protein (Browning et al., 1979), which was later identified as alpha subunit of pyruvate dehydrogenase (Browning et al., 1981). This was the first demonstration of the link between the electrophysiological phenomenon of LTP and the biochemical manifestation of LTP; synaptic protein phosphorylation. Since these initial insights into synaptic signalling, much attention has been focused on phosphorylation and its role in synapse function (Figure 1.5 for overview), a number of these key findings in glutamatergic synaptic signalling will be reviewed here.

1.4.1 Overview of glutamatergic signals

Glutamatergic synapses contain ionotropic (N-methyl-D-aspartate (NMDA) and -amino-3-hydroxy-5-methylisoxazole-4-propionic acid (AMPA)) receptors and metabotropic glutamate receptors (mGluRs), which exist as multiple subtypes (mGluR1-8). All of these receptors are activated by glutamate release from the presynaptic membrane; NMDA receptors allow influx of calcium ions, an important the second messenger, AMPA receptors are permeable to potassium and sodium ions, while metabotropic glutamate receptors are coupled to G-proteins and participate in regulation of intracellular stores of calcium. LTP is triggered by

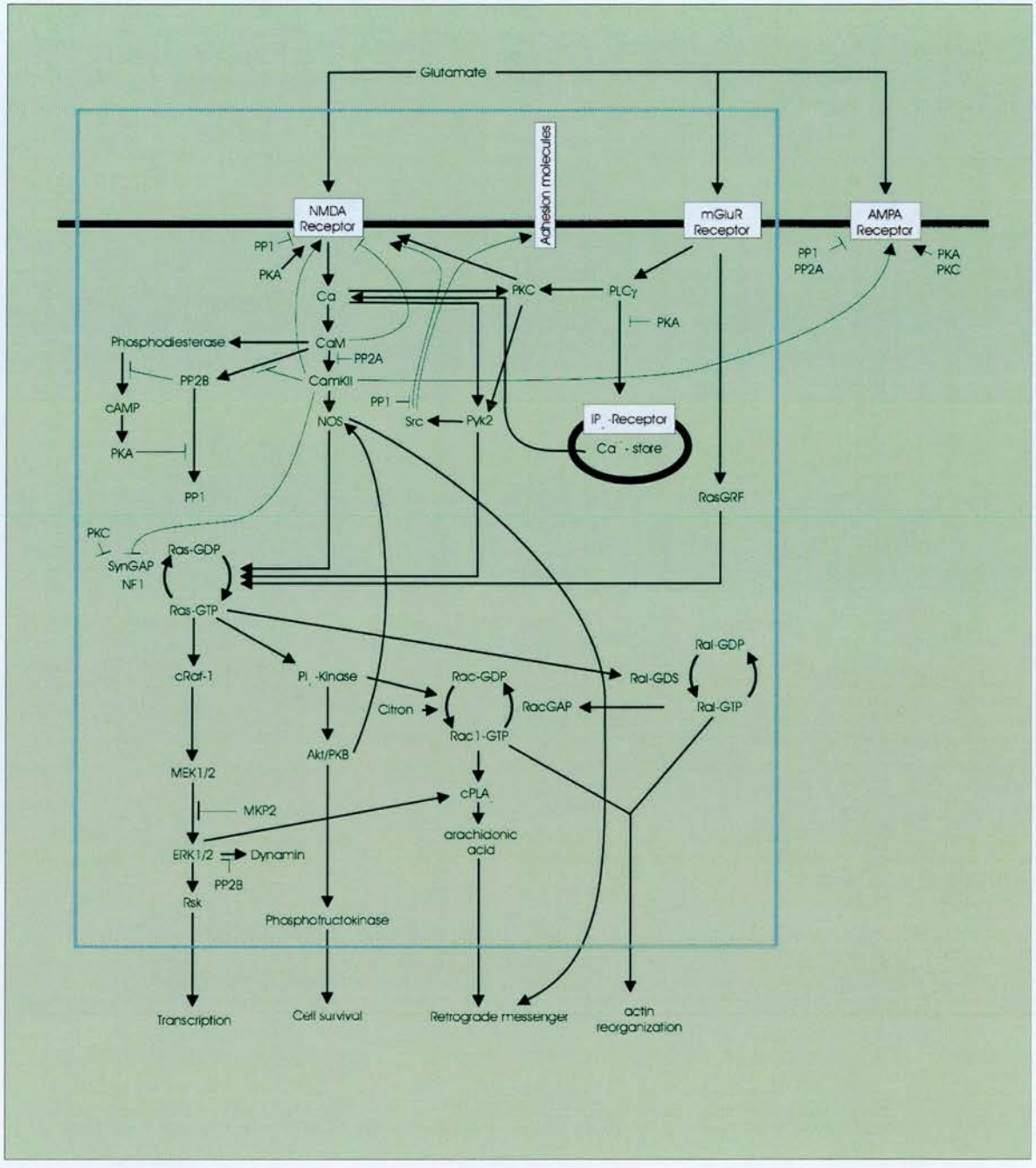


Figure 1.5 – An overview of signalling at the synapse. Activation of postsynaptic glutamate receptors causes increases in the level of intracellular calcium and other second messengers such as cAMP. A complex regulatory mechanism mediated by kinases, phosphatases and signalling molecules modulate channel function and signalling pathways which control processes such as transcription, cell survival and actin reorganisation.

Adapted from "Proteomics of the nervous system" Husi H, Grant SG. Trends Neurosci. 2001 24(5):259-66.

transient activation of NMDA receptors and is expressed as a persistent increase in synaptic transmission through AMPA receptors. High frequency synaptic stimulation causes calcium ion influx through the NMDA receptor and activation of CaM kinase II (CaMKII) and induction of LTP. On the other hand, low frequency stimulation causes moderate but prolonged increases in calcium levels which activate protein phosphatases and induction of LTD.

Analysis of the modulation of these receptors and associated proteins by phosphorylation, mediated by kinases and phosphatases, is essential to understand how signals are propagated and resolved by synapses.

1.4.2 NMDA receptor phosphorylation

The NMDA receptor NR1 and NR2 subunits are extensively phosphorylated, which permits modulation of NMDAR activity (see (Yamauchi, 2002) and (Salter and Kalia, 2004) for review), localisation (Tingley et al., 1997) and trafficking (Chung et al., 2004). The NR1 subunit is phosphorylated by PKA and PKC, which regulates its intracellular localisation (Tingley et al., 1997). Yotiao is a NR1-binding protein that also binds PKA and the constitutively active protein phosphatase PP1. Association of PP1 with this complex limits the channel activity but is opposed by the effects of PKA, which rapidly enhances NMDAR activity (Westphal et al., 1999). This is a classic example of how adapter proteins such as Yotiao can facilitate the localisation of a kinase and a phosphatase proximal to their substrate, in this case NR1.

The NR2 subunits have long C-terminal tails which contain multiple protein interaction sites that nucleate the complex of proteins attached to the receptor. The NR2B subunit is bound to and phosphorylated by CaMKII (Omkumar et al., 1996). CaMKII is activated by NMDAR receptor-mediated calcium influx and activated CaMKII binds to NR2B and results in sustained Ca2+/calmodulin (CaM)-independent (autonomous) kinase activity (Bayer et al., 2001). This activated

CaMKII-NR2B complex is also regulated by PP1 which causes dissociation of CaMKII from the PSD (Yoshimura et al., 1999). PSD-95, which is an important synaptic scaffolding protein involved in many synaptic processes (Kim and Sheng, 2004), binds to the C-terminal tail of NR2B. Recently, it was shown that phosphorylation of this binding site on NR2B by Casein Kinase 2 (CK2) disrupts the PSD-95-NR2B interaction (Chung et al., 2004). Phosphorylation of this site is regulated by NMDAR activity and Ca^{2++}/CaMKII and results in decreased surface expression of NMDA receptors (Chung et al., 2004).

The NR2 subunits are heavily tyrosine phosphorylated by Src (Yang and Leonard, 2001) and Fyn (Nakazawa et al., 2001) tyrosine kinases. Fyn phosphorylation of NR2B is significantly increased after induction of LTP in the CA1 region of the hippocampus and is reduced in Fyn mutant mice (Nakazawa et al., 2001). Src phosphorylates tyrosine residues in NR2A responsible for zinc binding and phosphorylation of these sites results in potentiation of the NMDA receptor by reducing tonic zinc inhibition (Zheng et al., 1998).

1.4.3 AMPA receptor phosphorylation

The AMPA receptor subunits are also heavily regulated by phosphorylation. Similar to the NMDA receptor, phosphorylation of the AMPA receptor modulates receptor function as well as protein interactions with the receptor. There are two major sites of phosphorylation on the GluR1 subunit, S831 and S845 that are phosphorylated by CaMKII and PKA, respectively (Lee et al., 2000). CAMKII phosphorylation of GluR1 is correlated with an increase in the single-channel current of the AMPA receptor and this phosphorylation occurs during LTP in the CA1 region of the hippocampus (Barria et al., 1997). Investigation of the modulation of GluR1 by PKA phosphorylation of Ser-845 has suggested that phosphorylation of this residue may underlie the PKA-induced potentiation of AMPA receptors in neurons (Roche et al.,

1996). Induction of chemical LTP results in the dephosphorylation of GluR1 at serine 845, (PKA substrate), but not at serine 831 (PKC and CaMKII substrate) (Lee et al., 1998). In fact, LTP activates CaMKII, which phosphorylates S845, whereas LTP activates protein phosphatases such as PP1/PP2A, which dephosphorylate S845. It would therefore seem that regulation of these sites would contribute to the change in the efficacy of synaptic transmission that occurs during LTP and LTD.

Phosphorylation of the GluR2 subunit modulates its interaction with PDZ-containing binding proteins. S880, which is phosphorylated by PKC, is located in the C-terminus of GluR2 and occurs in a binding site for PICK-1 and GRIP-1. Phosphorylation of S880 decreases GRIP-1, but not PICK-1 binding (Chung et al., 2000). It has been demonstrated that PKC activation in neurons increases the S880 phosphorylation and recruits PICK1 to excitatory synapses. This correlates with fast internalisation of surface GluR2 suggesting that phosphorylation of S880 in GluR2 is important in the regulation of the AMPA receptor internalisation during synaptic plasticity (Chung et al., 2000). In addition, induction of LTD in hippocampal slices increases phosphorylation of S880 suggesting that the modulation of GluR2 interaction with GRIP1 and PICK1 may regulate AMPA receptor internalisation during LTD (Kim et al., 2001). The use of a peptide that disrupts the GluR2-PICK1 interaction was shown to inhibit LTD in the hippocampus, further supporting the importance of PKC phosphorylation in synaptic plasticity (Kim et al., 2001).

1.4.4 Modulation of the postsynaptic scaffold

As discussed for both the NMDA and AMPA receptors, phosphorylation is employed to modulate channel function as well as to modulate protein interactions. Dynamic modulation of protein interactions is emerging as a theme common among postsynaptic proteins and is of particular interest as the postsynaptic scaffold is thought to harbour many signalling proteins. If the postsynaptic scaffold is dynamically modulated then it would be possible that it plays a significant role in

orchestrating signalling pathways from the postsynaptic membrane. A selected number of examples of this will be introduced to discuss this theme further.

One of the first examples of regulation of a kinase by a scaffolding protein was that of AKAP79 (A Kinase Anchor Protein). AKAP79 forms a ternary complex with PKA and the calcium-dependent phosphatase, Protein phosphatase 2B (PP2B) (Coghlan et al., 1995) and regulates the action of this kinase-phosphatase pair by compartmentalisation with preferred substrates. AKAP79 targeting to the plasma membrane is regulated by PKC phosphorylation and calmodulin binding (Gomez et al., 2002). AKAP79 localisation is regulated by NMDA receptors through CaN activation and F-actin remodelling (Gomez et al., 2002). NMDA receptor activation also regulates dendritic spine localisation of PKA and CaN and association of the AKAP79-PKA complex with PSD-95 family members (MAGUKs) (Gomez et al., 2002). Glutamate receptors and PKA are recruited into a macromolecular signalling complex through direct interaction between the MAGUK proteins, PSD-95 and SAP97, and AKAP79 and phosphorylation of AMPA receptors is enhanced by a SAP97-AKAP79 complex that directs PKA to GluR1 (Colledge et al., 2000). This AKAP79 signalling complex is able couple intracellular calcium levels (through PP2B) to the PKA phosphorylation state of GluR1. This integration of intracellular signals can be transduced to GluR1 by the AKAP79 signalling complex and, regulation of AKAP79/150 postsynaptic targeting may be important for synaptic plasticity.

PSD-95 is a member of the membrane-associated guanylate kinase protein family (MAGUK) and is the best characterised of the group of scaffolding proteins at the synapse. It contains three PDZ domains, an SH3 domain and a GK domain, which together mediate interactions with at least 54 proteins. PSD-95 binds to the NMDA receptor (Niethammer et al., 1996) and is a crucial component of the NMDA receptor complex (Husi et al., 2000). PSD-95 also forms oligomers and is thought to exist as a scaffold in association with the postsynaptic membrane (McGee and Bredt,

1999), (Topinka and Bredt, 1998). The protein-tyrosine kinase Fyn has been shown to phosphorylate the NR2A subunit of the NMDA receptor in vivo (Tezuka et al., 1999). PSD-95 binds Fyn as well as the NMDA receptor subunits and promotes Fyn-mediated phosphorylation of NR2A (Tezuka et al., 1999). Different regions of PSD-95 interact with Fyn and the NMDA receptor subunits and therefore can recruit Fyn and other members of the protein-tyrosine kinases family to the NMDA receptor to regulate channel activity (Tezuka et al., 1999). In fact, PSD-95 and SAP102 (another MAGUK), can recruit PYK2, (protein-tyrosine kinase) to the NMDA receptor (Seabold et al., 2003). Induction of LTP requires the Pyk2/Src signalling pathway, which up-regulates the activity of NMDA receptors, so this recruitment of tyrosine kinases is an important regulatory mechamism at the synapse (Seabold et al., 2003).

SAP97 is also a MAGUK and is involved in the trafficking and clustering of glutamate ionotropic receptors to the postsynaptic membrane (Mauceri et al., 2004). It directly interacts with the NR2A subunit of NMDA receptor and CaMKII phosphorylation of S232 on SAP97 prevents binding to NR2A (Gardoni et al., 2003). CaMKII phosphorylation of an additional site (S39) drives SAP97 into spines and regulates the association of SAP97 with the postsynaptic scaffold (Mauceri et al., 2004).

It can be seen that from these examples of glutamate receptors and synaptic scaffolders that phosphorylation, both tyrosine and serine/threonine, is extensively used for modulation of function, protein interaction and localisation. Two particularly interesting points concerning phosphorylation at the synapse are that firstly, adaptor/scaffold proteins regulate kinases and phosphatases by specific recruitment to their substrates and secondly, phosphorylation of these adaptor/scaffold proteins regulates their association with other proteins and their localisation. This system of regulation does not have a clear hierarchy; the mediators of phosphorylation (kinases and phosphatases) regulate, and are regulated by their substrates, in this case scaffolding proteins. This process, in turn, regulates

receptor/channel activity as do signals (e.g. intracellular calcium) coming from receptors/channels regulate this system of kinases, phosphatases and scaffolding proteins.

Although, we have enough data concerning synaptic phosphorylation to draw broad conclusions such as these, the exact mechanisms and organisation of this complex regulatory system are unknown. Systematic analysis of these components by proteomic approaches will enable identification of the proteins at the synapse, identification of sites of phosphorylation of these proteins and with further developments in the field of kinomics/phosphoproteomics, annotation of kinases and phosphatases to their cognate substrates. This will provide a comprehensive platform for tackling synaptic signalling as a whole. However, functional characterisation of phosphorylation sites such as that described above is still very labour intensive and technological advances are necessary to functionally characterise phosphorylation sites on a proteome-wide scale.

1.5 Approaches to probe the phosphoproteome

1.5.1 Technology

Traditional biochemical approaches to analyse phosphorylation usually involve focused study of one phosphoprotein and often of one particular phosphorylation site. Antibodies that specifically recognise the phosphorylated epitope can be made and used to purify the phosphoprotein, detect the phosphoprotein in western blots or for imunohistochemistry or immunocytochemisty analyses. Generation and validation of such phospho-specific antibodies is a time consuming process taking up to one year to have an antibody that can be used confidently. The use of ^{32}P labelled γ-ATP for *in vitro* kinase reactions has been used widely and separation of labelled phosphorylated peptides by TLC (Thin layer chromatography) provides information about the number phosphopeptides and the intensity of phosphorylation. Analysis of these phosphopeptides from TLC plates has been

performed using mass spectrometry (MALDI-ToF (Matrix-associated laser desorption ionisation-time of flight)) and Edman sequencing. However, this methodology is suited for single phosphoprotein studies and the use of ^{32}P is hazardous.

1.5.1.1 Phosphoproteome enrichment strategies

Development of mass spectrometry and in particular liquid-chromatography-coupled mass spectrometry (LC-MS/MS) in the late 1990's permitted identification of large numbers of peptides in proteome-scale experiments (Gygi and Aebersold, 2000). The use of mass spectrometry for characterisation of phosphorylation on a proteome-scale was hampered by characteristics of phosphorylation itself. It has been estimated that one-third of any given proteome is phosphorylated and the stoichiometry of phosphorylation is often low, so to study phosphorylation on a global level, phospho-enrichment is a prerequisite.

Digestion of a proteome or even a portion of a proteome can result in generation of tens of thousands of tryptic fragments. This number of peptides is overwhelming in an LC-MS-MS experiment and as the mass spectrometer is usually operated in a mode which only selects the most intense ions for sequencing, only the most abundant portion of the proteome is sampled. Many signalling proteins are present at low abundance in cells and the portion of these proteins that are phosphorylated at any one time may also be quite low. Therefore in a standard LC-MS/MS experiment, the chance of detecting and sequencing a phosphopeptide is very low indeed. A number of strategies to overcome these problems were developed, which involve either physically enriching for phosphorylated species or isolating phospho-peptide specific signals in the mass spectrometer (Figure 1.6).

A number of chemical modification strategies have been developed in which the phosphate group has been replaced with a moiety that is chemically more stable

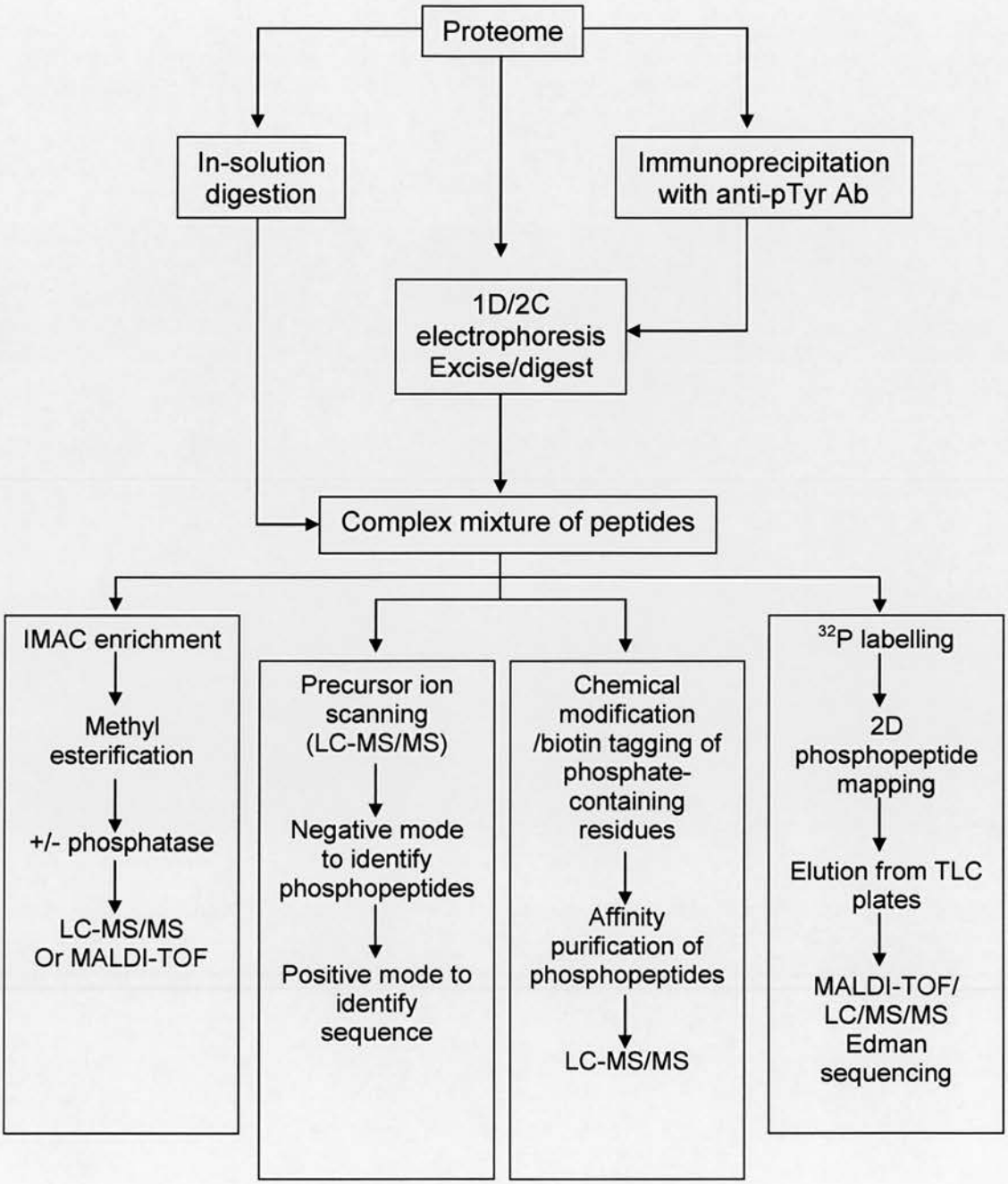
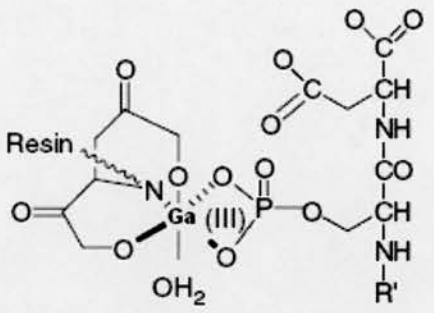
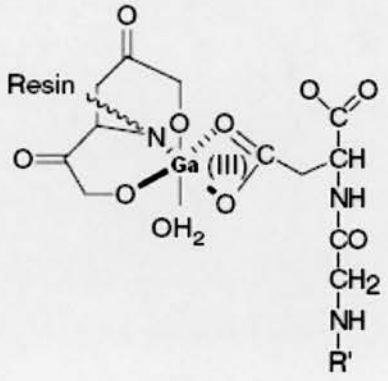


Figure 1.6 – Overview of techniques used to detect and map phosphorylation sites

A. Phosphopeptide enrichment by Gallium IMAC resin



B. Nonselective binding of carboxyl groups of peptides to the resin



C. Methyl-esterification of carboxyl group prevents non-specific binding

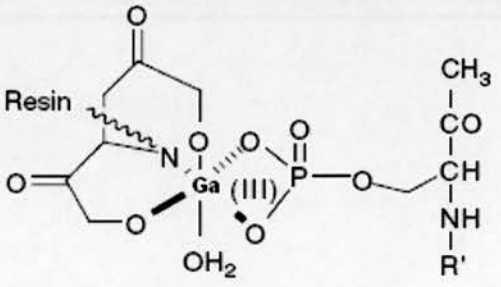


Figure 1.7 – Principle of immobilized metal affinity chromatography.

than phosphate. One such method employs β-elimination of the phosphate from phospho-threonine or serine and results in the formation of dehydroaminobutyric acid or dehydroalanine, respectively (Jaffe et al., 1998). The product can be detected using tandem mass spectrometry or in another method Michael addition is used to add a reactive thiol group after β-elimination to provide a linker for attachment of an affinity tag (Oda et al., 2001). This affinity tag is usually biotin and it permits purification of the chemically-modified (previously phosphorylated) peptides. This type of chemical modification of phosphate is not applicable to phospho-tyrosine residues and suffers from side-chain reactions in which cysteine and methionine can be modified. Other chemical modification protocols which are applicable to all phosphorylated residues have been developed (Zhou et al., 2001) but it involves many chemical reactions and have not been used successfully on real proteomic samples.

These chemical modification-based strategies were superseded in 2002 with the modification of the traditional IMAC (Immobilised metal-affinity chromatography methodology (Ficarro et al., 2002). IMAC purification of phosphopeptides is based on the attraction of negatively charged phosphate groups to positively charged immobilised (on a resin) metals (Figure 1.7). The main problem associated with this protocol is that acidic groups of aspartic and glutamic acid and electron donors (histidine) can bind to the positively charged metal, thereby contaminating the eluted sample. Ficarro *et al.* addressed this non-specific binding by chemical-esterification of acidic residues before IMAC enrichment. Application of this modified procedure to the yeast proteome permitted characterisation of 216 phosphopeptides (Ficarro et al., 2002).

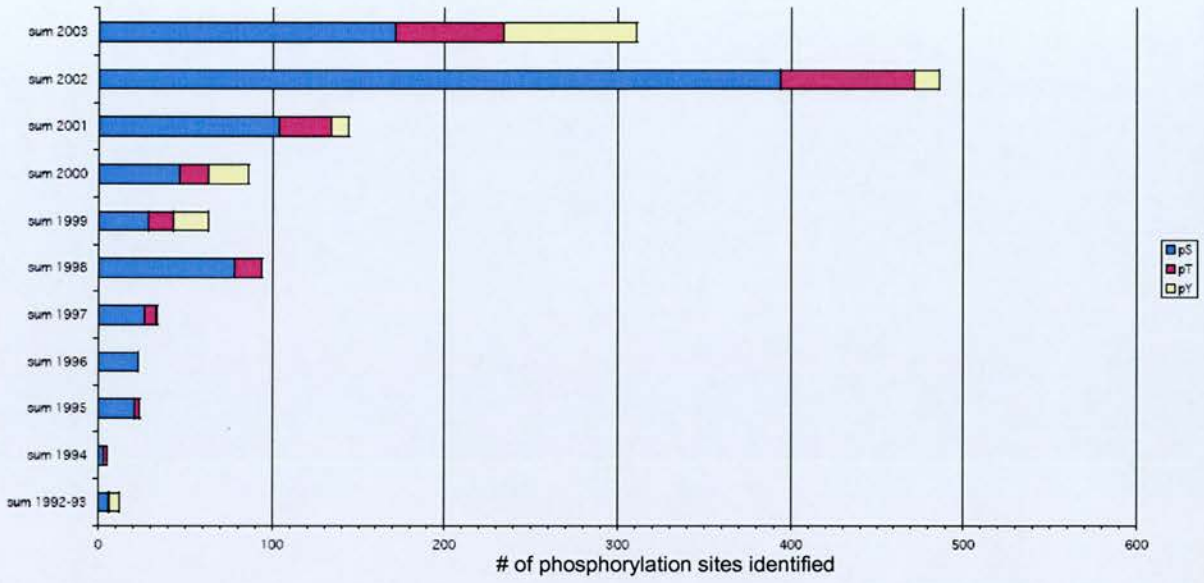


Figure 1.8 - Novel phosphorylation sites determined by mass spectrometry. The numbers of novel phosphorylation sites that have been identified by mass spectrometry as determined by literature mining are shown by year for the years 1992–2003. Contributions to the total number of sites per year of pSer (blue), pThr (red), and pTyr (yellow) are shown.

Adapted from "Mass spectrometric contributions to the practice of phosphorylation site mapping through 2003: a literature review." Loyet KM, Stults JT, Arnott D. Mol Cell Proteomics. 2005 Mar;4(3):235-45

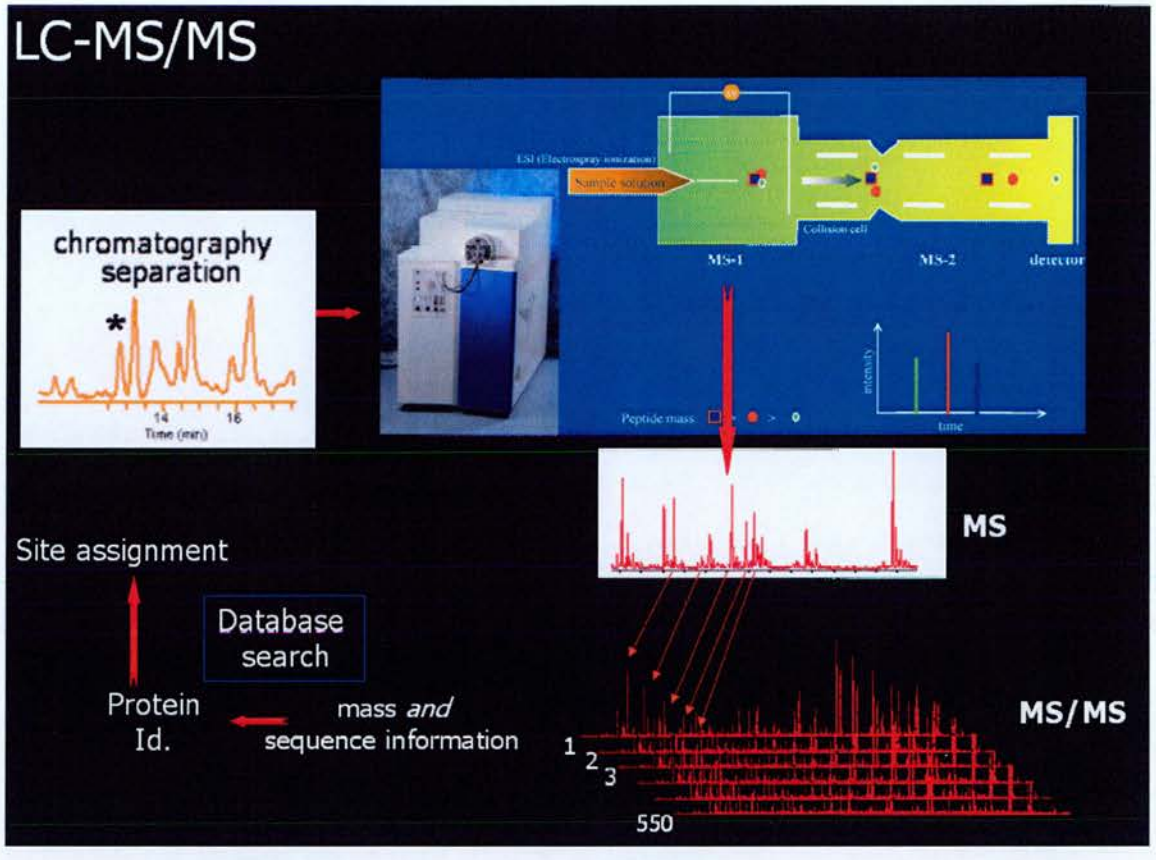


Figure 1.9 – Workflow of an LC-MS/MS experiment. A mixture of peptides from a protein sample digest is separated by reversed-phase chromatography on a nano-flow HPLC. The peptides elute from the RP column and are ionized by an electrospray source. In the first stage of mass spectrometry, m/z values and charge states for each precursor ion are determined and the most abundant precursor ions are selected for analysis in the second stage. The ions are then fragmented with by collision-induced dissociation (CID) with a gas to produce fragment ions which are detected. Using the mass (from MS-1) and sequence information (from MS-2) protein sequence databases are searched to provide peptide identifications and protein matches. Analysis of phosphopeptides is similar but specific scanning approaches are used in the mass spectrometer to detect the loss of phosphate during CID. Assignment of the exact position of phosphorylation is accomplished by manual interpretation of the mass spectra.

1.5.1.2 MS-based strategies for phosphorylation analysis

The use of mass spectrometry for phosphorylation site identification has increased greatly over the last decade superseding traditional approaches of 2D phosphopeptide mapping and Edman sequencing. In the period from 1992 to 2003 1281 phosphorylation sites were characterised by mass spectrometry in 203 publications (Loyet et al., 2005) (Figure 1.8). In the latter years of this period, advances in sample preparation have enabled hundreds of phosphorylation sites to be determined and reported in a single publication. In addition, the concomitant advances in technology of mass spectrometers have also facilitated the transition from single phosphoprotein analysis to analysis of phosphoproteomes.

A mass spectrometer consists of an ion source, a mass analyser that measures the mass to charge ratio (m/z) of peptides and a detector that counts the numbers of ions at each m/z value (Figure 1.9). ESI mass spectrometers use a heated capillary needle through which the peptide solution elutes and are subsequently ionised and introduced into the mass spectrometer. This type of ion source is readily compatible with liquid chromatographic separations resulting in an LC-MS/MS setup that was used in the analyses described in this thesis. In a typical LC-MS/MS experiment a peptide mixture is separated by analytical reversed-phase chromatography on a HPLC (High pressure liquid chromatography) and are introduced into the mass spectrometer via the ESI source. Multiply protonated peptides give rise to the first mass spectrum (MS1) and then the computer generates a prioritised list of the most intense precursor ions detected and are subsequently subjected to fragmentation by energetic collision with gas (CID (collision induced dissociation)) in the collision cell. This MS/MS spectrum contains data relating to the amino acid composition of the peptides and the pattern of fragmentation that occurred. This process of tandem mass spectrometry provides m/z values of the peptides and fragmentation data for each peptide. This data is then used by sequence database searching programs such

as MASCOT (Perkins et al., 1999) or SEQUEST (Ducret et al., 1998) and provides probability-based peptide matches and therefore protein identifications.

Identification of phosphopeptides and characterisation of phosphorylation sites by LC-MS/MS analysis is more difficult than standard protein identification experiments and can be achieved by special phospho-scanning functions in the mass spectrometer. Neutral loss scanning is a popular approach to detect phosphopeptides and it involves monitoring for the loss of ions characteristic to phosphorylated residues. When peptides containing phosphoserine or phosphothreonine are subjected to CID they usually undergo gas-phase β-elimination resulting in a neutral loss of phosphoric acid (-98 Da) or are dephosphorylated (-80 Da). The resultant mass spectrum indicates the position of the loss of phosphate by a spacing of 69 Da (dehydroalanine) or 83 Da (dehydroaminobutyric acid). This type of analysis requires prior knowledge of the charge state of the precursor ion before CID and thus is highly compatible with Q-Tof mass spectrometers operated in data-dependent acquisition mode. This approach is not as efficient for detecting phosphotyrosine as phosphotyrosine is more resistant to neutral loss. However, another scanning approach is suitable for detection of phosphotyrosine and is based on the production of reporter ions characteristic for the loss of phosphotyrosine. Immonium ions of phosphotyrosine have an m/z value of 216.043 and are generated by double cleavage of the peptide backbone. These can be detected with good sensitivity with detection possible from sub-picomole amounts of gel-separated proteins (Steen et al., 2001).

Additionally, when phospho-Ser/Thr/Tyr containing peptides are fragmented by collision-induced dissociation (CID) in the negative ion mode, a characteristic PO_3^- ion (m/z = -79) can be detected by precursor ion scanning. Once this ion is detected, the mass spectrometer can be set to automatically switch to positive ion mode to conduct MS/MS for sequencing.

1.5.2 Challenges

The main challenge for phosphoproteome analysis is enrichment of phosphorylated proteins and peptides to sufficient quantities so that they can be characterised by mass spectrometry. Preservation of phosphorylation during enrichment is essential, as any loss of phosphorylation due to phosphatase activity or by chemical dephosphorylation would significantly reduce the likelihood of detection of low abundance phosphorylation. Phosphoproteome enrichment methods should also be as straightforward as possible to minimise sample losses and to enhance reproducibility of the protocol. Another important challenge of proteomic or phosphoproteomic experiments is analysis of large volumes of data that can be produced. After identification of phosphorylation sites, comprehensive bioinformatic analysis is necessary to annotate phosphorylation sites. Extensive literature mining for functional information pertaining to phosphoproteins, phosphorylation sites and also to the sequence surrounding the phosphorylation sites should be performed. The use of prediction algorithms to annotate domains, amino acid sequence characteristics and potential kinases and binding sites is necessary. Without this kind of post-analysis annotation lists of phosphopeptides are of limited initial value. In a similar manner, generation of long lists of protein identifications in proteome profiling experiments presents the problem of data analysis. It is necessary to systematically annotate such lists with as much relevant data as possible to extract the value and biological significance of the data.

The strength of proteomics is that its can generate vast amounts of unbiased data and can allow the rapid characterisation of proteomes and phosphoproteomes, which would not be possible using conventional techniques.

1.6 Aims of the project

The synapse is a specialised cell-cell contact that contains presynaptic and postsynaptic membranes that can transduce electrical information into long-lasting molecular changes that constitute synaptic plasticity. It is now apparent that complex mechanisms of trafficking, signalling and reorganisation occur on both sides of the synaptic cleft and that these processes must be highly organised and regulated.

The aim of this project was to study the phosphorylation and the organisation of proteins at the synapse. To achieve this, novel and robust methodologies for phosphoproteome enrichment and analysis were developed and the application to synaptic and cytosolic phosphoproteomes is described in this thesis. In addition, the identification and integration of data from postsynaptic proteomes is described as well as analysis of this data to reveal trends of molecular organisation that may exist in the postsynaptic proteome.

CHAPTER 2

ROBUST ENRICHMENT OF PHOSPHORYLATED SPECIES IN COMPLEX MIXTURES BY SEQUENTIAL PROTEIN AND PEPTIDE METAL-AFFINITY CHROMATOGRAPHY AND ANALYSIS BY TANDEM MASS SPECTROMETRY

2. Robust enrichment of phosphorylated species in complex mixtures by sequential protein and peptide metal-affinity chromatography and analysis by tandem mass spectrometry

2.1 ABSTRACT

Reversible protein phosphorylation mediated by kinases, phosphatases and regulatory molecules, is an essential mechanism of signal transduction in living cells. Although phosphorylation is the most intensively studied of the several hundred known post-translational modifications on proteins, until recently the rate of identification of phosphorylation sites has remained low. The use of tandem mass spectrometry has greatly accelerated the identification of phosphorylation sites although this progress was limited by difficulties in phosphoresidue enrichment techniques. We have developed upon existing immobilised metal-affinity chromatography (IMAC) techniques for capturing phosphopeptides, to selectively purify phosphoproteins from complex mixtures. Combinations of phosphoprotein and phosphopeptide enrichment were more effective than current single phosphopeptide purification approaches. We have also implemented iterative MS-based scanning techniques to improve detection of phosphorylated peptides in these enriched samples. Here, we provide detailed instructions for implementing and validating these methods together with analysis by tandem mass spectrometry for the study of phosphorylation at the mammalian synapse. This strategy should be widely applicable to the characterisation of protein phosphorylation in diverse tissues, organelles and in cell culture.

2.2 INTRODUCTION

Protein phosphorylation is an essential regulator of protein function and signalling pathways in the cell. Phosphorylation is primarily used as a reversible tag of activity, by which information can be transduced from one protein to another in response to an appropriate stimulus. This process of signal transduction is fundamentally important to many cellular processes and as a consequence, is a highly complex and regulated system. It is now known that such signalling pathways are often mediated by multi-protein complexes which translate a stimulus (e.g. receptor stimulation) to downstream effectors. Multiprotein signalling complexes or "signalling machines" are responsible for orchestrating the complex signalling events at glutamate receptors (Husi et al., 2000), (Farr et al., 2004), ATP receptors (Kim et al., 2001), cytokine receptors (Bouwmeester et al., 2004), growth factor receptors (Blagoev et al., 2004) and G-protein coupled receptors (Becamel et al., 2002). They seem to share a common organisation of receptor, protein scaffold, (in which the signalling molecules are localised) and membrane to cytoskeletal interactions. The propagation of an appropriate response to receptor stimulation is highly regulated by phosphorylation cascades. This is exemplified by the intensively studied MAP kinase pathway, which ultimately leads to activation of transcription factors and modulation of gene expression. Phosphorylation is employed to modulate protein function and stability and to mediate phosphorylation dependent protein-protein interactions (e.g. SH2 binding of PI 3-kinase to NR2B) (Hisatsune et al., 1999) conferring a higher order level of regulation in such protein complexes.

Historically, synapse phosphorylation and its importance in regulating signal transduction and protein function has been studied at the level of single molecules (Greengard et al., 1993), (Sweatt and Kandel, 1989), but new proteomic strategies lend themselves to the global characterisation of the signalling properties of the proteome. Systematic characterisation of kinase and phosphatase substrates is

essential to extend current knowledge of known signalling pathways and will inevitably reveal many more. It is now apparent that the extent of phosphorylation in a proteome far exceeds original estimates, with the majority of proteins capable of being phosphorylated at multiple sites with a recent reports of 20-30 phosphoresidues in a single protein (Collins et al., 2005), (Zhang et al., 2004). A recent estimate of a possible 100,000 phosphorylation sites in the human proteome and a current knowledge of 3-4% of these, highlights the daunting task ahead (Zhang et al., 2002). The use of mass spectrometry to characterise phosphorylation has rapidly increased in recent years with a 5000% increase in phosphorylation sites defined by this approach from 1992 to 2002 (Loyet et al., 2005). Although MS-based strategies for phosphorylation site identification have clear advantages over traditional techniques (for review, see Mann et al. 2002) it is not yet routine to analyse complex phosphoprotein samples. Analysis of phosphoproteins is not straightforward for a number of reasons. First, the stoichiometry of phosphorylation is generally low. Second, the phosphorylated sites on proteins might vary, implying that any given phosphoprotein is heterogeneous. Third, many classes of proteins are present at low abundance within cells and, in these cases; enrichment is essential before analysis. Fourth, most analytical techniques used for studying protein phosphorylation have a limited dynamic range meaning that minor sites might be difficult to identify. Finally, endogenous phosphatases could dephosphorylate proteins/peptides unless their activity is inhibited during sample preparation.

Sample preparation is crucial for the success of any MS-based phosphoproteomic analysis and numerous approaches have been developed to overcome its associated problems. These include phosphoresidue enrichment strategies such as phosphospecific antibodies and IMAC (Immobilised metal affinity chromatography), and labelling strategies such as ^{32}P-ATP and chemical modification with subsequent biotin tagging (for review, see Mann et al. 2002). IMAC enrichment, which is based on the affinity of metals (usually Fe(III) or Ga(III) to negatively charged phosphate groups, has become the most common approach

for *in vivo* phosphorylation site identification (Loyet et al., 2005). This is largely due to improvements made in the specificity of IMAC brought about by chemical derivatisation of carboxylic acid groups to methyl esters, which decreases non-specific binding of acidic peptides (Ficarro et al., 2002). Recent phosphoproteomic studies have utilised peptide IMAC protocols together with methyl esterification to enrich for phosphopeptides prior to mass spectrometry. This has been used successfully to study phosphorylation in yeast (Ficarro et al., 2002), arabidopsis (Nuhse et al., 2003), and cell lines (Shu et al., 2004) however, the use of these IMAC approaches on complex mammalian sub-cellular organelles has not proved as successful (DeGiorgis et al., 2005), (Trinidad et al., 2005).

The MS strategy employed for phosphorylation analysis is mass spectrometer type-dependent and the most commonly used are ESI (electrospray ionisation) mass spectrometers (usually nanoelectrospray) in conjunction with 1 or 2D liquid chromatography separation (Loyet et al., 2005). Historically, triple quadrupole mass spectrometers were the first to be widely used for phosphorylation studies, particularly with the nano-electrospray source (Wilm and Mann, 1996). This instrument was particularly suited to linked scanning modes: neutral loss and precursor ion scanning were useful for detecting phosphopeptides in the presence of much larger amounts of unmodified peptide. The triple quadrupole for proteomics work was largely superseded by the more sensitive quadrupole time-of-flight (Q-Tof) geometry, which could not perform linked scans in hardware by the very nature of the geometry. Post-processing of the data could extract phosphopeptide information (and indeed information on many other types of post-translational modification), and the increased sensitivity of the Q-Tof made this trade-off acceptable. Recently, a new generation of instruments based on linear traps offer new scanning possibilities for phosphopeptides similar to those originally available from triple quadrupoles. It should be noted that no one mass spectrometer type will cover all cases. In particular, we point out that matrix assisted laser desorption ionisation (MALDI) still has a very useful role in studying

EX BIBL. UNIV. EDINBURGH

phosphopeptides, and with enrichment methods may be better-suited to the detection of multi-phosphorylated peptides (Larsen et al., 2005). The work described here was carried out on either a Q-ToF Ultima (Waters) or a 4000 Q-Trap (AB). Specific MS scanning modes such as neutral loss and precursor ion scanning have been developed to improve the detection of phosphorylation in highly complex peptide mixtures. Gas phase β elimination of phosphorylated serine and threonine residues produce a loss of phosphate that can be monitored in neutral loss mode and sequence information can be obtained in the same experiment. When phospho Ser/Thr/Tyr containing peptides are fragmented by collision-induced dissociation (CID) in the negative ion mode, a characteristic PO_3^- ion (m/z = -79) can be detected by precursor ion scanning. Once this ion is detected the mass spectrometer can be set to automatically switch to positive ion mode to conduct MS/MS for sequencing.

In this protocol we describe how we have improved traditional IMAC and LC-MS/MS based strategies to study the synapse phosphoproteome by implementing multistage proteome fractionation and phosphoproteome enrichment and iterative MS-based scanning approaches (Figure 2.1). This protocol should be useful for researchers pursuing such large-scale analyses of phosphorylation in complex samples.

2.3 MATERIALS

2.3.1 Chemicals

DTT

Acetic acid

Acetonitrile chromasolv (Fluka, 34851)

Acetyl chloride

Acrylamide

Ammonium bicarbonate

Ammoniumpersulfate

Aprotinin (Roche, 981532)

Bis-acrylamide

B-Mercaptoethanol

Bromophenol blue

CAPS

DMSO

ECL Plus Western Blotting detection reagents (Amersham Bioscience, RPN2132)

EDTA

Formic acid

Gallium chloride GaCL3 5g (Sigma-Aldrich, 25,419-3) or (Avocado Research Chemicals, 43879)

Glycerol

Glycine

H2O2

Iso-butanol

KCl

Leupeptin (Roche, 1034626)

Liquid nitrogen

Luminol

Na2HPO4

NaCl

p-coumaric acid

PMSF

Pro-Q Diamond Phosphoprotein gel stain kit (Molecular Probes, M-33306)

SDS

Sodium acetate

Sodium ortho-vanadate

Sucrose

TEMED

Thiourea

Tris-acetate

Tris-Base

Tween20

Urea-ultra grade (Sigma, U0631)

2.3.2 Enzymes/proteins

Trypsin Gold, Mass Spectrometry Grade (Promega) V5280

PeppermintStick phosphoprotein molecular weight standards (Molecular Probes, P-33350)

2.3.3 Plastics

38 ml polycarbonate centrifuge tubes (Nalgene, 306/0258/08)

1.5 ml Microcentrifuge tubes

2.3.4 Columns and Resins

Fast-flow chelating sepharose with iminodiacetic acid (IDA) (Amersham Biosciences, 17-0575-01)

Self Pack POROS 20 MC Media (Applied Biosystems, 1-5428-02)

Bio-Spin Chromatography Columns (BioRad, 732-6025)

Trap column:

PepMap C18 3 μm 100Å, 300 μm i.d. x 5mm (LC Packings)

BetaMax Neutral 5μm, 180 μm i.d. x 30mm (Thermo Hypersil-Keystone, 95005-030215)

Analytical column:

PepMap C18 3 μm 100Å, 75 μm i.d. x 15 cm (LC Packings, 160321)

Note: other type columns at similar performance may also be used.

C 10/10 chromatography column (Amersham Bioscience, 19-5001-01)

RESOURCE 15RPC 1ml Reversed Phase chromatography column (Amersham Bioscience, 17-1181-01)

Vivaspin 20 PES membrane 5000 MWCO (Vivascience, VS2011)

2.3.5 Equipment

Dounce Homogeniser (VWR, 406/0315/10)

Peristaltic Pump (Amersham Bioscience, 18-1110-91)

ÄKTA FPLC (Amersham Bioscience, 18-1900-256)

Mini-PROTEAN II Multiscreen apparatus (BioRad, 170-4017)

X-ray film developer

Microcentrifuge

Liquid nitrogen dewar flask

Roller shaker

37°C incubator

Vortexer

Sorvall Discovery™ 100SE ultracentrifuge with swinging bucket rotor AH-629(36ml) or equivalent

Savant SpeedVac (Thermo)

Typhoon™ Scanner (Amersham Bioscience, 9410-PC)

Nano flow HPLC system with an autosampler and a valve switch unit, such as

- Ultimate™ gradient pump equipped with FAMOS autosampler and Switchos II (LC Packings), or
- CapLC with nano stream selected module (Waters)

Mass spectrometer: Q-Tof Ultima API with nanospray source (Micromass)

Note: Other type Q-Tof API or tandem mass spectrometers which give similar performance can also be used.

4000 Q TRAP (Applied Biosystems).

Other

Hyperfilm™ ECL™ (Amersham Bioscience, RPN2103K)

2.4 RECIPES

Recipe 1: Synaptosome preparation buffer A

50 mM Tris-acetate pH 7.4 (0.5 M stock)

10% (w/w) sucrose

5 mM EDTA (0.5 M stock)

1 mM PMSF

2 µg/ml Aprotinin (2 mg/ml stock)

2 µg/ml Leupeptin (2 mg/ml stock)

1 mM sodium ortho-vanadate

Add H_2O to 50 ml

Combine components in a 50 ml tube, rotate to dissolve sucrose and chill at 4°C

Recipe 2: Synaptosome preparation buffer B

5 mM Tris-acetate pH 8.1 (0.5 M stock)

1 mM PMSF

2 µg/ml Aprotinin (2 mg/ml stock)

2 µg/ml Leupeptin (2 mg/ml stock)

1 mM sodium ortho-vanadate

Add H_2O to 50 ml

Combine components in a 50 ml tube and chill at 4°C

Recipe 3: Synaptosome preparation buffer C

50 mM Tris-acetate pH 7.4 (0.5 M stock)

28.5% (w/w) sucrose

Combine components in a 50 ml tube, rotate to dissolve sucrose and chill at 4°C

Recipe 4: Synaptosome preparation buffer D

50 mM Tris-acetate pH 7.4 (0.5 M stock)

10% (w/w) sucrose

Combine components in a 50 ml tube, rotate to dissolve sucrose and chill at 4°C

Recipe 5: 100 mM $GaCl_3$ Solution

Carefully break off the glass seal of the vial of $GaCl_3$ powder (5 g) and immerse in 284 ml of distilled H_2O in a glass beaker. Once dissolved store in at 4°C protected from light.

Note: Caution must be exercised when dissolving $GaCl_3$ in water, as it is an exothermic reaction, which sputters and evolves HCl gas; wear protective clothing and perform in a fume hood.

Recipe 6: Protein IMAC wash buffer

6 M Urea

50 mM Tris-acetate pH 7.4

Make up to 1 L and pass through a 0.2 micron filter

Recipe 7: Protein IMAC elution buffer

6 M Urea

50 mM Tris-acetate pH 7.4

100 mM EDTA

100 mM EGTA

Make up to 1 L and pass through a 0.2 micron filter

Recipe 8: SDS-PAGE (Laemmli) 5 X reducing sample buffer

4 ml H_2O

1 ml 0.5 M Tris pH 6.8

1 ml glycerol

1.6 ml 10% SDS

0.4 ml β-ME or 20 mM DTT (0.32 ml 0.5 M)

0.2 ml 0.1% bromophenol blue

Recipe 9: SDS-PAGE (Laemmli) gel solutions

Acrylamide/Bis	146 g acrylamide 4 g Bis-acrylamide H_2O to 500 ml, store at 4°C or at RT
1.5 M Tris pH 8.8	90.75 g Tris-Base H_2O to 400 ml, set pH to 8.8, H_2O to 500 ml, store at RT
0.5 M Tris pH 6.8	30 g Tris-Base H_2O to 450 ml, set pH to 6.8, H_2O to 500 ml, store at RT
10% SDS (w/v)	10 g SDS (Sodium salt)/ 100 ml H_2O, store at RT
10% APS (w/v)	0.1 g ammoniumpersulfate / 1 ml H_2O Store at RT, do not keep for more than 1 week
TEMED	TEMED, store at 4°C
Iso-butanol	Water saturated iso-butanol

Recipe 10: SDS-PAGE (Laemmli) running buffer

10 x Running buffer	30 g Tris-Base 144 g glycine 10 g SDS

H_2O to 1 l, check pH (should be self-adjusting) and set to 8.3

Recipe 11: Pro-Q Diamond Phosphoprotein gel fixing solution

Prepare a solution of 50% methanol and 10% acetic acid.

Recipe 12: Pro-Q Diamond Phosphoprotein gel destain solution

1 L of destain solution:

50 mL of 1 M sodium acetate, pH 4.0

750 mL of ultrapure water

200 mL of acetonitrile

Combine and mix thoroughly.

Recipe 13: Wash solution for SYPRO Ruby gel staining

Prepare a 100 ml solution of 10% methanol and 7% acetic acid.

Recipe 14: Transfer solution

10 mM CAPS pH11.0 at 4°C

10% (v/v) MeOH

Recipe 15: 20 x PBS

160 g NaCl

4 g KCl

12.2 g Na_2HPO_4 (anhydrous)

4 g KH_2PO_4

H_2O to 1 L

Store at 4°C or RT (might precipitate at lower temperature)

PBS/Tween PBS + 0.1% Tween20

Recipe 16: Chemiluminescence solution A

Stocks

250 mM Luminol (44.3 g/l) in DMSO, keep at -20°C

90 mM p-coumaric acid (14.8 g/l) in DMSO, keep at -20°C

Solution A

5 ml Luminol

5 ml p-coumaric acid

12.5 ml 1.5 M Tris pH 8.8

12.5 ml 0.5 M Tris pH 6.8

215 ml H_2O

Store at 4°C in the dark

Recipe 17: Chemiluminescence solution B

250 µl 30% H_2O_2

250 ml H_2O

Store at 4°C in the dark

Mix solutions A and B 1:1 just prior to blot development

Recipe 18: Tryptic digestion buffer

Dilute the sample to

1 M urea

0.125 M thiourea

25 mM NH_4HCO_3 (pH 8)

Trypsin was reconstituted with 50 mM acetic acid and was added to this solution

Recipe 19: Mobile phase A for reversed phase desalting of phosphoprotein digest

94.9% H_2O

5% acetonitrile

0.1% TFA

Recipe 20: Mobile phase A for reversed phase desalting of phosphoprotein digest

94.9% acetonitrile

5% H_2O

0.1% TFA

Recipe 21: 2 M Methanolic HCL

Add 160 µl of acetyl chloride with stirring to 1 ml of methanol slowly in a dropwise fashion.

Note: Caution must be exercised during this addition as this reaction generates heat. Perform operation in fume hood using safety glasses and gloves

Recipe 22: High organic peptide solubilisation/loading buffer

1:1:1 of acetonitrile, methanol and H_2O, pH 2.5-3

Recipe 23: Aqueous peptide solubilisation/loading buffer

0.3 % acetic acid, pH 2.5-3

Recipe 24: Mobile phase A for LC-MS/MS of tryptic digest of phosphoproteins

94.9% H_2O

5% CH_3CN

0.1% formic acid

Recipe 25: Mobile phase B for LC-MS/MS of tryptic digest of phosphoproteins

95% CH_3CN

4.9% H_2O

0.1% formic acid

Recipe 26: Mobile phase C for LC-MS/MS of tryptic digest of phosphoproteins

99.9% H_2O

0.1% formic acid

Recipe 27: Mobile phase A for LC-MS/MS of IMAC enriched phosphopeptides

99.9% H_2O

0.1% formic acid

Recipe28: Mobile phase B for LC-MS/MS of IMAC enriched phosphopeptides

70% CH_3CN

29.9% H_2O

0.1% formic acid

Recipe 29: Mobile phase B for LC-MS/MS of IMAC enriched phosphopeptides

99.5% H_2O

0.5% formic acid

Both CH_3CN and H_2O are at HPLC gradient grade in recipes 19 - 24.

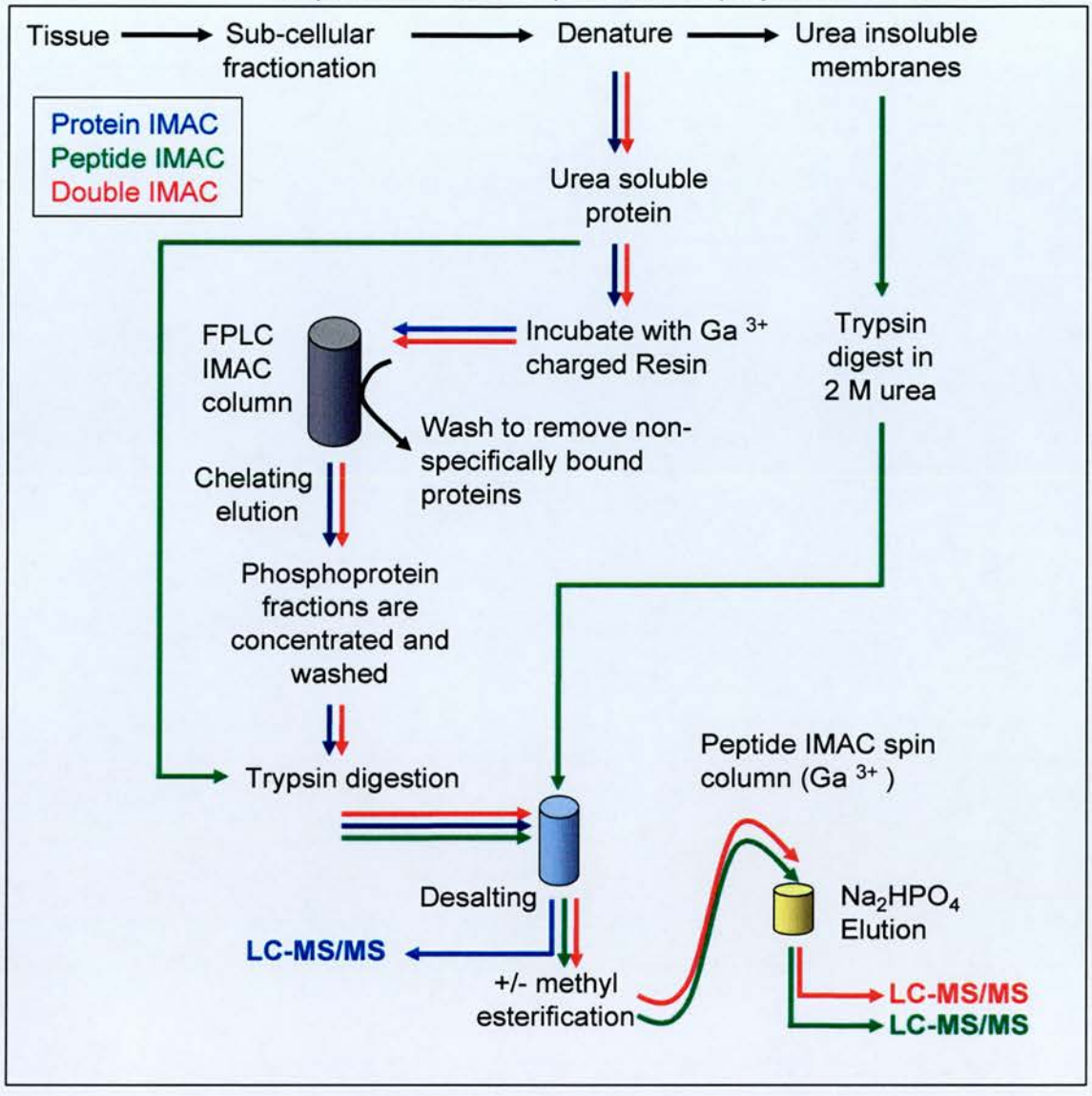


Figure 2.1 - Schematic overview of methodology.
The target proteome, in this case a sub-cellular fraction, is phase-partitioned by its differential solubility in 6 M urea. The soluble fraction can be processed in two ways. Traditionally this fraction is digested in solution, desalted and applied to a peptide IMAC column (green arrows) with or without methyl-esterification and analysed by LC-MS/MS. In addition we have developed top-down enrichment step where the soluble proteins are applied to a protein IMAC column (blue arrows) and the eluted phosphoproteins are digested and also analysed by LC-MS/MS. A novel combination of these two approaches (red arrow) in which a two-step enrichment of phosphoproteins and phosphopeptides and analysis by LC-MS/MS results in improved identification of phosphorylation sites. The urea-insoluble fraction which represents tightly bound or integral membrane proteins can be analysed by direct digestion in 2 M urea to release peptides from cytoplasmic domains of these proteins, which can be applied to a peptide IMAC column to isolate phosphopeptides. These combined strategies when applied to the synapse phosphoproteome yielded identification of over 650 phosphorylation events corresponding to 331 phosphorylation sites. This combined strategy of phase-partitioning and sequential enrichment of phosphoresidues allows comprehensive analysis of a complex sample.

2.5 INSTRUCTIONS

An overview of the purification strategies is shown in Figure 2.1. In this section the application of our protocols to the synapse phosphoproteome will be described but these should be readably applicable to other tissues, cells and organelles.

2.5.1 Purification of Synaptosomes by Sub-cellular Fractionation

This section briefly describes purification synaptosomes from mouse forebrains, this protocol has been well documented (Carlin et al., 1980) but the original protocol has a number of variations.

1. Mouse forebrains should be rapidly dissected, frozen immediately in liquid nitrogen and stored at -80°C.
2. Cool all buffers and the Homogeniser to 4°C.
3. Collect 16 forebrains and add to homogenisation buffer (Recipe 1) (9 ml/g of buffer) in a 50 ml Dounce Homogeniser.
4. Homogenise with 15 strokes.
5. Collect the extract and centrifuge at 800 x g for 20 min (all centrifugation steps performed at 4°C).
6. Centrifuge the supernatant at 16,000 x g for 30 min.
7. Re-suspend the pellet in Recipe 2 (5 ml/g of buffer).
8. Homogenise with 3 strokes.
9. Incubate the extract in the Homogeniser at 4°C for 45 min.
10. Homogenise with 10 strokes.
11. Add sucrose to 34% (w/w).
12. Prepare sucrose gradients by addition of 10 ml of Recipe 4 (10% sucrose (w/w) to 38 ml polycarbonate centrifuge tubes, followed by careful introduction of 10 ml of recipe 3 (28.5% sucrose (w/w) from the bottom of the tube.

13. Add 10 ml of extract (in 34% sucrose (w/w)) to the bottom of the tube, such that the gradients are 10, 28.5 and 34% sucrose (w/w) from the top to the bottom.
14. Centrifuge at 60,000 x g for 120 mins.
15. Collect the protein-containing band that forms between the 34% and 28.5% sucrose gradients and dilute with 50 mM Tris-acetate pH 7.4 to 10% sucrose.
16. Centrifuge at 48,000 x g for 30 min, remove supernatant and re-suspend this synaptosome pellet in 50 mM Tris-acetate pH 7.4.
17. Store at -80°C.

2.5.2 Protein IMAC Sample Preparation

In our protein IMAC protocol, a cell lysate or sub-cellular fraction is solubilised in 6 M urea and incubated with a metal-charged sepharose-based resin. Specific elution is achieved by a stopped-flow elution with chelating buffer and eluted phosphoproteins can be concentrated and washed. This method when performed with an FPLC system permits large scale purification which is necessary for subsequent phosphopeptide analysis by mass spectrometry (Figure 2.2). Batch purification in microcentrifuge tubes can also be formed which can be useful for rapid purification of smaller quantities of phosphoprotein for gel or western blot analysis.

1. Strip the Fast-flow chelating sepharose (IDA) of any bound metal by incubation with 100 mM EDTA and 100 mM EGTA for 1 hour with gentle agitation.
2. Clamp a C 10/10 chromatography column to a retort stand and connect to a Peristaltic Pump.
3. Add 5 ml (bed volume) of chelating resin to the column.
4. Wash the resin with 10 column volumes of H_2O using the pump at a flow rate of 1 ml/min.

5. Remove the resin and add to 50 ml of Recipe 5 (100 mM $GaCl_3$ Solution) and incubate at room temperature on a roller-mixer over night and protected from sunlight.
6. Wash the charged resin with 10 resin volumes of H_2O using the pump at a flow rate of 1 ml/min.
7. Store the charged resin (50% H_2O) at 4°C, protected from sunlight and use within 1 week.
8. Solubilise the protein sample (Synaptosomal proteins) (12.5 mg/10 ml) in 6 M urea for 20 min with mixing, centrifuge at 13,000 rpm for 5 min and retain the supernatant.
9. Equilibrate 1 ml of gallium charged resin with Recipe 6 (protein IMAC wash buffer) by a 5 resin volume in-column wash.
10. Incubate the urea-soluble supernatant (~ 9.5 ml) with 1 ml of resin for 1 hour at room temperature on a roller shaker and protect from light.
11. Equilibrate an FPLC (ÄKTA) system with elution buffer (Recipe 7) and then wash buffer (Recipe 6) at room temperature.
12. Pack the resin/protein mixture into a C 10/10 column and attach to the FPLC, avoiding introduction of air bubbles into the system.
13. Wash unbound protein from the column with 20 column volumes of wash buffer (100% A) at a flow rate of 1 ml/min until the A280 > 0.25 mAU.
14. Switch the flow to 100% buffer B (recipe 7 (elution buffer)) and run at 0.5 ml/min and start to collect 1 ml fractions.
15. Once the A280 reaches ~ 10 mAU, change to 0% B for 10 min to allow a stopped flow elution which results in a discrete elution peak, see Figure 2.2.
16. Pool 4-5 fractions surrounding the phosphoprotein peak and wash with buffer B (4 times dilution of the sample) in a Vivaspin 6 PES membrane spin column and concentrate the sample to 200 µl.

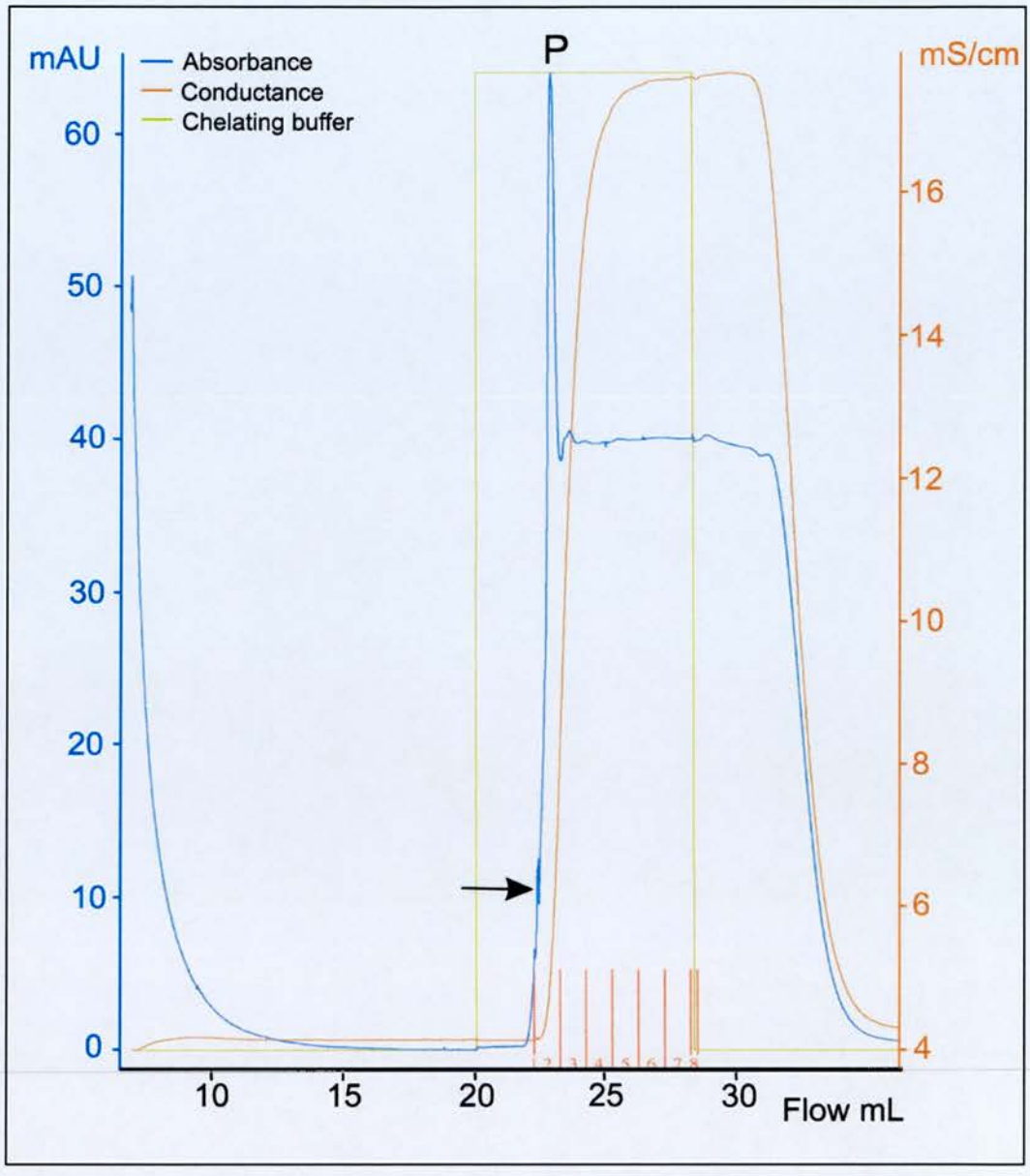


Figure 2.2 - FPLC chromatogram of a protein IMAC purification
Un-bound protein is washed off the IMAC column with 20 column volumes of wash buffer (100% A) at a flow rate of 1 ml/min until the A280 > 0.25 mAU. Then the flow is switched to 100% buffer B (elution buffer) and run at 0.5 ml/min. A stopped-flow elution with a chelating buffer is performed to elute the phosphoproteins in a discrete peak (P).

2.5.3 Validation of Phosphoprotein Purification

This section describes validation of the protein IMAC enrichment which is advisable before further purification or analysis by mass spectrometry. Validation can be achieved by phosphoprotein staining of protein IMAC purified phosphoproteins on an SDS-PAGE gel (Figure 2.3). Alternatively western blotting with phosphospecific antibodies or antibodies to known phosphoproteins should show clear enrichment in the protein IMAC purified sample (Figure 3.3). As elements of these two techniques are routine they will be described briefly.

2.5.3.1 Electrophoresis and Phosphoprotein Staining of SDS-PAGE gels

1. Boil 10% of the eluted phosphoprotein and molecular weight marker in tightly closed micro-centrifuge tubes in 5x Sample (Laemmli) Buffer (Recipe 8).
2. Separate on an SDS-PAGE (Recipe 9, 10) mini-gel for 50 min at 200 V.
3. All incubation steps are performed at room temperature with gentle agitation.
4. Fix the gel in 100 ml of fixing solution (Recipe 11) and incubate for 30 min and repeat this step again to remove all SDS.
5. Wash the gel three times in 100 ml H_2O for 10 min each.
6. Stain the gel with 50 ml of Pro-Q Diamond phosphoprotein gel stain in the dark for 90 min.
7. Destain the gel three times with 100 ml of Destain solution (Recipe 12) for 30 min.
8. Wash the gel twice with H_2O for 5 min each.
9. Scan the gel with a Typhoon scanner using excitation maximum at ~555 nm and an emission maximum at ~580 nm.

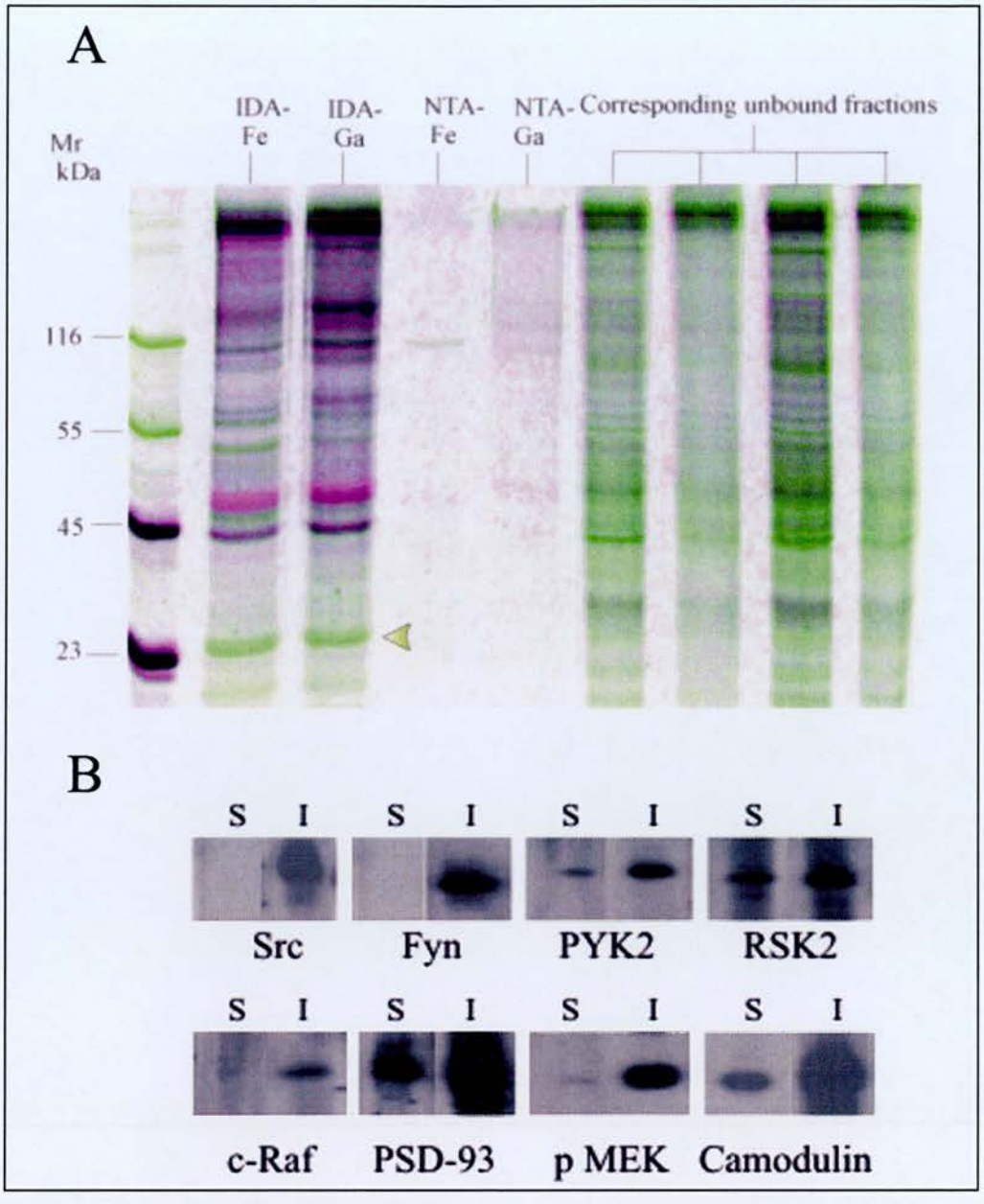


Figure 2.3 - Validation of IMAC phosphoprotein isolation strategy
A. Overlay of Pro-Q diamond phosphoprotein/ Ruby total protein stains of an SDS-PAGE gel for different resin/metal IMAC combinations used to selectively isolate phosphoproteins from urea-soluble synaptosomal extract. The corresponding unbound fractions are included for comparison. 20% of eluted protein from each purification was separated on a 12% SDS-PAGE gel and stained and imaged sequentially. Phosphorylated protein appears pink to black, whereas unphosphorylated protein appears green. Lane 1 shows Peppermint Stick molecular weight markers in which proteins of molecular weight 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 resins due to ion-exchange effects, and therefore removed from the dataset.**B.** Western blots of equal amounts of IMAC enriched (I) versus synaptosomal (S) protein for 8 synaptic phosphoproteins. A significant enrichment for all phosphoproteins was observed in the protein IMAC sample, and a specific enrichment for the phosphorylated form of MEK was seen using a phospho-specific antibody.

10. Rinse the gel twice with H_2O for 5 min each.
11. Incubate the gel in 50 ml SYPRO Ruby gel stain solution overnight.
12. Wash in Recipe 13 for 30 min and then twice with H_2O for 5 min each.
13. Scan the gel with a Typhoon scanner using excitation maximum at ~280/450 nm and an emission maximum at ~610 nm.
14. Overlay the phosphoprotein and total protein images of the gel using Totallab software (Nonlinear Dynamics).

2.5.3.2 Western Blotting

Western blotting of protein IMAC enriched and whole synaptosome samples for a selected set of phosphoproteins was carried out using a Mini-PROTEAN II Multiscreen apparatus (Figure 3.3). This apparatus allows screening of up to 40 different antibodies on a single membrane and requires much less of each individual antibody compared to traditional western blotting. This offers a high-throughput approach for validating a set of proteins but a conventional western blotting apparatus will suffice in its absence.

1. Adapt an SDS-PAGE gel-casting comb to cast a single well across the gel.
2. Separate the samples on a 12% SDS-PAGE mini-gel and transfer to PVDF membrane at 4°C for 90 min at 75 V in transfer solution (Recipe 14).
3. Block the membrane with 1% BSA in PBS/Tween (Recipe 15) overnight.
4. Wash the membrane with 3 x 20 ml of 1 x PBS/Tween for 5 min each.
5. Place the membrane onto the base of the Multiscreen apparatus, tighten screws to clamp.
6. Use optimal dilutions of primary antibodies in PBS/Tween, 550 µl per well and incubate for 2 hours with gentle agitation at room temperature.
7. Wash each well with 3x 0.5 ml of 1 x PBS/Tween for 5 min each.
8. Add 0.6ml of 1/5000 dilution of peroxidase-linked secondary antibody (in 1% BSA/PBS/Tween and incubate for 2 hours at room temperature.

9. Wash each well with 3x 0.5 ml of 1 x PBS/Tween for 5, 10 and 15 min each.
10. Signals were detected using enhanced chemiluminescence using Recipes 16 and 17, or alternatively with ECL Plus Western Blotting detection reagents (Amersham Bioscience).

2.5.4 In Solution Protein Digestion

1. Dilute the sample to 1 M urea, 0.125 M thiourea, 25 mM NH_4HCO_3 (pH 8) to a final volume of 400 µl (Recipe 18).
2. Add Trypsin Gold, mass spectrometry grade (Promega) in a ratio of 1:20 in 50 mM acetic acid.
3. Incubate at 37°C for 3 hours.

Note: We recommend Trypsin Gold, mass spectrometry grade as it has high activity and good specificity. Trypsin, sequence grade (Roche) can also be used but an overnight incubation is necessary for complete proteolysis. We have also used Sequencing grade modified porcine trypsin (Promega) but in our experience it produces semi-tryptic ends but this can be accommodated by specifying "semi-trypsin" in Mascot's search parameters.

4. Terminate the reaction by freezing in liquid nitrogen and lyophilise.

Note: 6 M Urea insoluble membranes can be digested by increasing the urea concentration in the digestion buffer to 2 M. This will result in reduced activity of trypsin but will allow better solubility and thus digestion of these proteins.

2.5.5 Reversed Phase Desalting of Protein Digests

Protein samples, whole lysate or protein IMAC enriched, for further processing (peptide IMAC) are desalted on a 1 ml bed volume RESOURCE 15RPC Reversed Phase chromatography column. Peptides for analysis by mass spectrometry are desalted on-line as described in the LC-MS/MS section to follow.

1. Wash and equilibrate the column with Buffer A (99.9% H_2O/0.1% FA, Recipe 19) and Buffer B (95% Acetonitrile/4.9 % H_2O/0.1% FA, Recipe 20).
2. Load the trypic digest into the injection loop and introduce to the column at 0.5 ml/min.
3. Wash with Buffer A until baseline is reached, ~ 10 column volumes.
4. Elute desalted peptides with a 0-95% gradient of Buffer B over 5 min.
5. Collect fractions, pool and lyophilise.

2.5.6 Peptide IMAC Sample Preparation

Peptide IMAC was carried out as described (Ficarro et al., 2002) with some modifications. There are a number of conditions in this protocol which can affect the yield and population of phosphopeptides obtained. Firstly, the choice of metal used to perform peptide IMAC can affect specificity and recovery. The most commonly used metals are Fe^{3+} and Ga^{3+} (Haydon et al., 2003), (Ficarro et al., 2002), (Posewitz and Tempst, 1999) and display different efficiencies of capture and release of phosphopeptides (Haydon et al., 2003). However, we found that Ga^{3+} performed the best in protein IMAC experiments (Collins et al., 2005), so we recommend to also use Ga^{3+} in the peptide IMAC step of a sequential phosphoprotein and phosphopeptide purification to minimise losses due to differences in specificity. It has been stated many times that acidic peptides affect the specificity of peptide IMAC purifications and methyl-esterification of peptide side chains has been reported to improve the specificity (Ficarro et al., 2002). In agreement with this we have observed a 30% reduction in the number of acidic amino acids in phosphopeptides which have been detected in peptide IMAC of methyl-esterified samples compared to unmodified samples. However, other studies have found that contaminating peptides do not have any different physiochemical properties than phosphopeptides in the sample but were actually from highly abundant proteins in the sample (Haydon et al., 2003).

In addition, peptide chemical modification reactions do not go to completion and we have observed that methyl-esterification in particular never occurs on 100% of peptides. This have the effect of diluting any particular sample into populations of modified and unmodified peptides which effectively reduces the amount of any given peptide and increases the sample complexity overall. It is recommended to perform peptide IMAC on methyl-esterified and unmodified versions of a sample to increase overall phosphopeptide coverage.

Next, the buffer in which peptides are solubilised and applied to the column affects the population of phosphopeptides observed. In sequential protein and peptide IMAC experiments we used a high organic (equal volumes of acetonitrile, methanol and water pH 2.5-3) and an aqueous (0.3 % acetic acid, pH 2.5-3) buffer to Solubilise and load peptides to the peptide IMAC column. Phosphopeptides found in both experiments represented a small fraction of the total phosphopeptides detected, with 37 and 60% of the double IMAC dataset found in the high organic and aqueous experiments respectively. The reason for this differential appearance based of the solvent used can be explained by looking at the GRAVY (Kyte and Doolittle, 1982) (Grand Average of Hydropathy) values of these peptides (Figure 2.4). Phosphopeptides found in the aqueous solubilised samples are overall less hydrophobic than those found in the high organic solvent solubilised samples. This would indicate that in previous studies such as that of the yeast phosphoproteome (Ficarro et al., 2002), in which only a high organic loading buffer was used, a significant sub-population of phosphopeptides may have been lost at this step. It is clear that in any complex sample there are populations of peptides which favour different solubilisation conditions depending on their hydrophobicity, so it is recommended to use different conditions to achieve greater coverage of the peptide population.

1. For methyl-esterification, desalt and lyophilise peptide digest and reconstitute with 0.5 mg/ml 2 M Methanolic HCL (Recipe 21).

2. Incubate with gentle agitation for 2 hours at room temperature.
3. Quench with 2 volumes of H_2O and lyophilise.
4. Take 0.5 ml (bed volume) of POROS MC and incubate with 100 mM EDTA and 100 mM EGTA on a roller mixer overnight to remove any trace metals from the resin.
5. Transfer the resin to a C 10/10 column and wash with 10 volumes of H_2O.
6. Incubate with 5 ml of Recipe 5 (100 mM $GaCl_3$ Solution) at room temperature on a roller mixer over night and protect from sunlight.
7. Wash the charged resin with 10 resin volumes of H_2O. Store the charged resin (50% H_2O) at 4°C, protect from sunlight and use within 1 week.
8. Reconstitute the peptide sample (methyl-esterified/unmodified) in either a high organic loading (equal volumes of acetonitrile, methanol and water pH 2.5-3, Recipe 22) or an aqueous loading buffer (0.3 % acetic acid, pH 2.5-3, Recipe 23).

   Note: For phosphoprotein digests, 3 mg of digested phosphoprotein was applied in 1 ml of loading buffer to a 100 µl bed volume of gallium charged POROS MC. A higher concentration of whole cell lysate may be applied to a similar resin volume as the percentage phosphopeptide content will be lower, for example 6 mg of digested whole synaptosomes was also applied in 1 ml of loading buffer to a 100 µl bed volume of gallium charged POROS MC.

9. Centrifuge the peptide solution for 10 min at 13, 000 rpm to remove any un-solubilised material.
10. Incubate the solubilised peptides with 200 µl of pre-equilibrated gallium charged POROS MC slurry (100 µl bed volume) at room temperature for 1 hour with gentle agitation.
11. Load the peptide/resin slurry into a pre-equilibrated spin column and centrifuge briefly to elute the unbound peptides.
12. Wash by addition and removal (by centrifugation) with 10 resin volumes of the loading buffer.

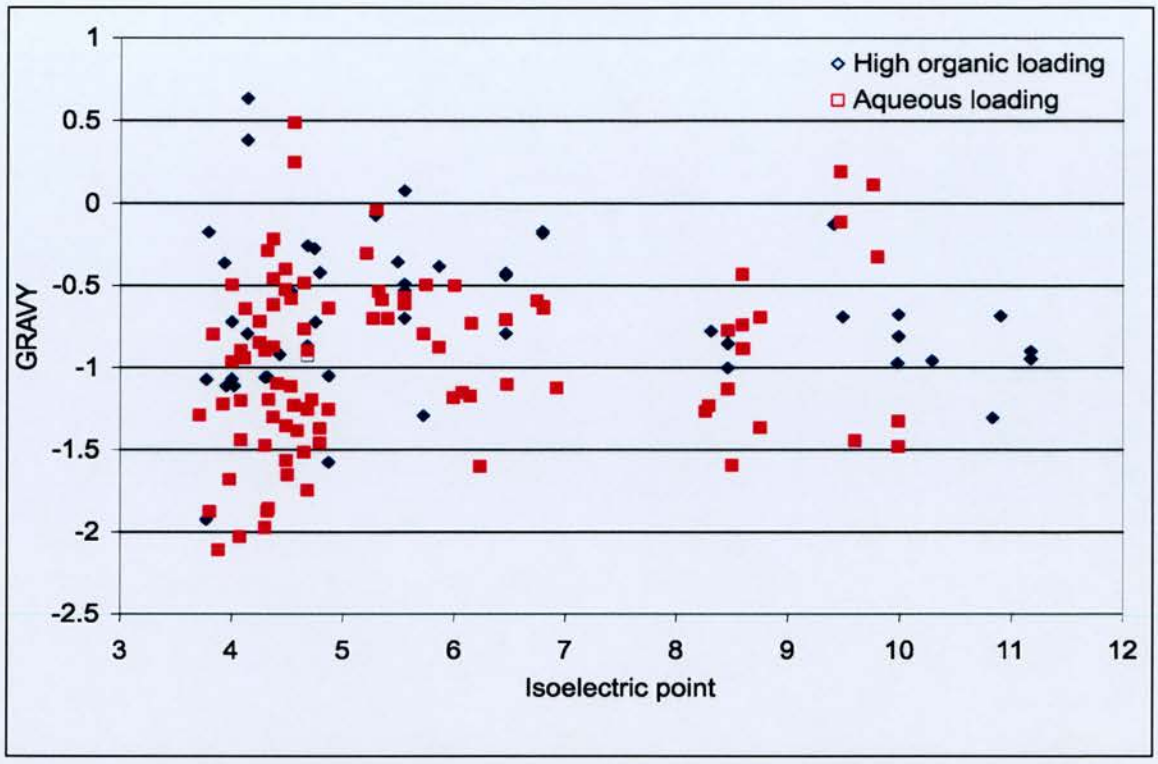


Figure 2.4 - Comparison of peptide IMAC solubilisation/loading buffers
GRAVY (Grand Average of Hydrophobicity) versus isoelectric point for phosphopeptides found in 2 peptide IMAC experiments are plotted. In one double IMAC experiment peptides were solubilized and loaded onto the peptide IMAC column in a high organic buffer (1:1:1 of acetonitrile/methanol/H_2O) and in the other peptides were solubilised and loaded in an aqueous buffer (0.3% acetic acid). These two buffers enable identification of two populations of phosphopeptides differing in their hydrophobicity.

13. Add 100 µl of 200 mM Na_2HPO4 and incubate for 5 min, spin and collect eluted phosphopeptides, repeat this elution again.

Note: The concentration of Na_2HPO4 used to elute phosphopeptides from the peptide IMAC column may be reduced to 50 mM, if volume reduction by partial lyophilisation is necessary.

14. Rapidly freeze at -80°C.

2.5.7 Double IMAC Sample Preparation

This protocol is a combination of protein and peptide IMAC steps described above and provides the obvious advantage of two rounds of enrichment for phosphorylated molecules. The application of a single protein IMAC strategy to the synaptic phosphoproteome resulted in the detection of 105 phosphorylation events but selective enrichment of the phosphopeptides derived from these purified phosphoproteins improved the detection of phosphorylation three fold, with the detection of 308 phosphorylation events.

1. Digest 3 mg of protein IMAC purified phosphoprotein with 1:20 ratio of Trypsin Gold, mass spectrometry grade for 3 hours in Recipe 18.
2. Desalt the digest as described in the Reversed Phase Desalting of Protein Digests section.
3. Perform methyl-esterification if required (see Peptide IMAC Section).
4. Reconstitute in loading buffer (high organic or aqueous).
5. Incubate 1 ml of this peptide mixture with 200 µl of pre-equilibrated POROS-Ga slurry (100 µl bed volume) for 1 hour at RT.
6. Load the resin into a spin-column and wash with 10 volumes of loading buffer.
7. Elute phosphopeptides with 2 x 100 µl of 200 mM Na_2HPO_4.

2.5.8 LC-MS/MS Analysis

For LC-MS/MS analysis, a nano flow HPLC system with valve switch unit is needed, so that the trap column can be placed for loading and desalting sample efficiently. This is particularly true for desalting phosphopeptide samples in high salt (0.2 M Na_2HPO4) with a large volume size (100-200 µl). Using a flow at 10 µL/min, we found it took up to 1 hour to desalt the sample. In our lab, we used Ultimate™ nano flow gradient pump equipped with FAMOS autosampler and Switchos II (LC Packings), or CapLC with nano stream selected module (Waters) coupled to Q-Tof hybrid tandem mass spectrometer. We found it is very helpful to have good mass accuracy in MS/MS spectra in assigning the phosphorylation sites.

Before sample analysis on LC-MS/MS, a systematic check is always performed in our lab and we would like to recommend the following:

1. Optimise the nanospray position and instrument parameters with 500 fmol/µL (for Q-Tof) or 100 fmol/µL (for Q-Tof Ultima) of [Glu1]^1-fibrinopeptide B in 25% CH_3CN, 74.9% H_2O, 0.1% formic acid.
2. Calibrate the instrument with the [Glu1]^1-fibrinopeptide B solution in step 1.
3. Systematically perform LC-MS/MS system checks with a tryptic digest of BSA, such as chromatographic resolution, mass accuracy and sequence coverage of BSA. For Q-Tof, use either 100 or 200 fmol BSA; for Q-Tof Ultima, 100 fmol.

   Note: If the analytical sample is at a low quantity, then the system should be tested using BSA at 10 fmol or lower.

2.5.9 LC-MS/MS analyses of a tryptic digest of IMAC purified phosphoproteins on a Q-Tof Ultima.

1. Loaded the protein digest via the autosampler to the trap column on the valve switch unit then and desalt with 0.1% formic acid (Recipe 26) at a flow rate of 10-20 µl/minute.

Note: When using a PepMap C18 3 µm 100Å, 300 µm i.d. x 5mm (LC Packings) trap column, the flow rate can be at 20 µl/minute. If the trap is BetaMax Neutral 5µm, 180 µm i.d. x 30mm (Thermo Hypersil-Keystone), the flow rate should be set at 10 µl/minute.

2. Switch valve after 3 minutes to connect the trap column on-line with analytical column.
3. Run a gradient as show in Table 2.1 to reverse elute peptides from the trap and separate on the analytical column using mobile phases (Recipes 24 and 25). The flow rate through the columns should be 250 – 300 nl/min.

MS acquisition starts at t = 3.0 min when desalting is finished. The instrument is operated in automated Data Dependent Acquisition (DDA) acquisition mode using MassLynx. In DDA mode, the instrument continues to scan at Tof-MS survey and only switches to Tof-MS/MS scan when multiply charged ions (precursor ions) in the Tof-MS survey spectrum meet the criterion for fragmentation. The majority of ions generated from tryptic peptides are at a charge status of at least 2. Details for the instrument settings are described below.

Time (min)	% A (Recipe 24)	% B (Recipe 25)	Comment
0.1	94	6	Trap and analytical column off-line; sample loading and desalting
3.0	94	6	
180.0	70	30	Trap and analytical column on-line; gradient separation; MS acquisition in this section
240.0	50	50	
242.0	20	80	
247.0	20	80	
248.0	94	6	Trap and analytical column off-line to equilibrated separately
300.0	94	6	

Table 2.1 - Liquid chromatography conditions for analysis of a phosphoprotein digest.

1. Acquire Tof-MS survey over the mass range m/z 400-1500 at 1 sec per scan.
2. The switching criterion from Tof-MS to Tof-MS/MS is dependent on precursor ion intensity and charge state (+2 and +3 selected), each for 5 sec.
3. The collision energy used to perform Tof-MS/MS is varied according to the m/z and the charge state of the precursor ion.
4. Acquire data for Tof-MS/MS over the mass range m/z 50-2000.
5. Perform Tof-MS/MS on up to three multiply charged ions of the highest intensity with a dynamic exclusion time of 120 sec.

   Note: depending on the complexity of the sample, more precursor ions can be selected from Tof-MS survey spectrum to subject to Tof-MS/MS.

6. For a highly complex sample, two LC-MS/MS runs with the same LC gradient may be used.
7. After the first standard LC-MS/MS run is finished (Run 1), examine the survey data.
8. Select up to 5 multiply charged ions above the intensity threshold from each survey spectrum that had not been subjected to MS/MS in Run 1 and incorporate into an inclusion list for the second run (Run 2).
9. Perform the Run 2 analysis using the same conditions as Run 1 except that the MS switching to MS/MS should be directed to analyse only multiply charged ions from the "including list".

   Note: We term this LC-MS/MS approach, incorporating a primary analysis followed by a repeat experiment based on an inclusion list, Targeted Repeat Analysis (TRA).

10. Process both LC-MS/MS data to pkl files using Proteinlynx in Masslynx and then merge to a single mgf file with the software (merge) provided by Matrix Science and submitted to Mascot search.

2.5.10 LC-MS/MS analyses of a tryptic digest of IMAC purified phospho-proteins on a 4000 Q TRAP

Although we have found that the majority of proteins in protein IMAC samples are phosphorylated, the molar ratio of phosphorylated to non-phosphorylated species is reduced upon trypic digestion. An alternative to performing a second round of enrichment of phosphopeptides is to selectively monitor phosphopeptides in the mass spectrometer. The specific behaviour of phosphopeptides in a mass spectrometer can be exploited in a number of ways for analysis of very complex peptide mixtures with a low abundance of phosphorylated peptides.

- Neutral loss of 98 in MS/MS for those phosphopeptides containing phosphoserine and phosphothreonine in positive electrospray ionisation mode.
- 216 fragment ion for phosphopeptides containing phosphotyrosine in positive electrospray ionisation mode.
- 79 fragment ion for phosphopeptides in negative electrospray ionisation mode.

By setting the mass spectrometer to specifically monitor precursor ions of the fragment ions, phosphopeptides can be selectively identified. In our experience, we found that monitoring the precursor of 79 fragment ion in the negative mode was the most efficient approach and the use of the 4000 Q TRAP which is capable of fast auto-switching between negative and positive ionisation mode is well suited for this.

Two precursor-scanning approaches based on monitoring the precursor of 79 fragment ion can be adopted on the 4000 Q TRAP to selectively analyse phosphopeptides. The LC conditions are the same as that used for IMAC purified phosphoprotein digest, described above.

Approach 1: Conventional precursor of 79 fragment ion scanning

1. The instrument is set in negative ion mode (-ESI) to scan for the precursors of m/z 79 (PO_3^-);
2. When detected precursor ions for 79, the instrument is automatically switched to positive mode (+ESI) then MS/MS fragmentation the detected precursors.
3. The steps are repeated.

Approach 2: TRA precursor of 79 fragment ion scanning

1. Analyse one-third of the sample in negative mode for detecting precursors of m/z 79.
2. Examine data to generate the inclusion list for the second run in positive ion mode (+ESI).
3. Analyse the remaining two-thirds of the sample in Information Dependent Acquisition (IDA) mode with the above inclusion list.

2.5.11 LC-MS/MS analyses of IMAC enriched phosphopeptides.

This is performed in a similar way to that described above for digested IMAC purified phosphoproteins, with the following exceptions.

1. Load phosphopeptides to the trap column and desalt with 0.5% formic acid (Recipe 29) at a flow rate of 10 µl/minute for 30-60 min depending on salt amount in the sample.

   Note: We suggest the trap column here is BetaMax Neutral 5 µm, 180 µm i.d. x 30 mm (Thermo Hypersil-Keystone).

Time (min)	% A (Recipe 27)	% B (Recipe 28)	Comment
0	99	1	Trap and analytical column on-line for gradient separation MS acquisition in this section
30	98	2	
75	97	3	
105	95	5	
125	90	10	
135	80	20	
145	50	50	
150	5	95	
160	5	95	
161	99	1	Trap and analytical column off-line to equilibrate separately
180	99	1	

Table 2.2 - Liquid chromatography conditions for analysis of a phosphopeptide mixture isolated by a peptide IMAC step.

2. Separate phosphopeptides with the gradient shown in Table 2.2, using mobile phases described in Recipes 27 and 28 (A as 99.9% H_2O-0.1% formic, B as 70% CH_3CN-29.9% H_2O-0.1% formic acid). The flow rate is 150 nl/min

   Note: This is due to the high pressure when the trap and analytical columns are setup on-line.

The majority of phosphopeptides elute at very low organic percentage. A comparison of liquid chromatography separations of a phosphoprotein digest and an IMAC enriched phosphopeptide mixture is shown in Figure 2.5.

2.6 Data Analysis

We recommend a Mascot V2.0 (Matrix Science) server for iterative searching of a custom, non-identical, combined human, mouse and rat IPI database (EBI).

1. Select trypsin with a maximum of 2 miscleavage sites when using Trypsin Gold from Promega or sequence grade trypsin from Roche; Select semi-trypsin with maximum 2 miscleavage sites for sequence grade modified trypsin from Promega as this trypsin produce some semi-tryptic fragments.
2. Select correct database with the appropriate taxonomy.
3. Set following variable modifications:
   Carbamyl (N-term)
   Carbamyl (K)
   Methyl ester (C-term) – when peptides are methyl esterified
   Methyl ester (DE) – when peptides are methyl esterified
   Oxidation (M)
   Phospho (ST)
   Phospho (Y)
4. Select monoisotopic mass
5. Set peptide tolerance at 0.4 Da and MS/MS tolerance at 0.3 Da though majority MW errors will be less than 0.2 Da if calibration is correct.
6. Select Data format as Micromass (.pkl) file when data file is generated by Proteinlynx in Masslynx; for merged file as described previously select "Mascot generic".
7. Protein identifications with peptides matching details will be given by Mascot as well as the possible phosphorylation sites.

Verify the phosphorylation sites by manually checking the spectra using PEAKS software. An example of a mass spectrum of a methyl-esterified di-phosphorylated peptide is shown in Figure 2.6.

2.7 Troubleshooting

We have found the protein IMAC protocol to be very reproducible, but in case of problems there are a number of parameters which should be checked. The most likely reason for failure or low yields in protein IMAC purifications is sample or buffer contamination. Small amounts of chelating chemicals such as EDTA or EGTA will strip the metal from the resin and contamination of the buffer with the metal used as the affinity reagent will result in reduced or no phosphoprotein yield. In the event of a failed IMAC purification, we recommend that in the first instance, that all buffers should be replaced and the efficiency monitored by western blotting or phosphoprotein staining. We recommend a working concentration of approximately 1 mg/ml for the protein extract and a ratio of 10:1 of extract to resin to ensure specific binding of phosphoproteins. Reduction of the protein extract concentration or extract to resin ratio may result in non-specific binding. The applicability of protein IMAC to integral membrane or membrane associated proteins is limited by their solubility in 6 M urea. We recommend that partial proteolysis with trypsin or endoproteinase Lys-C be carried out to release cytoplasmic domains of such proteins for subsequent protein and or peptide IMAC purification. Phosphoproteins may be eluted from the resin by boiling in at 97°C for 5 min in 5 X reducing sample buffer for subsequent SDS-PAGE separation but it should be noted that after specific chelating elution, low levels of non specific binding proteins are retained and would thus contaminate the sample prepared by a boiling elution. The peptide IMAC protocol is also subject to similar effects from contaminating chelators and metals and the same precautions should be taken to avoid this. There are two main problems associated with purification of phosphopeptides by IMAC: specificity of binding and elution efficiency. Traditional peptide IMAC suffers to varying degrees, from non specific binding of acidic non-phosphorylated peptides. This has been improved by the conversion of peptides to corresponding methyl-esters, but as

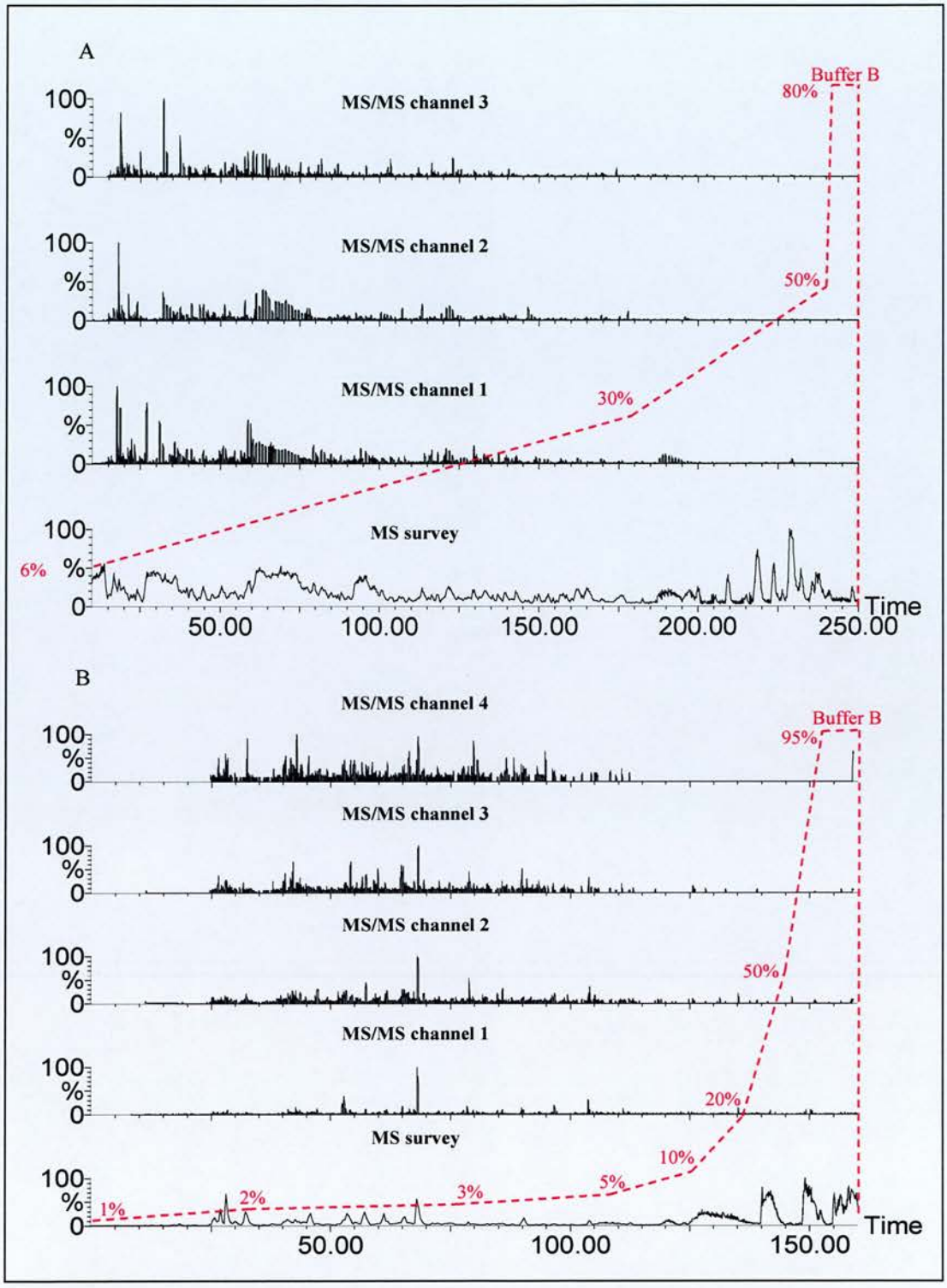


Figure 2.5 – Liquid chromatography desalting and separation of peptides
BPI (base peak intensity) chromatographs of Run 1 for the protein IMAC digest (A) and for the peptide IMAC purified phosphopeptides (B). The red dashed line represents the gradient plot in each chromatogram. A standard gradient was used for the protein IMAC digest (A) with the majority of peptides eluted by 30 % buffer B (32% organic). As most phosphopeptides isolated in a peptide IMAC step are multi-phosphorylated and hence quite hydrophilic, a shallow gradient of buffer B is used (B). It can be seen that the majority of phosphopeptides are eluted before 5% B (3.5% organic).

this reaction does not usually go to completion, this introduces other problems such as dilution of a peptide species into modified and unmodified forms. Phosphoproteins may be eluted from the resin by boiling in at 97°C for 5 min in 5 X reducing sample buffer for subsequent SDS-PAGE separation but it should be noted Non-specific interactions due to hydrophobic and ionic effects can be reduced by performing IMAC at pH 2.5-3, at which most glutamic and aspartic acid residues are protonated and by including up to 30% acetonitrile in the loading buffer. We recommend these later conditions in conjunction with and without methyl-esterification for greatest coverage of a phosphopeptide population in a sample.

2.8 Notes and remarks

It is generally accepted that currently no one approach for enrichment and analysis of phosphoproteins or phosphopeptides will enable identification of all phosphorylation sites (Steen et al., 2002). We support the idea that a multi-faceted approach such as that we have described in this protocol is necessary to achieve higher phosphoproteome coverage. It is also clear that a portion of the phosphoproteome is hidden from analysis in basal state conditions and those phosphoproteomic approaches which have been developed in the last number of years, should be applied to numerous activation/treatment states in a given system to reveal more of the low-abundance and transient phosphoproteome.

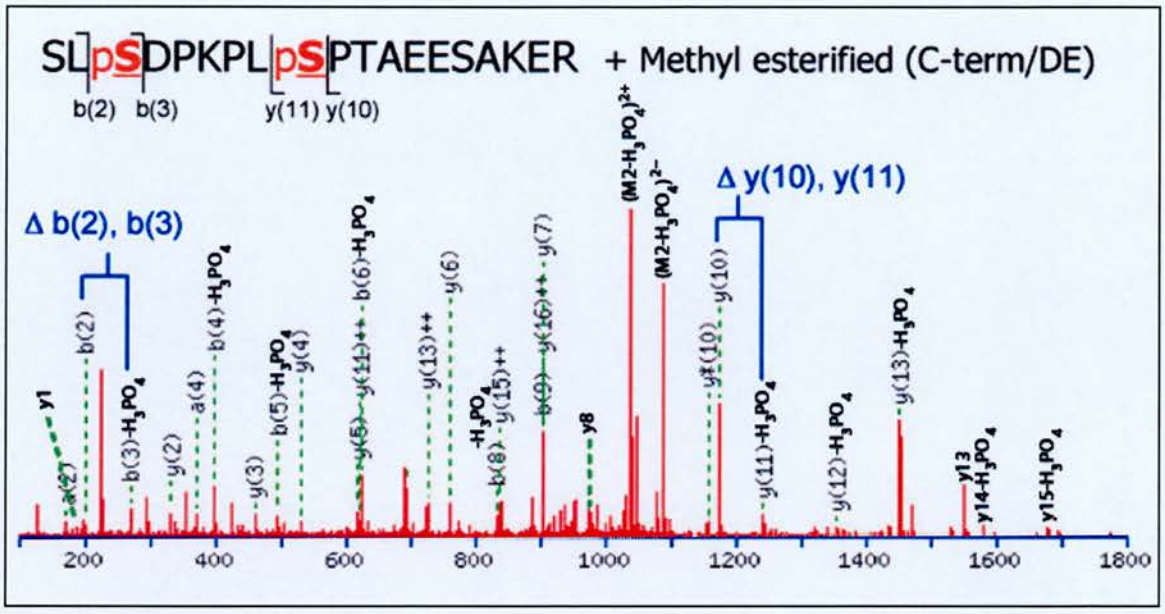


Figure 2.6 - Mass spectrum of a methyl-esterified phosphopeptide derived from Bassoon protein (Swiss-Prot accession O88737) showing phosphorylated serine residues 2860 and 2866. Phosphorylation of serine 2860 was identified by the mass difference between the b(2) and b(3) ions caused by neutral loss of phosphate in the mass spectrometer. 167 daltons (phosphoserine) minus 98 daltons (H3PO4) results in a mass difference of 69 daltons, which is characteristic of dehydroalanine, and therefore phosphoserine prior to neutral loss. Similarly, phosphorylation of serine 2866 was identified by the mass difference between the y(10) and y(11) ions that was caused by neutral loss of phosphate, resulting in a mass difference of 69 daltons.

CHAPTER 3

PROTEOMIC ANALYSIS OF *IN VIVO* PHOSPHORYLATED SYNAPTIC PROTEINS

3. Proteomic Analysis of in Vivo Phosphorylated Synaptic Proteins

3.1 ABSTRACT

In the nervous system, protein phosphorylation is an essential feature of synaptic function. Although protein phosphorylation is known to be important for many synaptic processes and in disease, little is known about global phosphorylation of synaptic proteins. Heterogeneity and low abundance make protein phosphorylation analysis difficult, particularly for mammalian tissue samples. Using a new approach, combining both protein and peptide Immobilised metal-affinity chromatography (IMAC) and mass spectrometry (MS) data acquisition strategies; we have produced the first large-scale map of the mouse synapse phosphoproteome. We report over 650 phosphorylation events corresponding to 331 sites (289 have been unambiguously assigned), 92% of which are novel. These represent 79 proteins, half of which are novel phosphoproteins, and include several highly phosphorylated proteins such as MAP1B (33 sites) and Bassoon (30 sites). An additional 149 candidate phospho-proteins were identified by profiling the composition of the protein IMAC enrichment. All major synaptic protein classes were observed, including components of important pre- and post-synaptic complexes as well as low abundance signalling proteins. Bioinformatic and in vitro phosphorylation assays of peptide arrays suggest a small number of kinases phosphorylate many proteins and each substrate is phosphorylated by many kinases. These data substantially increase existing knowledge of synapse protein phosphorylation and support a model where the synapse phosphoproteome is functionally organised into a highly interconnected signalling network.

3.2 INTRODUCTION

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 (Sheng, 2001). In recent years with the advent of proteomic technologies a coherent map of the mammalian synapse proteome is emerging. Mass spectrometry (MS) based analysis of the postsynaptic density (PSD) has established for the first time a detailed list of its molecular components (Yoshimura et al., 2004), (Peng et al., 2004), (Li et al., 2004). Systematic analysis of functional multiprotein complexes embedded in the PSD (Husi et al., 2000) has also added to our knowledge of the overall organisation of the postsynaptic proteome.

Central to the functioning of signalling complexes and indeed the most basic signalling 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 diverse signalling 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. SH2 binding of PI 3-kinase to NR2B) (Hisatsune et al., 1999) 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 (Greengard et al., 1993), (Sweatt and Kandel, 1989), but new proteomic strategies lend themselves to the global characterisation of the signalling properties of the synapse proteome. Phosphopeptides can be purified from complex protein mixtures using immobilised metal affinity chromatography (IMAC) (Andersson and

Porath, 1986) and identified using MS (Ficarro et al., 2002). Recent phosphoproteomic studies have utilised 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 (Ficarro et al., 2002), arabidopsis (Nuhse et al., 2003), and cell lines (Shu et al., 2004). However, the application of these approaches to complex mammalian sub-cellular organelles such as the synapse has yet to be established.

MS, although a powerful tool for analysis of phosphorylation analysis, 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 as well as 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 (Beausoleil et al., 2004), (Chang et al., 2004).

Here, we describe the use of a combination of cellular fractionation procedures with large-scale IMAC phosphoprotein and phosphopeptide enrichment protocols and complementary MS analytical strategies to characterise mouse forebrain synaptosomes. This has resulted in the unambiguous identification of 331 sites of phosphorylation in 79 synaptic proteins involved in important pre- and postsynaptic multiprotein complexes and signalling pathways. Large-scale in vitro phosphorylation screening on peptide arrays for 95 sites with 7 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 signalling in a particular system.

3.3 MATERIALS AND METHODS

3.3.1 Isolation of Synaptosomes

See section 2.5.1: Purification of Synaptosomes by Sub-cellular Fractionation.

3.3.2 Phosphoprotein/peptide purification

3.3.2.1 Protein IMAC of urea soluble synaptosomal fraction

Fast-flow chelating sepharose with iminodiacetic acid (IDA) (Amersham Biosciences) or nitrilotriacetic acid (NTA) (Qiagen) chelating groups were charged with GaCl3 or FeCl3. Synaptosomal proteins (12.5 mg) were solubilised in 6 M urea and the supernatant was removed and incubated with 2 ml of the metal charged resin with mixing for 1 hour at room temperature (RT). The unbound protein was washed with buffer A (6 M urea/50 mM tris-acetate) to baseline and the phosphoproteins were specifically eluted with buffer B (6 M 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 μg protein was diluted with buffer C (1 M urea, 0.125 M thiourea, 5% (v/v) CH3CN and 0.1 M NH4HCO3) to a final volume of 400 μl. Trypsin (sequence grade, Roche) was added to the sample in a ratio of 1:20 and the mixture was incubated at 37°C for 2 hours and then dried in a SpeedVac (Thermo Life Science). The sample was reconstituted in 50 mM NH4HCO3 and 1/4 was injected for each on-line LC-MS/MS analysis.

3.3.2.2 Double IMAC of urea soluble synaptosomal fraction

3 mg of protein IMAC purified sample was digested with modified porcine or gold trypsin (Promega) in a ratio of 1:20 in buffer D (1 M urea and 25 mM NH4HCO3) at 37°C for 4 hours. The resultant digest was desalted and dried and methyl esterification was performed with 2 M methanolic HCl, (10), when needed. Self Pack

POROS® 20 MC Media (Applied Biosystems) for phosphopeptide purification was charged with GaCl3 as described above for the IDA/NTA resins. Peptide digests (with and without methyl esterification) were reconstituted in buffer E (equal volumes of acetonitrile, methanol and water pH 2.5-3). 1 ml of this peptide mixture was incubated with 200 µl of POROS-Ga slurry for 1 hour at RT. The resin was then loaded into a spin-column and washed with 10 volumes of buffer E. Phosphopeptides were eluted with 2 x 100 µl of 200 mM Na2HPO4.

3.3.2.3 Peptide IMAC of whole and urea insoluble synaptosomal fractions

2.5 mg of the 6 M urea insoluble pellet left after removal of the supernatant was digested with trypsin in buffer F (2 M urea and 25 mM NH4HCO3). The resultant digest was desalted, methyl-esterified and subjected to peptide IMAC step as described previously. Similarly, 6 mg of whole synaptosomal proteins were digested in buffer F, desalted, methyl-esterified and subjected to peptide IMAC step.

3.3.3 Phosphoprotein staining and image analysis

IMAC enriched phosphoprotein samples were separated on a 12 % SDS-PAGE gel and sequentially stained with Pro-Q diamond (phosphoprotein) and SYPRO Ruby (total protein) stains (Molecular Probes). Peppermint Stick 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 (Nonlinear Dynamics).

3.3.4 Mass spectrometry

Mass spectrometry was performed by Lu Yu (The Wellcome Trust Sanger Institute)

3.2.4.1 Online nano LC-MS/MS

On-line nano LC-MS/MSA nanoflow HPLC 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 (FA) 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 FA. Peptides on the trap were back-flushed to, and separated on the analytical column (PepMap C18, 75 μm id x 15 cm, LC Packings). The gradients are shown Table 3.1.

3.2.4.2 Analysis of IMAC enriched phosphoprotein digests

In the LC-MS/MS analysis of the phosphoprotein digest, the Q-Tof Ultima was operated in automated Data Dependent Acquisition (DDA) mode. Each cycle had a 1 sec 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 sec. 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 with 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).

A

Time (min)	0	180	240	242	247
A (%)	94	30	50	20	20
B (%)	6	70	50	80	80

B

Time (min)	0	30	75	105	125	135	145	150	160
A (%)	99	98	97	95	90	80	50	5	5
B (%)	1	2	3	5	10	20	50	95	95

Table 3.1. Liquid chromatography gradients

A. LC gradient for separation of phosphoprotein digest. Solvent A was 95% H2O-5% ACN/0.1% FA, solvent B was 95% ACN-5% H2O / 0.1% FA. Flow rate through the columns in the separation was 200-250 nL/min.

B. LC gradient for IMAC enriched phosphopeptides analysis. Solvent A was 100% H2O/0.1% FA, solvent B was 70% ACN-30% H2O/0.1% FA. Flow rate through the columns in the separation was 150 nL/min.

3.2.4.3 Analysis of IMAC enriched phosphopeptides

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 analyse phosphopeptides. In the first approach the instrument was used in negative ion mode (-ESI) to scan for the precursors of m/z 79 (PO3-) with automatic switching to positive mode (+ESI) for MS/MS for the detected precursors. In the second approach, we applied TRA strategy. First, one-third of the sample was analysed 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 (IDA) mode. Raw data was processed to give a peak list file and submitted to a local Mascot V2.0 (Matrix Science) server for iterative searching on a custom, non-identical, combined human and mouse IPI database (EBI). Assignment of phosphorylation sites was verified manually with the aid of PEAK Studio (Bioinformatics Solutions) software.

3.3.5 Peptide array phosphorylation assays

Peptide array data was generated by Marcelo Coba (The Wellcome Trust Sanger Institute)

Jerini Phosphosite detectorTM peptide arrays (Jerini Peptide Technologies, Gmbh, Germany) were used to determine which of the MS identified phosphorylation sites (selected 95 sites) could be phosphorylated by 7 kinases in an in vitro assay. 15 amino acid long peptides which encompassed the selected sites were synthesised on cellulose membranes in a parallel manner using SPOT technology (Wenschuh et al., 2000) deposited to glass slides and were covalently Immobilised to the glass slide surface. Each peptide was present in triplicate on the chip, and 7 full length proteins which are capable of being phosphorylated were also included. Negative control peptides for each phosphorylation site were included, replacing serine or threonine

with alanine or valine, respectively and positive control peptide sequences for each kinase were present.

3.2.5.1 Kinase reactions

Peptide arrays were sealed with Gene-Frame TM incubation chambers (Abgene house, Surrey, UK) 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 DTT, 100 µM ATP, 15 mM MgCl2 and 10 µM [γ-32] ATP). Recombinant active kinases (Upstate), (3 µg of PKA catalytic subunit (with 2 µM cAMP), 2 µg of Akt1 (delta 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 DTT, 0.5 mM CaCl2, 0.1 mg/ml phosphatidylserine, 25 µg/ml diacylglycerol).

After a 45 minute incubation at 32 °C the peptide microarrays were washed 6 times, alternating between 0.1 M phosphoric acid and distilled water.

3.2.5.2 Peptide array image analysis

γ-32 P incorporation in the Immobilised peptide spots was detected on a Typhoon 8600 Phosphor imager (50 µm resolution), (Amersham Biosciences). Image analysis and signal quantification was carried out using ImageQuant TL (Amersham Biosciences) and positive signals were defined after background subtraction.

3.3.6 Bioinformatic analysis

All proteins detected in this study were classified according to Swiss-Prot keywords (http://ca.expasy.org/sprot/). Known phosphoproteins in Supplementary Table 1 were annotated by extensive literature mining of PubMed. Scansite

(http://scansite.mit.edu/) was used in high stringency mode to predict if proteins in Supplementary Table 1 were likely to be phosphorylated and responsible kinases are indicated. TMHMM Server v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM/) was used to predict the occurrence of transmembrane helices. Scansite and NetPhosK (http://www.cbs.dtu.dk/services/NetPhos/) were both used for predicting phosphorylation sites by cognate kinases, and on those peptides where the phosphorylation sites could not be unambiguously assigned by MS data (Supplementary Table 3). Scansite was also used to check the presence of phosphorylation sites in Pfam protein domains and to predict if phosphorylation sites were localised in phospho-dependent interaction domains (Supplementary Table 3). Graphical illustrations of the location of identified sites in relation to protein domain structure, 3D structure and information on homologs and paralogs of identified phosphorylation sites can be found at http://www.ppo4.org/phospho/synaptosome.html. Network construction: Protein-protein interactions for NMDA receptor complex proteins were obtained from PPID (www.PPID.org). Each protein (node) which had interaction (edge) information was used to plot a network graph using InterViewer, a network graphing program. Kinase-substrate edges were manually superimposed onto the protein-protein interaction network.

3.4 RESULTS

3.4.1 Methodological overview

Several approaches capable of identifying large numbers of in vivo phosphorylated proteins have been used to functionally annotate the mouse synaptic proteome. A summary of this is shown in Figure 1A. In the first approach, a mouse synaptosomal preparation was digested with trypsin and phosphopeptides captured by peptide IMAC. The enriched sample was analysed by nano-flow LC-MS/MS. Despite the enrichment, the majority of identifications corresponded to unphosphorylated

peptides. In the second approach, the synaptosomes were first divided into urea soluble and insoluble fractions, the latter containing mostly membrane bound components. Peptide IMAC and MS was performed on the insoluble fraction in a comparable manner to the first approach. For the urea soluble fraction a new strategy was adopted, entailing an additional stage of enrichment by capture of the phosphoproteins. A two-pronged route was taken for subsequent analysis of the trypsin digest of this enriched fraction. The first route is based on rigorous MS using a) LC-MS/MS peptide identification based on a targeted repeat analysis (TRA) approach described in the methods section, and b) precursor ion scanning to identify phosphopeptides prior to sequencing. The second route used a further peptide IMAC followed by LC-MS/MS. Using these combined approaches, and implementing a previously unreported strategy of sequential protein and peptide IMAC, we have identified 228 potential synaptic phosphoproteins (Supplementary Table 1) and characterised over 350 phosphopeptides containing 331 sites of phosphorylation (Supplementary Table 2, 3).

3.4.2 Synaptic phosphoprotein analysis

3.4.2.1 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 NTA (a quadradentate ligand) and two metal ions, FeCl3 and GaCl3 (Figure 3.2A), to isolate phosphoproteins from urea soluble preparations. The IDA resin with both metal ions showed marked enrichment indicated by specific phosphoprotein staining on SDS-PAGE gels. The Ga3+ resin showed more effective depletion of phosphoprotein from the unbound fraction together with stronger phospho-staining of the purified sample (Figure 3.2B), and therefore was used in all subsequent protein IMAC experiments.

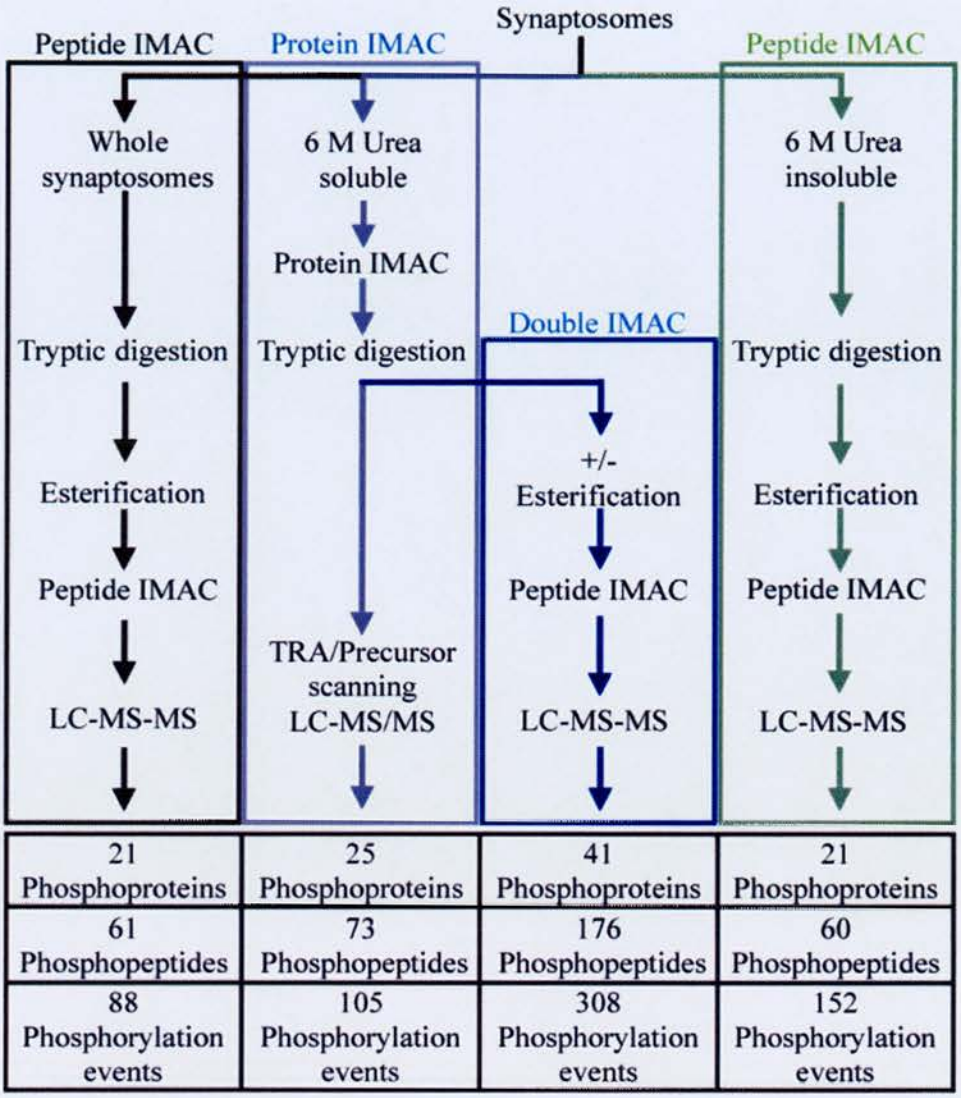


Figure 3.1 - Schematic representation of protein and peptide purification strategies.
Alternative starting material was derived from the differential solubility of synaptosomes in 6 M urea. Large-scale protein identification was achieved by TRA LC-MS/MS analysis of the IMAC purified phosphoproteins from the urea soluble fraction with some of the most abundant phosphopeptides characterized at this step. Further phosphopeptide identification was achieved by sequential protein and peptide IMAC purification, with esterification and single peptide IMAC with esterification of the urea insoluble fraction and of whole synaptosomal digests.

Analysis of tryptic peptides from the protein IMAC enrichment by LC-MS/MS identified 152 proteins and 19 phosphopeptides (Supplementary Tables 1, 2). TRA extended this by an additional 30 protein identifications and 21 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 in the methods section. 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 (Figure 3.3). This suggests that whilst a small proportion may represent contaminating proteins, the majority are likely phosphoproteins.

The identified components represent a diverse range of protein classes (Figure 3.3B, Supplementary Table 1) 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 characterised MS analysis of the phosphoprotein mixture cover single phosphorylation events (Figure 3.4), and thus is complementary to measurements made from the peptide IMAC approach that represents highly phosphorylated peptides better. A further important benefit is that it provides scope for two stages of sample enrichment, at the protein and subsequently at the peptide level.

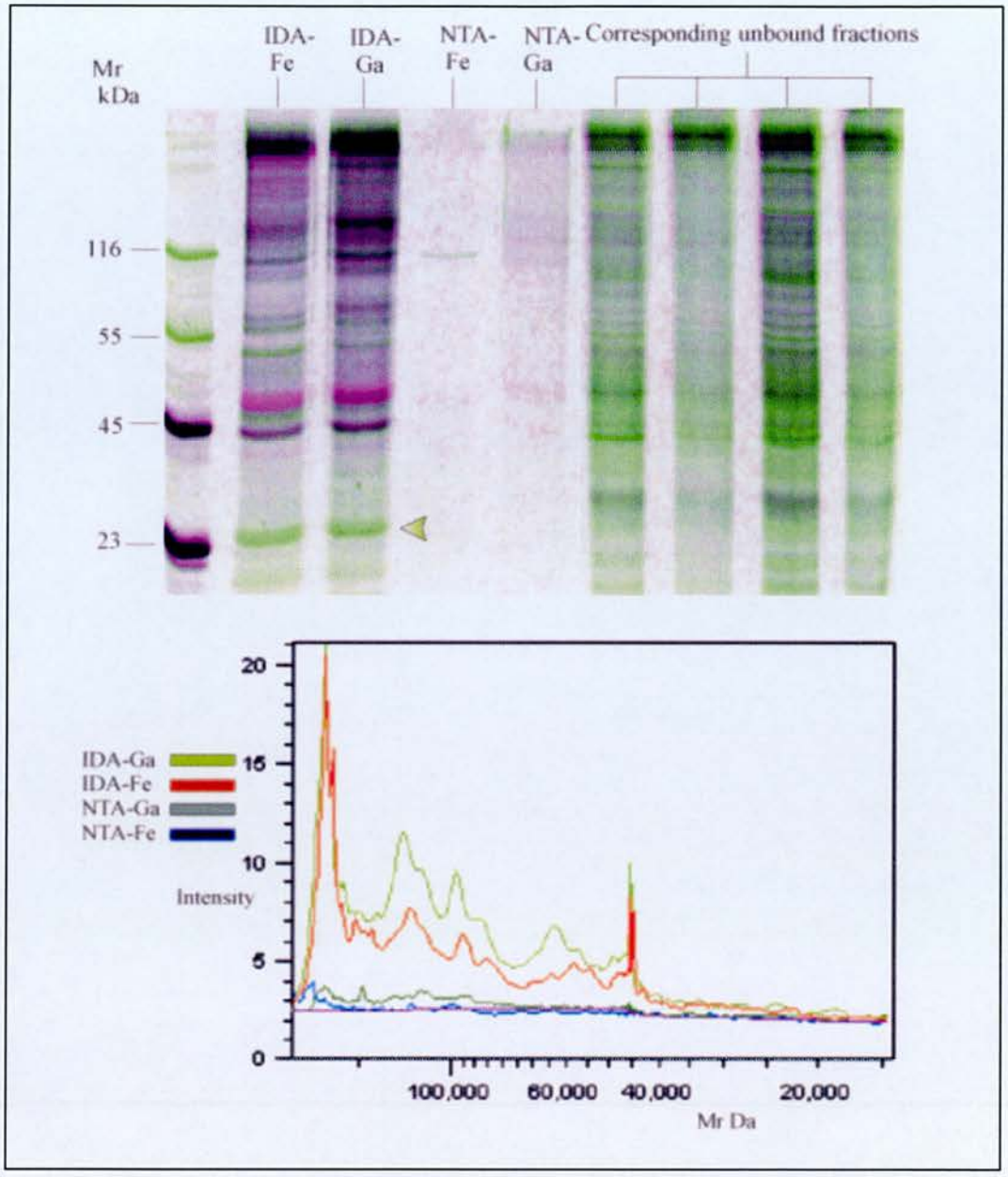


Figure 3.2 - Development of IMAC phosphoprotein isolation strategy.
A. Overlay of Pro-Q diamond phosphoprotein/ Ruby total protein stains of resin/metal IMAC combinations used to selectively isolate phosphoproteins from urea soluble synaptosomal extract. The corresponding unbound fractions are included for comparison. 20% of eluted protein from each purification was separated on a 12% SDS-PAGE gel and stained and imaged sequentially. Phosphorylated protein appears pink to black, whereas unphosphorylated protein appears green. Lane 1 shows Peppermint Stick molecular weight markers in which proteins of molecular weight 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 resins due to ion-exchange effects, and therefore removed from the dataset. **B.** Relative signal intensities of the Pro-Q diamond phosphoprotein stained image used in the overlay in Figure 3.2A. Phosphoprotein signals for lanes 2-5 in A are shown and plotted versus molecular weight. IDA-Ga has the highest overall signal intensity and thus the most enriched in phosphoprotein.

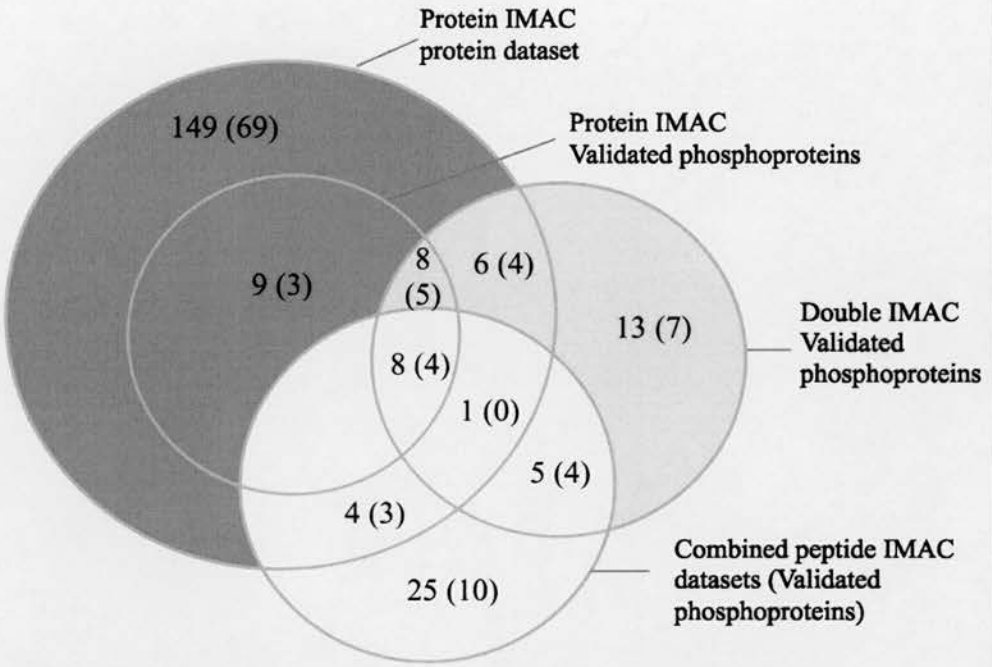
3.4.2.2 Double IMAC protocol

LC-MS/MS analysis of phosphopeptide from double IMAC yielded identification and characterisation of 176 phosphopeptides (Figure 3.3A, Supplementary Table 2) 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 7 proteins detected in the protein IMAC sample. Second, 19 phosphoproteins containing 37 phosphopeptides that were not characterised 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 (Supplementary Table 3). It is evident that combination of protein and double IMAC strategies enables identification and characterisation of phosphoproteins across a wider range of protein abundance and phosphorylation states.

3.4.3 Synaptic membrane phosphoprotein analysis

Integral membrane proteins are usually difficult to analyse since detergent based extraction methods are not readily compatible with subsequent purification strategies and LC-MS/MS analysis. However, for a phospho-proteomic analysis, cytoplasmic domains of integral membrane proteins are sufficient, since they contain the phosphorylation sites involved in intracellular signalling.

A



B

Protein class	PSD Proteins	NRC Proteins	Validated Phosphoproteins	Total
Channels/receptors	4	0	7	7
Transporters/transporting ATPases	5	2	9	12
Scaffolding and adaptors	17	10	17	21
Kinases	14	5	17	19
Phosphatases	4	1	2	4
Small G-proteins and modulators	12	0	8	18
Other signaling molecules	6	5	7	10
Other enzymes	12	5	8	20
Cell adhesion and cytoskeletal	55	16	51	69
Nuclear/transcriptional	0	0	6	10
Translational	2	0	2	3
Novel/uncharacterised proteins	3	0	14	35
Total	**134**	**44**	**148**	**228**

Figure 3.3 Summary of Datasets

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 6 MS based studies of the postsynaptic density and with the NMDA-receptor complex (NRC). Proteins in which we found phosphorylation sites (Experimental) and known phosphoproteins (Literature) are indicated.

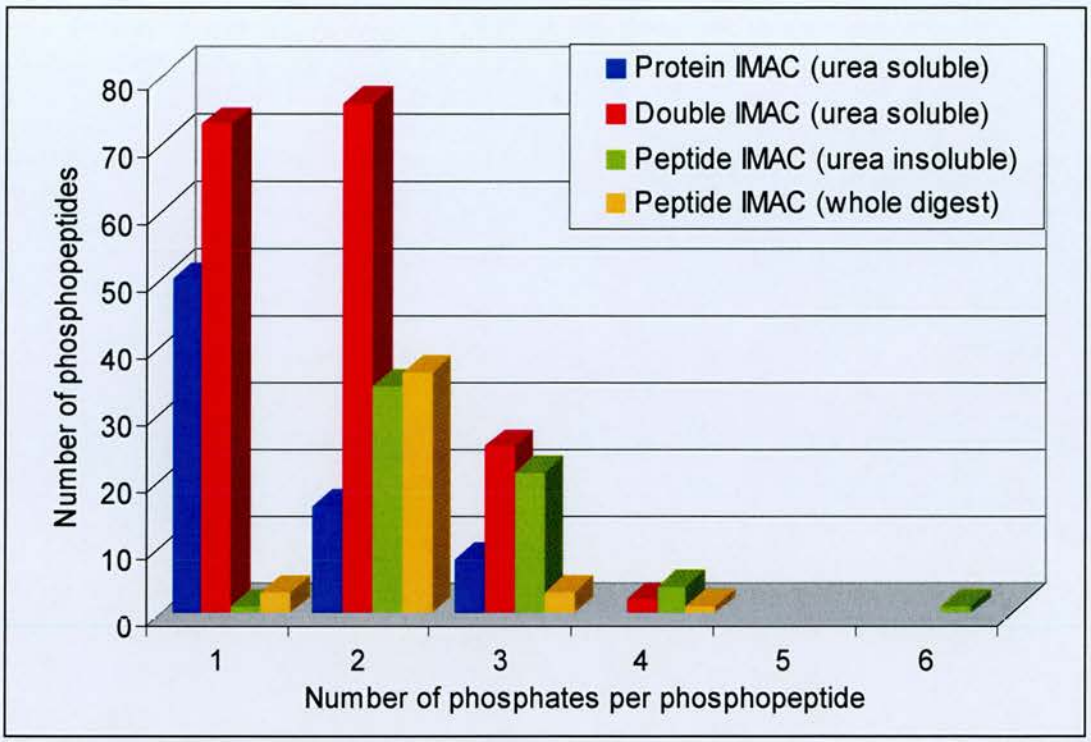


Figure 3.4 Distribution of the number of phosphates per phosphopeptide identified. The values are representative of all phosphopeptides detected by each approach and therefore contain some peptides covering overlapping sequences in proteins but with differing numbers of phosphates attached.

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 (Figure 3.1) This approach allowed identification and characterisation of 60 phosphopeptides from 31 proteins (Supplementary Table 2), of which 12 are predicted to be integral membrane proteins, and many others such as PSD-95 and Adapter-related protein complex-2 alpha-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 (Figure 3.3A).

3.4.4 Complementarity of analytical approaches

Inspection of the summariseddata (Figure 3.3A, 3.4) shows that the combination of sub-cellular fractionation and IMAC protocols with MS analyses are complementary, both in terms of the protein identification, as well as phosphorylation site characterisation. Overall, we sequenced 350 phosphopeptides covering 653 phosphorylation events. These correspond to a total of 331 phosphorylation sites of which 289 were localised unambiguously (Figure 3.1). The phosphorylation sites were found in 79 synaptic proteins and we identified a further 149 candidate phosphoproteins. Interestingly, the large numbers of phosphorylation events characterised 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 phosphorylation of peptides detected using the different protocols (Figure 3.4). 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 favouring detection of mono-phosphorylated peptides. The combination

of protein and peptide IMAC results in more representative coverage, with very similar preferences for mono- and di-phosphorylated 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 (Ballif et al., 2004). This approach yielded peptides with an average of 1.15 phosphates per peptide, compared to our 1.85 per 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 multiple phosphorylated peptides (up to 6 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 are greater that that observed for the protein IMAC (1.44) and double IMAC (1.75) analysis of the urea soluble fraction (Figure 4.3). The enrichment of multi-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, 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 previously reports in other large-scale studies (11), (12), (13). Since most of the phospho-tyrosines we characterised are on multi-phosphorylated 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 multi-phosphorylated peptides.

3.4.5 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 (Figure 3.3 and Supplementary Table 1). Systematic PubMed searching on the identified components revealed that 110 have been described in the literature as phosphoproteins (Supplementary Table 1 and Figure 3.3). Of the 79 synaptic phosphoproteins characterised 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 149 putative phosphoproteins, which we have characterised 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 further mapping of these proteins will produce confirmation of their phosphorylation (Figure 3.3A) and further validation of this approach. Within our data set we have also identified phosphorylation sites in 14 novel or uncharacterised 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 utilised 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 44-IG(S)(T)(T)NPFLDIPHDPNAAVYK-64 from mKIAA0942 protein was detected in the protein IMAC analysis and contained 1 p(S/T) phosphorylation site. Scansite predicted that the most likely site to be phosphorylated in this sequence (by PKC delta) was S46 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 localisation of the remaining 43 ambiguous phosphorylation sites (Supplementary Table 3).

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 3). All sites could be accounted for by a minimum set of 23 kinases (Figure 3.5) and 14 of these are shown in the literature to phosphorylate proteins in our dataset. The predicted kinases emerged as two groups; those that phosphorylate multiple sites on proteins (phosphate/substrate ratio > 1 (9 kinases)) (left of dashed line, Figure 3.5) and those that phosphorylate a single site per protein (phosphate/substrate ratio = 1 (14 kinases)). 9 kinases could each phosphorylate 10 or more sites on multiple substrates and together could account for 258 of the observed sites and 8 kinases could phosphorylate ten or more distinct synaptic substrates. In agreement with another report, we noted that the majority of sites cluster outside characterised protein domains (Nuhse et al., 2004) (Supplementary Table 3). 25 phosphorylation sites corresponded to residues predicted to mediate phosphorylation-dependent protein-protein interactions such as 14-3-3 and SH2 domains.

3.4.6 In vitro phosphorylation assays

Peptide array technology was employed to evaluate kinase specificity and multiplicity inferred from bioinformatic analysis of the synaptic phosphorylation dataset. 190 peptides were synthesised 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. 7 kinases (PKA, Akt, PKC (α, β, γ), CK2, p38 MAPK, ERK1, Cdk5) were used

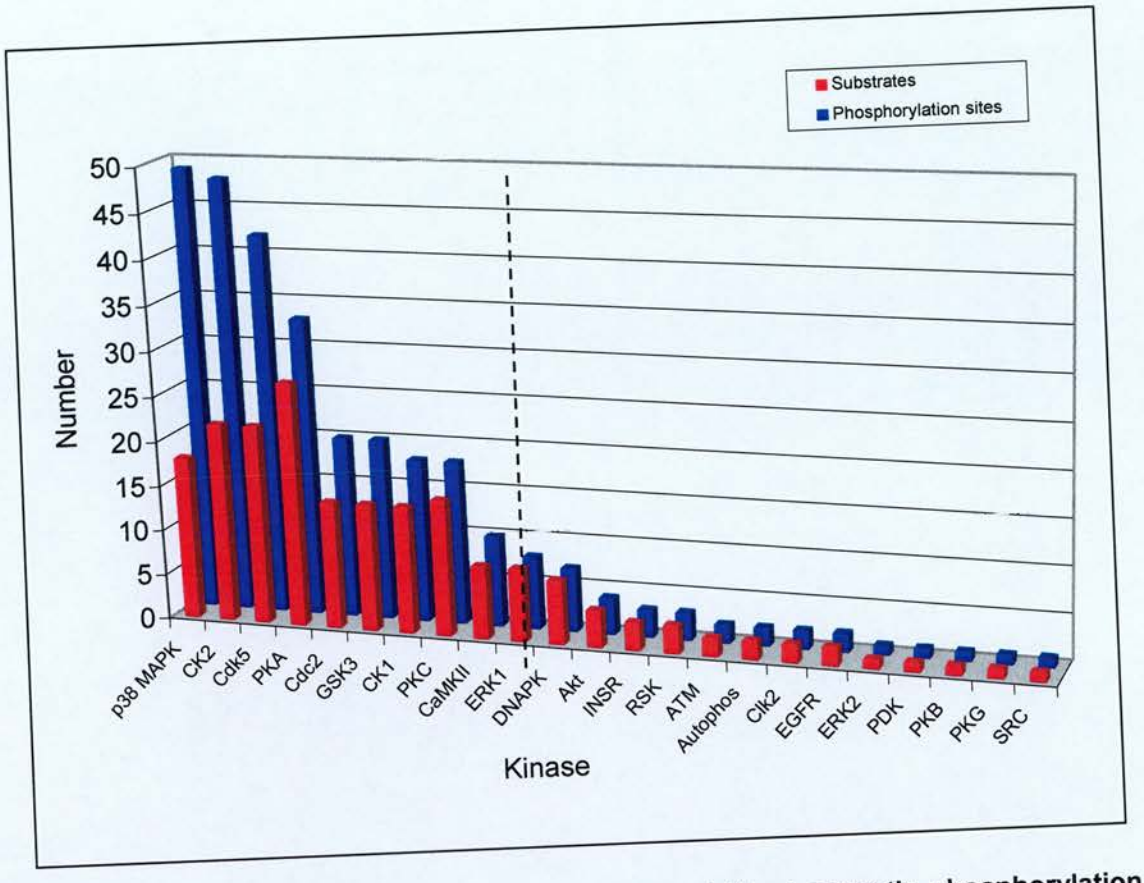


Figure 3.5 - Predicted kinase multiplicity for 289 defined synaptic phosphorylation sites. Scansite and NetPhosK kinase predictions and in a few cases data from the literature were used to evaluate the number of phosphorylation sites and distinct substrates for 23 kinases on 289 sites in synapse proteins. The total number of predicted sites (blue) and substrate proteins (red) are shown on the Y-axis for a given kinase (X-axis). The dashed vertical line between CamKII and ERK divides the kinases into two groups; the left-hand set has kinase/substrate ratio of >1 and the right-hand set has a kinase/substrate ration of 1.

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 (Figure 3.6) 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 7 kinases used on the 28 positive sites in the peptide array experiments. Scansite and NetPhosK predicted 58 % and 63 % of sites, respectively (Supplementary Figure 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 dataset.

Accession	Phosphoprotein	Residue	Cognate kinases determined by the peptide arrays						
			PKA	Akt1	PKC	CK2	p38	ERK1	Cdk5
Q9Z2H5	Band 4.1-like protein 1	T650				X			
O88737	Bassoon	S2122	X						
O88737	Bassoon	T2908					X	X	
O88737	Bassoon	S3213	X						
O88737	Bassoon	S3217					X	X	X
P28652	Calcium/calmodulin-dependent protein kinase type II beta	S315	X						
Q9R196	Calcium-activated potassium channel alpha subunit 1	S711						X	X
Q8K1M3	cAMP-dependent protein kinase type II-alpha reg chain	S77				X			
O35927	Delta catenin-2	S531	X		X				
Q8BJ42	DISKS large-associated protein 2	S440	X	X	X				
Q8BJ42	DISKS large-associated protein 2	S450					X	X	
AAO89219	DISKS large-associated protein 3	S58					X	X	X
Q8BJY3	Glutamate receptor, metabotropic 5	S135					X	X	X
Q8BJY3	Glutamate receptor, metabotropic 5	S137					X	X	X
P20357	Microtubule-associated protein 2	S1800			X				
P14873	Microtubule-associated protein 1B	S1775					X	X	X
P10637	Microtubule-associated protein Tau	S493							X
BAC65687	mKIAA0942 protein	S46	X		X				
P04370	Myelin basic protein (generic)	S144	X		X				
O35413	nArgbp2	S844	X						
Q9Z0P4	Paralemmin	T141					X	X	X
Q9Z0P4	Paralemmin	T145							X
Q9Z0P4	Paralemmin	S157					X	X	X
Q99569	Plakophilin 4	S510	X						
P70207	Plexin 2	S1620					X	X	
Q91XM9	PSD-93/Chapsyn-110	S414	X						
Q62108	PSD-95/SAP90	S418	X						
Q9JIR4	Rab3 effector, RIM-1	S895						X	X

Table 3.2. MS-identified phosphorylation sites with their cognate kinases as assessed by in vitro phosphorylation assays on immobilized peptide arrays. Detected signals which were positive after background subtraction are indicated with "X". Cognate kinase predictions for these selected phosphorylation sites were carried out using Scansite and with NetPhosK, and are indicated by yellow and green boxes respectively. Blue boxes indicate when both programs predicted the same kinase for a given site.

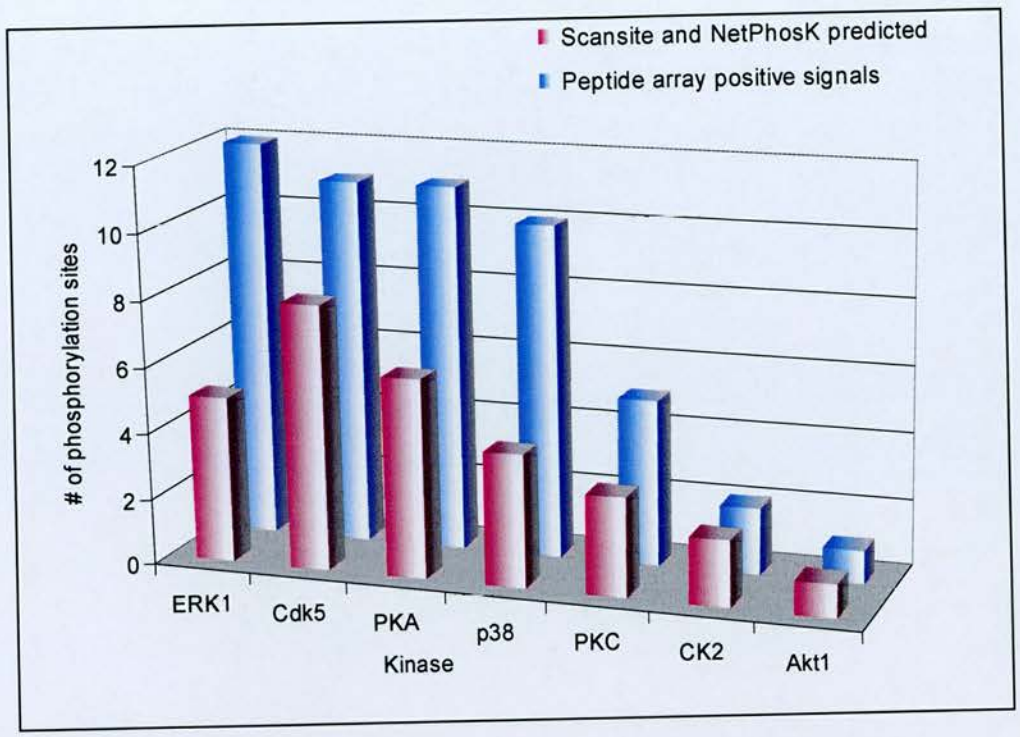


Figure 3.6 - Comparison of cognate kinase identifications for 28 phosphorylation sites screened by peptide arrays

The number of kinases and binding sites with known consensus sequences constrains global predictions, however in combination with datasets 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.

3.5 DISCUSSION

3.5.1 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 dataset contained postsynaptic components including the postsynaptic density (134 proteins) and NMDA-receptor associated proteins (44 proteins) (Figure 3.3) and presynaptic components, of which the most striking examples were Piccolo-Bassoon transport vesicles.

3.4.1.1 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 (Sudhof, 2004). It is well known that phosphorylation is critical to the regulation of calcium regulated synaptic vesicle exocytosis (Turner et al., 1999) and we have identified a number of phosphoproteins known to play a role in this process (Figure 3.7). Piccolo-Bassoon transport vesicles (PVT) have been shown to be putative precursor vesicles in the active zone (Shapira et al., 2003), and contain a number of proteins functionally coupled to synaptic vesicle exocytosis. Bassoon, a presynaptic protein with homology to Piccolo, has is intimately involved in vesicle release (Shapira et al., 2003) and together with Piccolo and Rim-1/2, constitute half

of the protein content of PTVs. 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 which is associated with PTVs (Shapira et al., 2003). PKA phosphorylation of an isoform of Rim1 regulates presynaptic LTP (Lonart et al., 2003) highlighting the regulatory importance of such presynaptic molecules. In addition to the main constituents of PVTs, we mapped 10 phosphorylation sites on Munc-18 (Figure 3.9). Phosphorylation of Munc-18 by PKC has been shown to disassemble Munc-18-syntaxin1A complexes and may have a regulatory role in presynaptic plasticity (Craig et al., 2003). 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 (Figure 3.7).

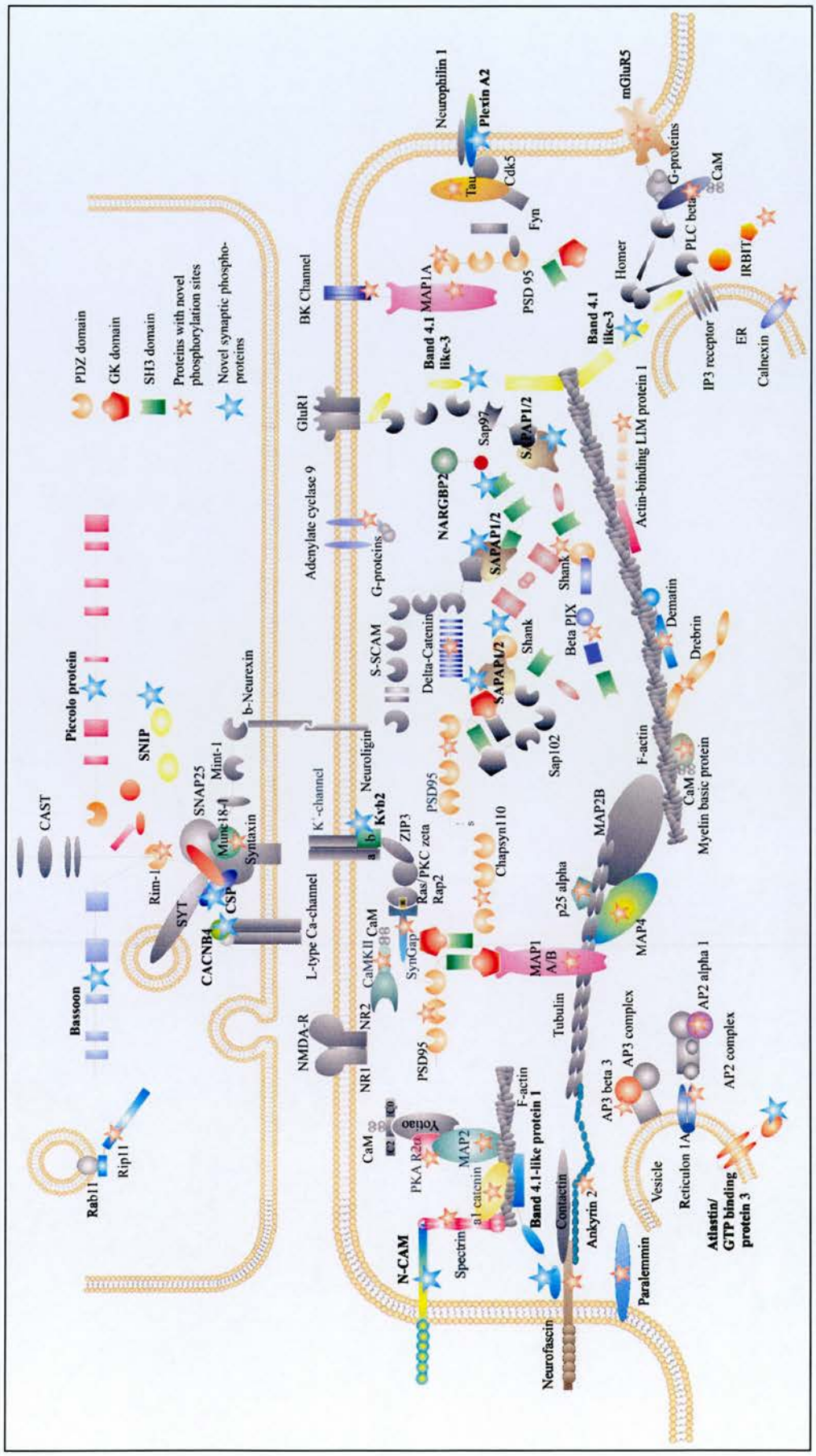


Figure 3.7 - Schematic illustration of phosphoproteins and their interaction partners at the synapse. Selected phosphoproteins characterised in this study were used to construct a synaptic interaction diagram. Protein-protein interaction information was gathered from PPID (www.ppid.org). All phosphoproteins are coloured and proteins shaded in grey are known synaptic proteins and are included for protein-protein interaction completeness. Proteins with novel phosphorylation sites are marked with a red star and novel phosphoproteins are marked with a blue star.

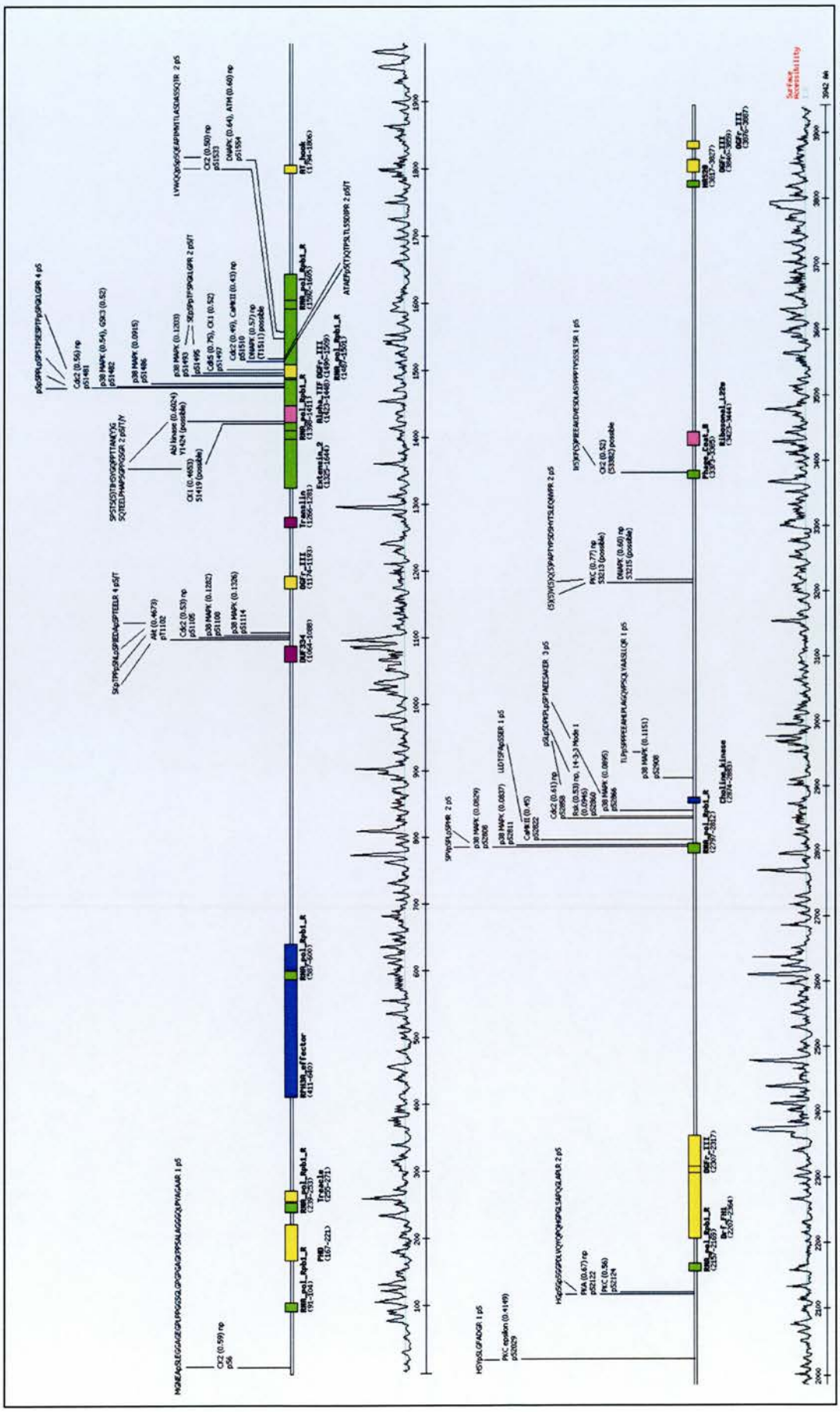


Figure 3.8 – Protein domain map of Bassoon protein. The location of 30 phosphorylation sites and the phosphopeptides they were found in are indicated. Scansite was used to predict the most likely kinases to phosphorylate these sites and if any of these sites could be potential phosphorylation dependent binding sites for example 14-3-3 binding.

3.4.1.2 Postsynaptic multiprotein complexes and pathways

The neurotransmitter glutamate activates synaptic plasticity primarily via the ionotropic N-methyl-D-aspartate (NMDA) receptor and metabotropic (mGluR) receptors. This leads to Ca2+ elevation in the dendritic spine and signal transduction to AMPA (alpha-amino-3-hydroxy-5-methylisoxa-zole-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 (Husi et al., 2000), (Grant, 2003).

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 (Figures 2, 4). Phosphopeptides derived from three PDZ domain-containing proteins (Shank1, and PSD-93/Chapsyn110 and PSD-95) were detected in this study. PDZ domain-containing are important components of the NRC and the postsynaptic density and form protein scaffolds which support and regulate many essential synaptic signalling processes (Nourry et al., 2003). We found a phosphorylation site in PSD-93 (S414) which was also found in MS analyses of purified NRC complexes (unpublished results). As 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 (Nada et al., 2003). Three phosphorylation sites were mapped in PSD-95 within a 20 amino acid stretch, 2 pSer just before its SH3 domain and a pTyr (Y432) located at the start of its SH3 domain in the β1 sheet (Figures 3.10, 3.11). Structural analysis of the SH3 and Guanylate kinase (GK) domains of PSD-95 (McGee et al., 2001), (Tavares et al., 2001) showed that there is an intermolecular interaction between the SH3 and GK domains involving the formation of a β-sheet including residues N- and C- terminal to the SH3 domain.

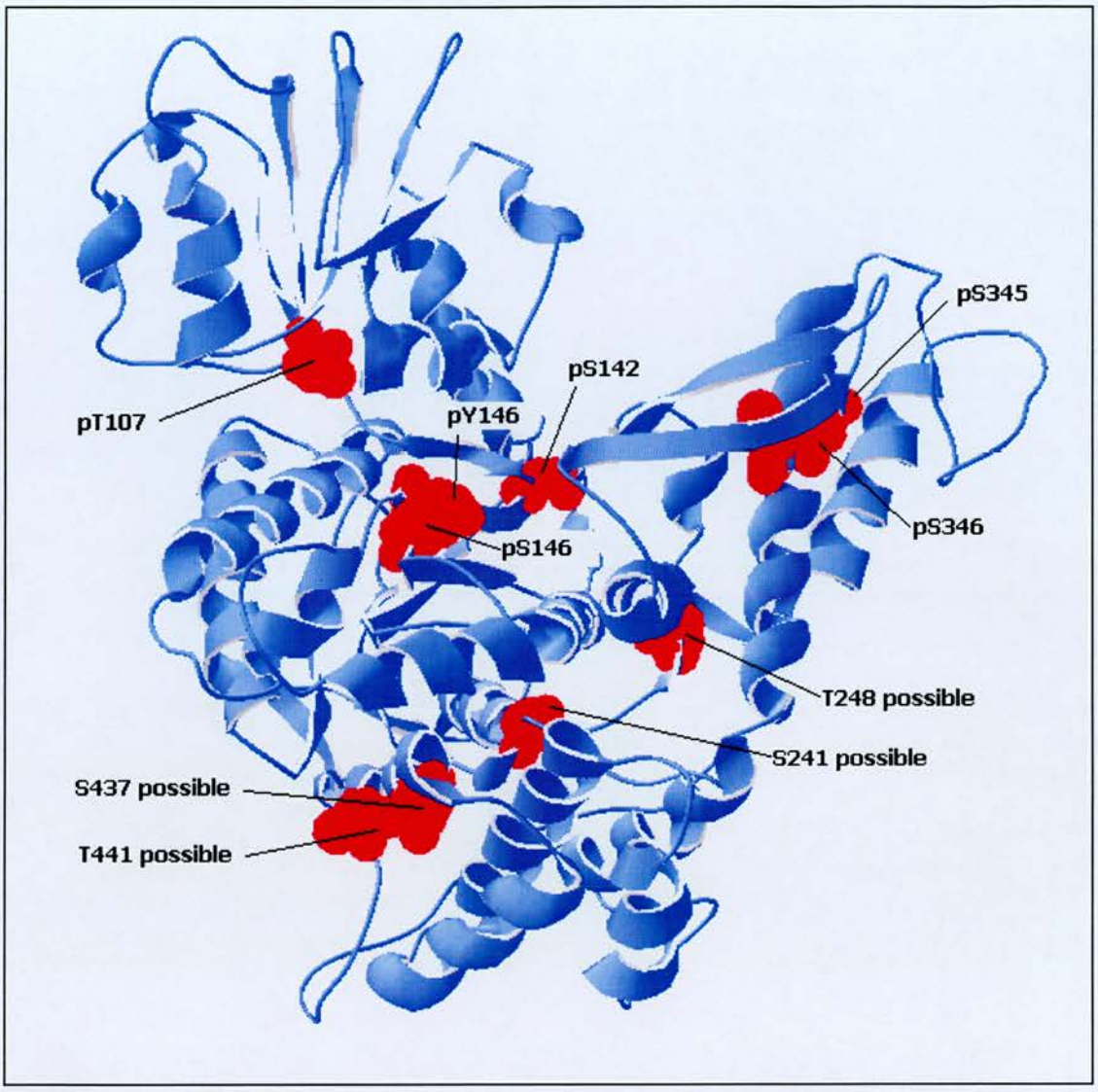


Figure 3.9 – Predicted 3D structure of Munc-18-1. The structure of Munc-18-1 was predicted by SwissModel, a Comparative Protein Modelling Server using 5 known template structures chosen by primary sequence homology. The location of 6 defined phosphorylation sites and 4 phosphorylation sites which were not exactly defined are indicated on the structure. It can be seen that all of these phosphorylation sites are in accessible regions of low complexity.

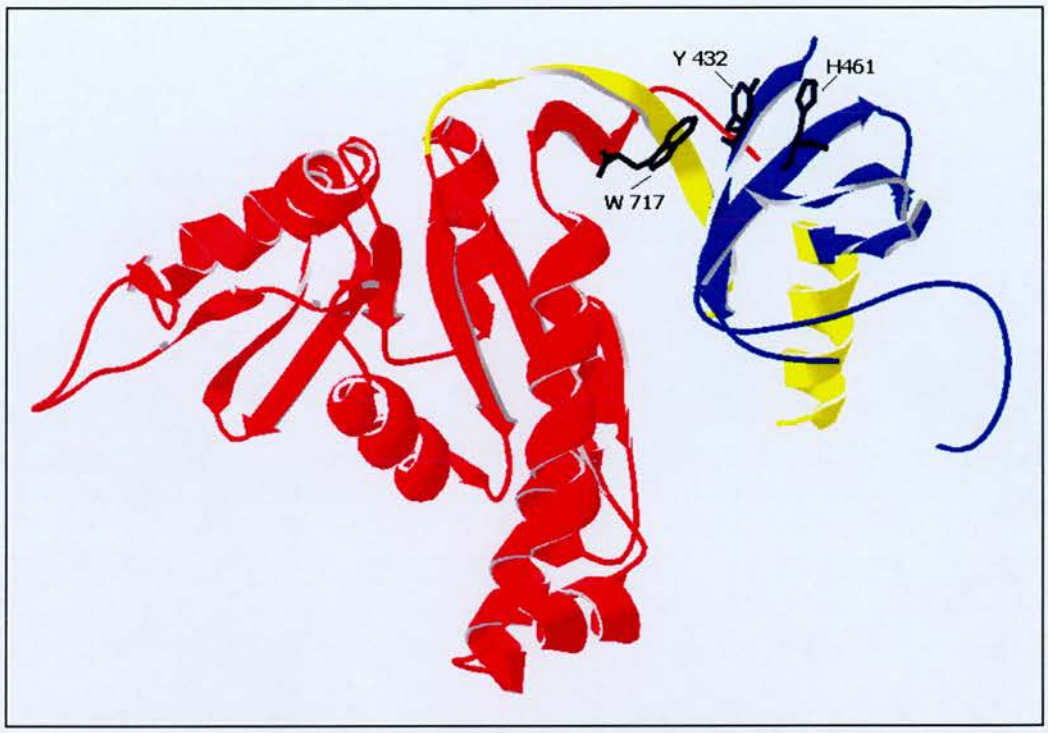


Figure 3.10 – Known partial 3D structure of PSD-95. 3D structure of the SH3 and GK domains of PSD-95 (1JXO.pdb) showing the identified phospho-tyrosine-432 and its neighbouring residues tryptophan-717 and histidine-461, all of which have been implicated in the proposed mechanism of oligomerization of PSD-95 (Tavares, G. A., Panepucci, E. H. & Brunger, A. T. (2001) Mol Cell 8, 1313-25.)

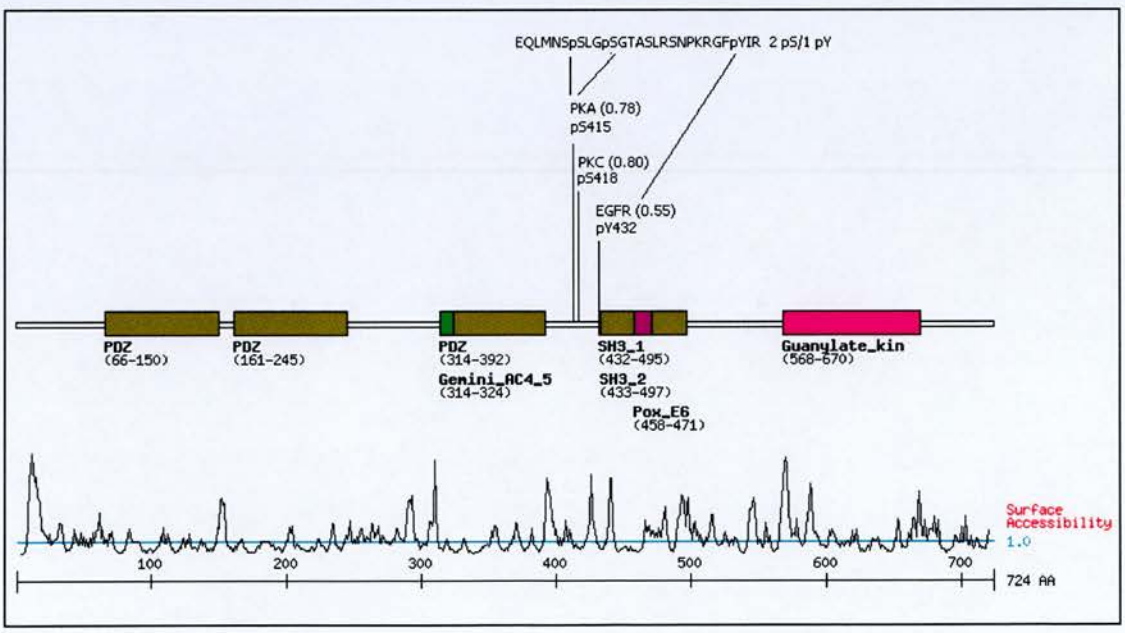


Figure 3.11 – Protein domain structure of PSD-95. The location of 3 phosphorylation sites and the phosphopeptides they were found in are indicated. Scansite was used to predict the most likely kinases to phosphorylate these sites.

Y432 is located in this hydrophobic core mediating this interaction and is itself, interacting with H461. Introduction of a negative charge by phosphorylation of Y432 could disturb this interaction and may have a regulatory role in the formation of an SH3 and GK intermolecular interaction and also affect the interaction of other proteins such as AKAP79 which bind to this region (Colledge et al., 2000) (Figure 3.10). It has been proposed that this intermolecular interaction is a conserved feature of MAGUK (Membrane-associated guanylate kinase) superfamily based on mutational data available for different MAGUK proteins (McGee et al., 2001). These phosphorylated sequences in PSD-95 are also present in two paralogous proteins; SAP97 and SAP102, representing a possible conservation of regulatory functions (Table 3.3).

Mouse	Human	Rat	Name	Phosphopeptide sequences	Residues
Q62108			**PSD-95**	**EQLMNSpSLGpSGTASLRSNPKRGFpYIR**	**409-434**
		P31016	PSD-95	**EQLMNS SLG SGTASLRSNPKRGF YIR**	409-434
	P78352		PSD-95	**EQLMNS SLG SGTASLRSNPKRGF YIR**	452-477
		Q62936	SAP102	**EQMMNS SMS SGSGSLRTSEKRSL YVR**	500-525
P70175			SAP102	**EQMMNS SMS SGSGSLRTSEKRSL YVR**	500-525
	Q92796		SAP102	**EQMMNS SMS SGSGSLRTSEKRSL YVR**	482-507
	Q12959		SAP97	**EQMMNS SIS SGSGSLRTSQKRSL YVR**	562-587
Q8CGN7			SAP97	**EQMMNS SVS SGSGSLRTSQKRSL YVR**	529-554
		Q62696	SAP97	**ETMMNS SVS SGSGSLRTSQKRSL YVR**	561-586

Table 3.3 – Paralogous phosphorylation sites in MAGUK proteins. Two phospho-serines and one phospho-tyrosine residue found just before and at the start of the SH3 domain in PSD-95 are completely conserved in two other paralogous proteins; SAP102 and SAP97

There is a growing body of evidence that scaffold proteins are regulated by phosphorylation (Gardoni et al., 2003), which is consistent with the number of identified scaffold phosphoproteins identified in this study. In fact, we have identified two phosphorylation sites on PKA-RII alpha, which map to a defined 15 amino acid sequence, which has been shown to interact with MAP2 (Figure 3.12) (Scott et al., 1990). This interaction is believed to influence cytoskeletal localisation of the PKA holoenzyme (Scott et al., 1990) and phosphorylation of the binding site could modulate this interaction and ultimately localisation of PKA.

The usefulness of a differential urea extraction strategy is shown by identification of phosphorylation sites in a number of integral membrane or membrane-associated proteins. We have found sites in cell adhesion molecules such as NCAM 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 signalling complex associated with mGluR5 (Farr et al., 2004) (Figure 3.7). We have identified phosphorylation sites in mGluR5 (group I metabotropic glutamate receptor) and IRBIT (membrane associated IP3 (inositol triphosphate receptor) binding protein)) in the urea insoluble synaptosomal fraction and PLC-β in the urea soluble fraction. Activation of mGluR5 results in the hydrolysis of membrane phosphatidyl-inositol bisphosphate (PIP2) to diacylglycerol (DAG), which activates PKC, and inositol triphosphate (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 3 novel phosphorylation sites, and phosphorylation in this region has been postulated to regulate this interaction (Ando et al., 2003). mGluR5, G-proteins and PLC-β form a multiprotein complex with the IP3 receptor and IRBIT via the scaffolding properties of Homer (Tu et al., 1998). This complex seems to facilitate a signalling pathway from mGluR5 to modulation

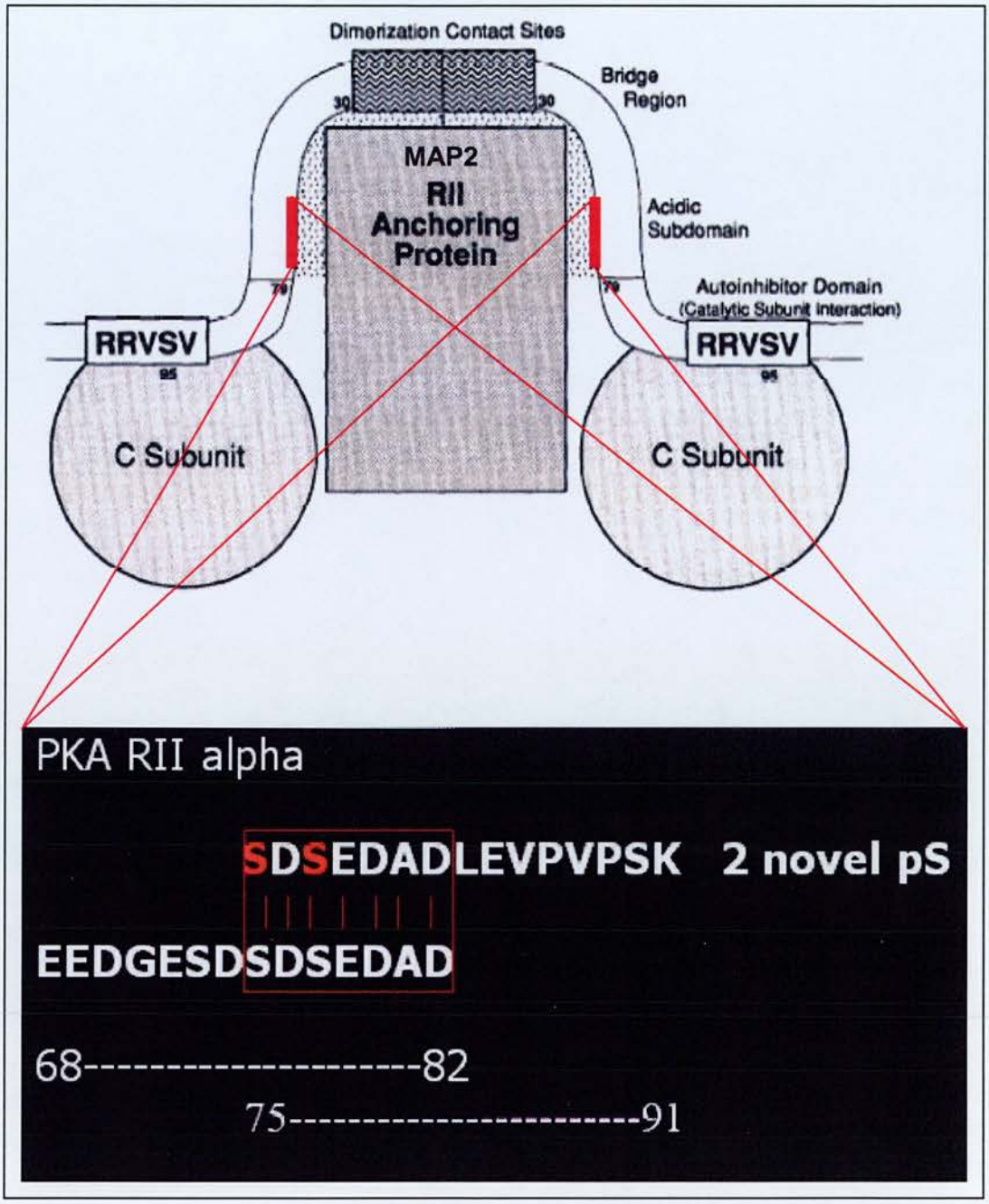


Figure 3.12 – Novel phosphorylation sites in the type two alpha regulatory subunit of cAMP-dependent kinase. Discovery of 2 phosphorylation sites in a region of PKA-R2A which was previously identified as important for binding MAP2. Introduction of negatively charged phosphate groups may regulate this interaction and thereby the localisation of PKA-R2A by MAP2. Top schematic adapted from Scott JD, 1990.

There is a growing body of evidence that scaffold proteins are regulated by phosphorylation (Gardoni et al., 2003), which is consistent with the number of identified scaffold phosphoproteins identified in this study. In fact, we have identified two phosphorylation sites on PKA-RII alpha, which map to a defined 15 amino acid sequence, which has been shown to interact with MAP2 (Figure 3.12) (Scott et al., 1990). This interaction is believed to influence cytoskeletal localisation of the PKA holoenzyme (Scott et al., 1990) and phosphorylation of the binding site could modulate this interaction and ultimately localisation of PKA.

The usefulness of a differential urea extraction strategy is shown by identification of phosphorylation sites in a number of integral membrane or membrane-associated proteins. We have found sites in cell adhesion molecules such as NCAM 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 signalling complex associated with mGluR5 (Farr et al., 2004) (Figure 3.7). We have identified phosphorylation sites in mGluR5 (group I metabotropic glutamate receptor) and IRBIT (membrane associated IP3 (inositol triphosphate receptor) binding protein)) in the urea insoluble synaptosomal fraction and PLC-β in the urea soluble fraction. Activation of mGluR5 results in the hydrolysis of membrane phosphatidyl-inositol bisphosphate (PIP2) to diacylglycerol (DAG), which activates PKC, and inositol triphosphate (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 3 novel phosphorylation sites, and phosphorylation in this region has been postulated to regulate this interaction (Ando et al., 2003). mGluR5, G-proteins and PLC-β form a multiprotein complex with the IP3 receptor and IRBIT via the scaffolding properties of Homer (Tu et al., 1998). This complex seems to facilitate a signalling pathway from mGluR5 to modulation

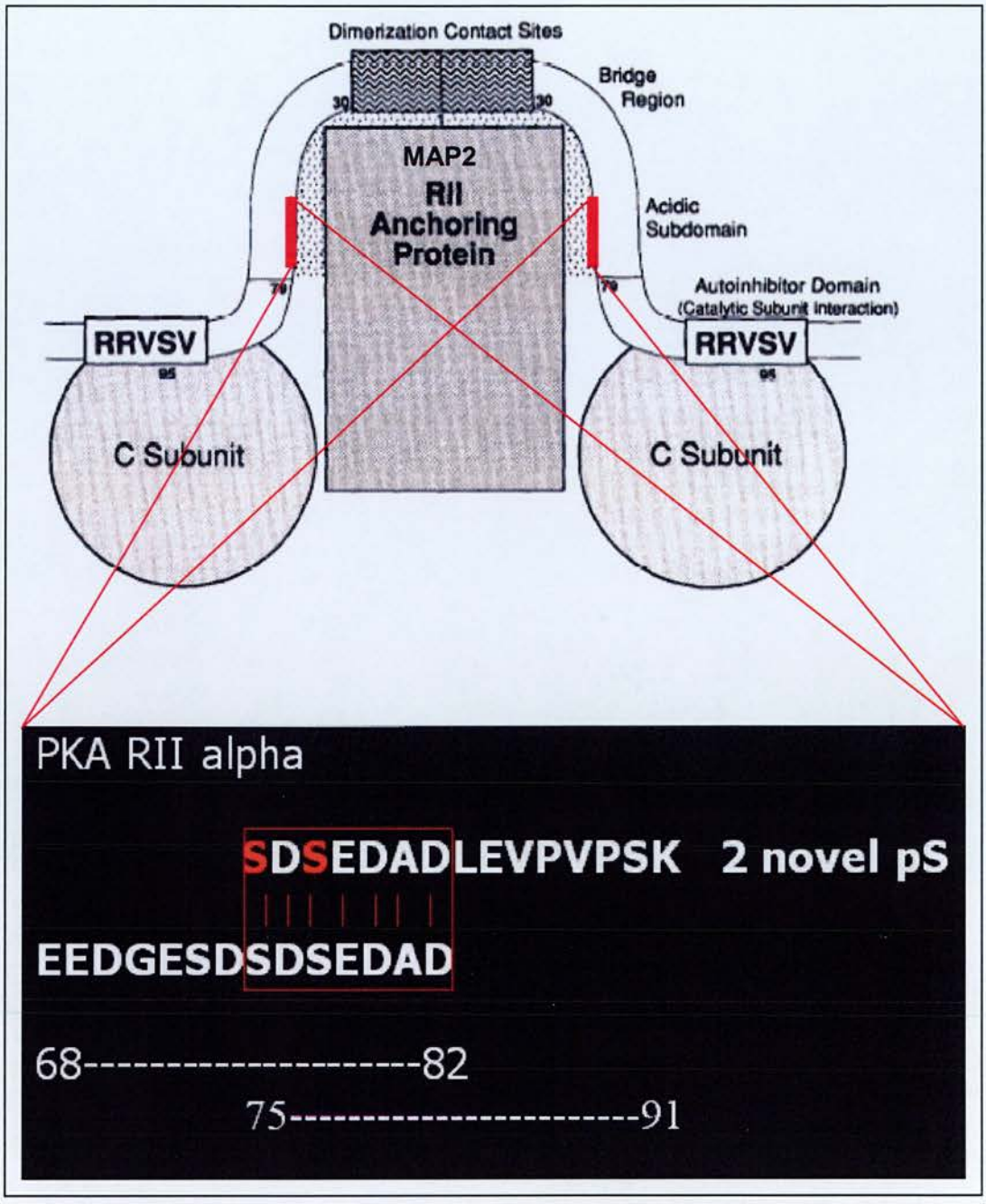
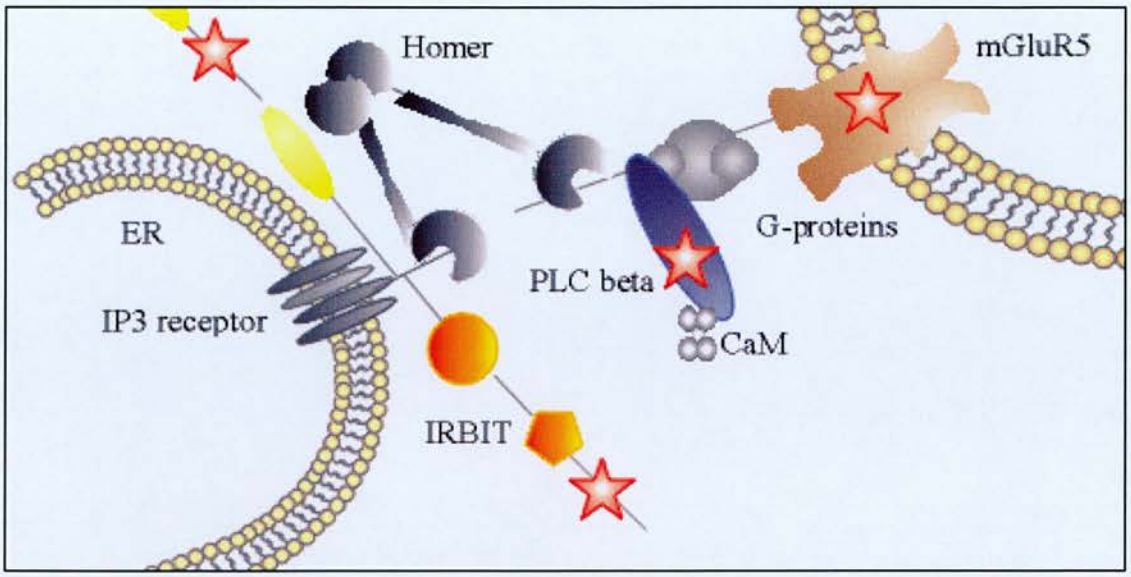


Figure 3.12 – Novel phosphorylation sites in the type two alpha regulatory subunit of cAMP-dependent kinase. Discovery of 2 phosphorylation sites in a region of PKA-R2A which was previously identified as important for binding MAP2. Introduction of negatively charged phosphate groups may regulate this interaction and thereby the localisation of PKA-R2A by MAP2. Top schematic adapted from Scott JD, 1990.

A



B

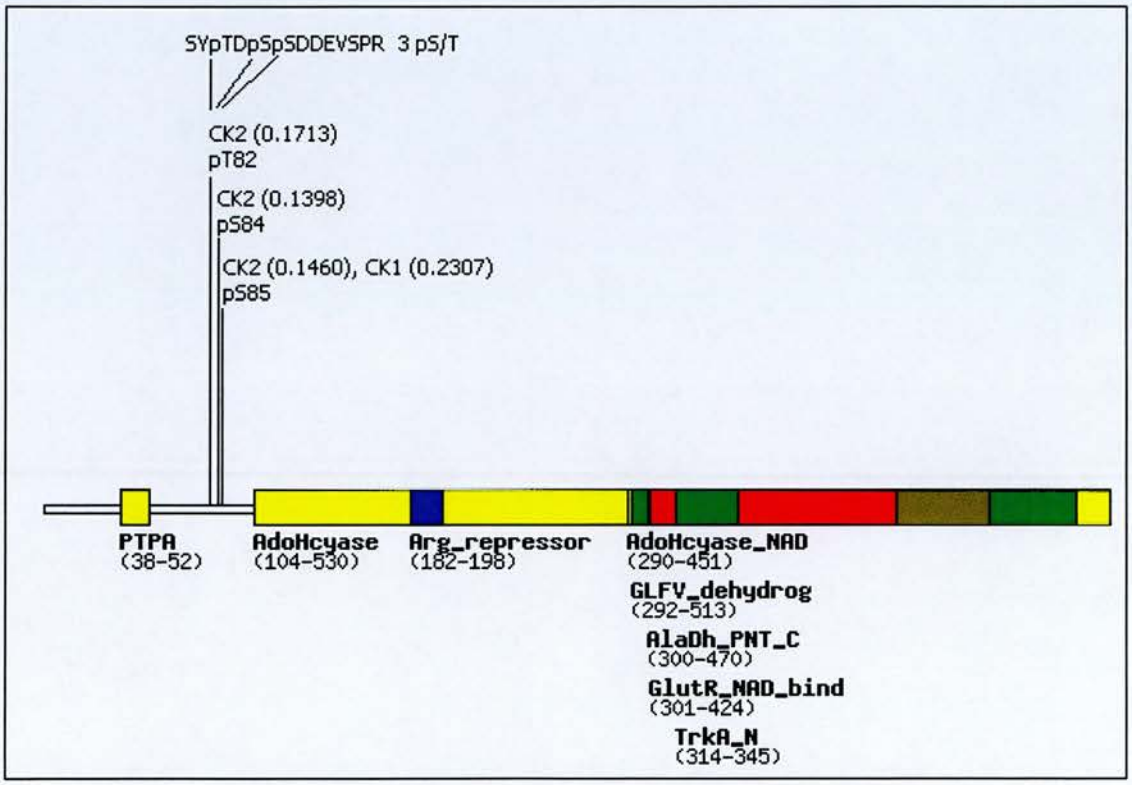


Figure 3.13 – Identification of phosphorylation sites in a membrane to membrane signalling pathway. A. Activation of mGluR5 results in the hydrolysis of membrane phosphatidyl-inositol bisphosphate (PIP2) to diacylglycerol (DAG), which activates PKC, and inositol triphosphate (IP3), which in turn activates the IP3 receptor to release intracellular calcium (33). We have identified novel phosphorylation sites in proteins marked with are red star. **B.** Domain map of IRBIT with N-terminal phosphorylation sites. 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 3 novel phosphorylation sites, and phosphorylation in this region has been postulated to regulate this interaction.

of intracellular calcium, which is known to regulate a myriad of intercellular signalling activities.

mGluR5 has a critical role in NMDAR 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 (Kotecha et al., 2003) and this process requires IP3R-mediated mobilisation 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 (Husi et al., 2000), NR2A (an NMDA receptor subunit) is present in the mGluR5 receptor complex (Farr et al., 2004) and 20 other proteins are common to these complexes. Function 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.

3.5.2 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 (Obenauer et al., 2003). 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 (Nuhse et al., 2004), 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 Immobilised peptide array experiments is consistent with known kinase consensus sequences, (Lizcano et al., 2002), (Rychlewski et al., 2004) and is therefore 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 2 phosphorylation sites screened on peptide arrays was available in the literature. Microtubule-associated protein Tau was shown to be phosphorylated at Ser-493 by Cdk5 (Cruz et al., 2003). This site was identified and was hyper-phosphorylated only in p25 over expressing mice (increased activation of Cdk5) and along with other hyper-phosphorylated residues contributed to neurodegeneration and formation of neurofibrillary tangles (Cruz et al., 2003). 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 Ser-144 by PKA and PKC (Kishimoto et al., 1985), 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 recognise that this in vitro data is not sufficient on its 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 firstly, assign kinases to phosphorylation sites using peptide array screening, which is a 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. As 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 prioritising phosphorylation sites or phosphoproteins for further functional investigation.

3.5.3 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 micro-environment of the substrate. Kinases can be targeted to their substrates via adapter proteins, for example, a MAGUK-AKAP79 complex recruits PKA to the postsynaptic membrane to phosphorylate the AMPA receptor subunit GluR1 (Colledge et al., 2000). 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 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 was superimposed onto this network (Figure 3.13). Six kinases were shown to phosphorylate 9 NRC components on the peptide array. It should be noted that three of these kinases (AKT1, p38 MAPK 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 (Figure 3.13B). PKA is linked to four interacting substrates (PSD-95, PSD-93, SAPAP1 & 2) via AKAP79, a kinase anchor protein which 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

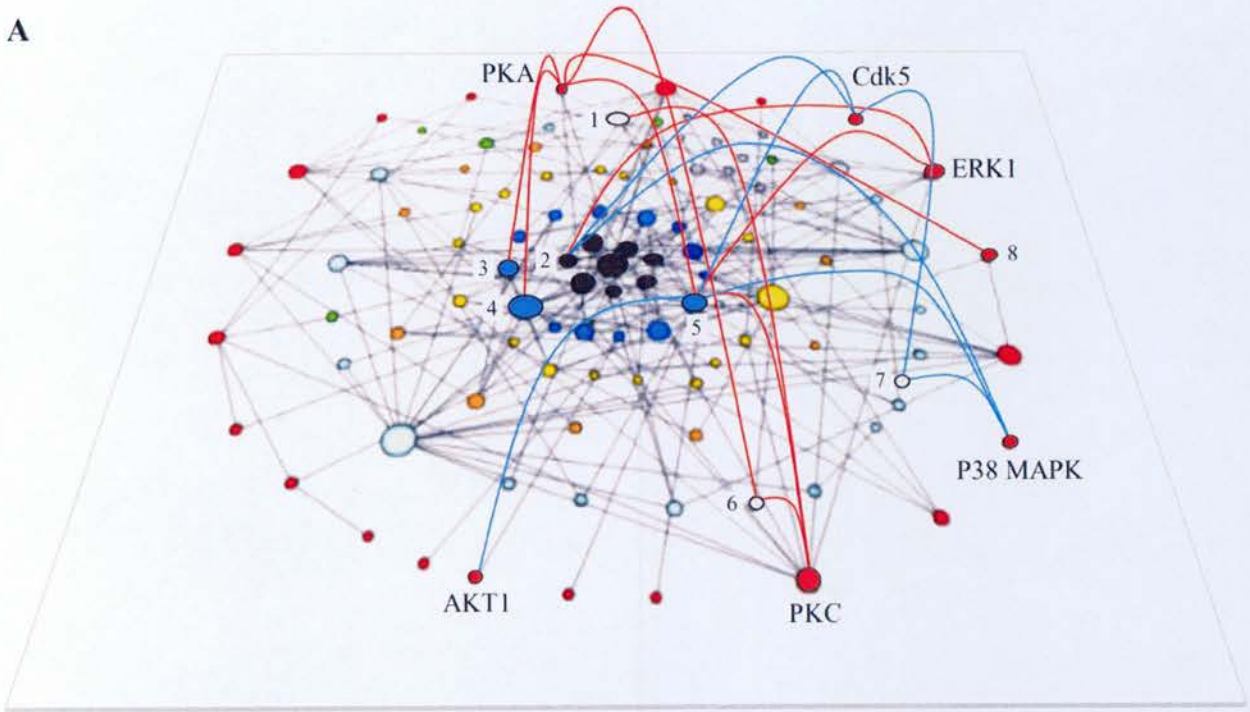


Figure 3.13 - Network illustration of protein-protein interactions and kinase-substrate interactions in the NMDA receptor complex (NRC).
A. Protein-protein interaction data (248 binary interactions from PPID) for 104 NRC proteins was used to construct a protein-protein interaction network. Each node (coloured circle representing a protein) is colour-coded according to its protein class. Kinase-substrate data from MS identified phosphorylation sites and peptide array experiments, for 8 NRC proteins, were superimposed onto the protein-protein interaction network. Red network edges (lines connecting 2 proteins) represent kinases present within the NRC, which phosphorylate NRC substrates and blue network edges represent kinases not presently known to be present within the NRC, which phosphorylate NRC substrates. Substrates are numbered as followed; 1: MAP2, 2: mGlur5, 3: PSD-93/Chapsyn110, 4: PSD-95, 5: SAPAP 1 & 2, 6: Myelin basic protein, 7: Bassoon, 8: CaMK2B. **B.** Illustration of path-lengths from PKA to its NRC substrates (PSD-95, PSD-93/Chapsyn110 and SAPAP 1 & 2). These PKA substrates are important synaptic scaffolders, which are closely associated with each other and their proximity to this kinase within the NRC would imply a physiological context for their phosphorylation.

therefore should be prioritised 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 synaptic signalling pathways

3.5.4 Synaptic plasticity and disease

In addition to identifying and characterising 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 (Supplementary Table 4). It is generally accepted that phosphorylation is important in the regulation of synaptic plasticity, a process, which is believed to be involved in learning and memory (Tokuda and Hatase, 1998), drug addiction (Yao et al., 2004) and pain (Garry et al., 2003). Functional studies have shown that 7 phosphoproteins we identified are involved in synaptic plasticity and perturbation of 5 in rodents, show impairments in learning and memory. Also, 7 out of the top 11 kinases predicted to phosphorylate our total phosphoprotein dataset are known to be intimately involved in synaptic plasticity. We have identified 8 phosphoproteins, which have been linked to Schizophrenia and other mental disorders. Finally, we have detected at least 7 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 (Cruz et al., 2003).

3.5.5 Conclusions

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 characterised, and in particular it can be used to derive a picture of the global organisation of the synapse phosphoproteome. The establishment of phosphorylation maps will provide the basis for numerous functional studies, which will aid the understanding of complex signalling 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.

CHAPTER 4

PHOSPHOPROTEOMIC ANALYSIS OF THE MOUSE FOREBRAIN: INVESTIGATION OF THE RELATIONSHIP BETWEEN INTRINSIC PROTEIN DISORDER AND PHOSPHORYLATION IN THE BRAIN.

4. Phosphoproteomic analysis of the mouse forebrain: Investigation of the relationship between intrinsic protein disorder and phosphorylation in the brain.

4.1 ABSTRACT

We have analysed the mouse forebrain cytosolic phosphoproteome by using a tandem (protein and peptide) immobilised gallium affinity chromatography purification strategy and analysis by tandem mass spectrometry. This resulted in identification of 485 phosphorylation events corresponding to 321 unique phosphorylation sites on 67 cytosolic proteins. We found that in a single 4-hour reversed-phase LC run with a PepMap column we could identify 401 phosphorylation events on 173 phosphopeptides, highlighting the efficiency of our approach. Intrinsic protein disorder, which is characterised by lack of any regular structure and enrichment of specific disorder promoting amino acids, is becoming increasingly recognised as important for protein function. We have investigated the role of protein disorder and its relationship with protein phosphorylation in the brain. We found that the cytosolic phosphoproteome dataset presented here is rich in sequence disorder and contains many putatively disordered proteins. The majority of these have RNA or tubulin binding functions and the combination of disorder and phosphorylation is likely to be essential for their function. We have observed that the majority of phosphorylation sites occur in regions of sequence disorder and that disordered proteins may produce more tryptic fragments than ordered regions, with implications for phosphoproteomic analyses. We identified 37 phosphorylation sites on nine protein kinases, the majority of which are novel and are not located in kinase domains indicating secondary regulatory functions. Finally, in our global study of cytosolic phosphorylation we have characterised a number of phosphoproteins with disease-associations, which may have related roles in the pathogenesis of neurodegenerative disease.

4.2 INTRODUCTION

Protein phosphorylation is an essential regulator of cellular function and the complement of phosphorylation sites in a proteome (phosphoproteome) is dynamic and highly complex. Recent phosphoproteomic analyses of numerous biological samples have together identified thousands of phosphorylation sites (Loyet et al., 2005), indicating that the phosphoproteome is much larger and complex than original estimates. Annotation of large numbers of phosphorylation sites is a difficult task as the amount of associated biological information is usually low in large-scale experiments. Phosphorylation sites can be mapped to functional protein domains, assigned putative kinases by prediction programs such as Scansite (Obenauer et al., 2003), assigned as putative binding sites, for example for 14-3-3 proteins, by using the ELM server (Puntervoll et al., 2003). Phosphorylation sites can also be assigned putative functions by retrospective annotation from the literature, for example, deletion studies that map a binding, catalytic or regulatory function to a portion of a protein sequence. In the absence of direct functional data relating to datasets of phosphorylation sites, such sequence- or literature-based annotation is important to identify potentially important phosphorylation sites for more focused study.

As phosphorylation can affect the conformation of a protein (Johansen and Ingebritsen, 1987), it occurs on accessible regions in the three dimensional structure (Ridsdale et al., 1997) and is thought to occur mostly in flexible regions in a protein structure (Shen et al., 2005) annotation of phosphorylation sites with structural information may be useful. The three-dimensional structure of a protein is encoded in its amino acid sequence and can take a number of forms along a continuum of increasing content of stable tertiary structure. Figure 4.1 shows examples of types of protein structure from completely unstructured proteins, molten globules, linked-folded domains to mostly folded proteins. The extent of folding upon target binding reduces with

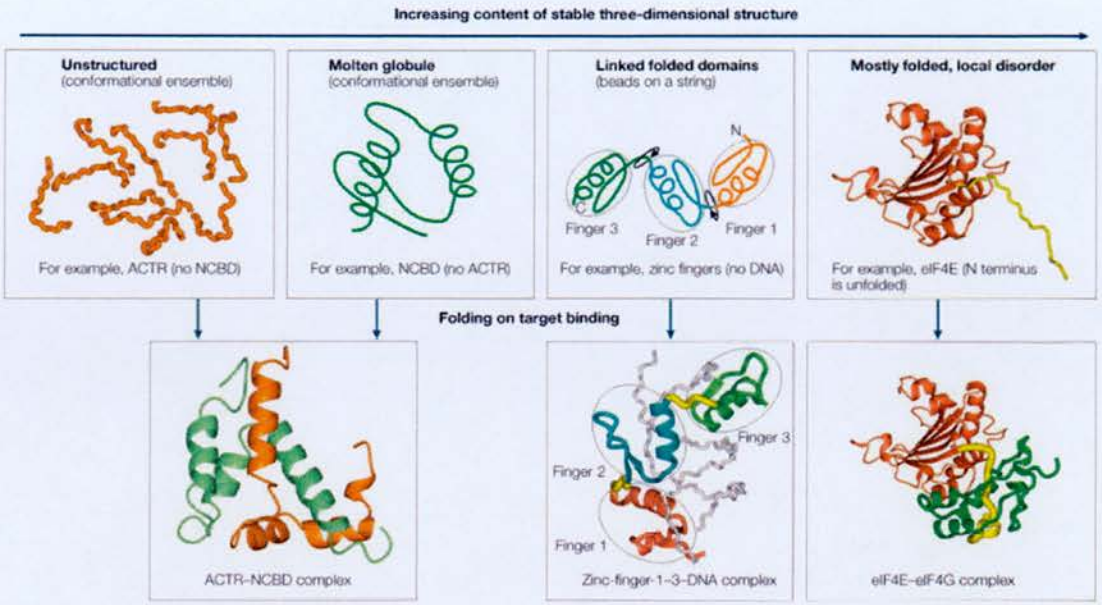


Figure 4.1 – Types of protein structure. Examples of the types of protein structure from completely unstructured proteins, molten globules, linked-folded domains to mostly folded proteins. The extent of folding upon target binding reduces with increasing content of stable three-dimensional structure. Adapted from "Intrinsically unstructured proteins and their functions", Dyson HJ, Wright PE. Nat Rev Mol Cell Biol. 2005 Mar;6(3):197-208

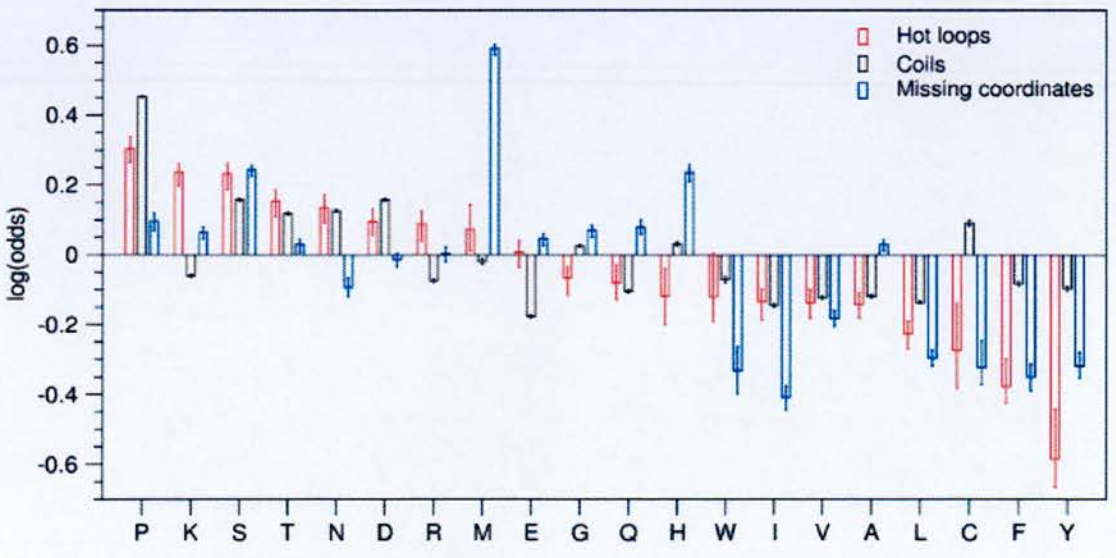


Figure 4.2 - Propensities for amino acids to be disordered.
Three different definitions of protein sequence disorder were used to define the propensities of amino acids to be in disordered sequence. Adapted from "Protein disorder prediction: implications for structural proteomics". Linding R et al. Structure. 2003 Nov;11(11):1453-9.

increasing content of stable three-dimensional structure. The of structure or lack of protein structure can be observed indirectly by X-ray crystallography, NMR-, Raman and CD-spectroscopy and by hydrodynamic measurements (Linding et al., 2003). The extent of disorder in a protein sequence will affect the likelihood of protein crystallisation and the success of subsequent X-ray structure determination. Therefore, disorder prediction algorithms have been developed to improve the success of structural proteomic studies by providing a way of screening proteins that are not likely to be amenable to such analysis. One such disorder prediction server is DisEMBL, which uses three different definitions of proteins disorder; presence of Loops/coils, Hot loops which are loops with a high degree of mobility as determined from temperature factors (B factors) and finally missing co-ordinates in X-ray structure that are defined as remark465 entries in PDB (Linding et al., 2003). The hot loop definition is probably the most reliable as an indication of protein disorder as the missing co-ordinates in PDB definition can be misleading as missing co-ordinates can be caused by protein truncations which result in abnormally disordered sequence.

Figure 4.2 shows the amino acid propensities determined from the definitions of protein disorder. It can be seen that generally hydrophobic residues promote order according to all three definitions and disorder-promoting residues include lysine, proline, arginine, serine, threonine and methionine. As serine and threonine are enriched in disordered sequence and it is known that phosphorylation sites tend to be located in flexible regions of proteins, it is likely that intrinsic protein disorder and protein phosphorylation are intimately linked. In fact, DISPHOS, which is a phosphorylation site predictor, uses protein disorder information to improve discrimination between phosphorylated and non-phosphorylated sites.

In recent years, mass spectrometry has been applied to biological samples to identify phosphorylation sites, usually in combination with a phospho-enrichment step. We have developed a tandem Immobilised gallium-affinity (IMAC) protocol in

which phosphoproteins are specifically purified from a biological sample, tryptically digested and the resultant peptide mixture applied to another IMAC column to specifically enrich for phosphopeptides (Collins et al., 2005), (Collins et al., 2005). This tandem-affinity protocol and single protein and peptide IMAC protocols were used to identify phosphorylation sites in urea-soluble synaptic proteins. In, total 650 phosphorylation events, corresponding to 331 unique phosphorylation sites were identified, with the largest portion (308 phosphorylation events) identified in the tandem-IMAC experiment.

Here we describe application of this tandem-IMAC approach to the mouse forebrain cytosolic phosphoproteome. In a single four-hour LC run we could identify up to 401 phosphorylation events on 173 phosphopeptides. In total, we identified 485 phosphorylation events corresponding to 321 unique phosphorylation sites on 67 phosphoproteins. We have investigated the relationship between protein disorder, proteolytic cleavage and phosphorylation in the cytosolic phosphoproteome. The cytosolic extraction procedure was enriched for very soluble proteins i.e. hydrophilic proteins, which tend to be disordered. Phosphoproteomic analysis of this fraction identified at least 16 putative intrinsically disordered phosphoproteins, which have mainly RNA- or Tubulin-binding function. The majority of identified phosphorylation sites in the total dataset are located in regions of sequence disorder indicating a function associated with phosphorylation-dependent binding or regulation of conformational state. We suggest that protein phosphorylation is heavily employed in intrinsically disordered regions to regulate the increased number protein interactions (Wright and Dyson, 1999) mediated by intrinsically disordered proteins. In addition, the use of hyperphosphorylation as a means of creating local negative charge to regulate, for example, protein-RNA/DNA backbone interactions, may explain the high stoichiometry of phosphorylation observed for some proteins. Finally, we have characterised a number of phosphoproteins with disease-associations, which may have related roles in the pathogenesis of neurodegenerative disease.

4.3 MATERIALS AND METHODS

4.3.1 Protein extraction

Mouse forebrains were rapidly dissected, frozen immediately in liquid nitrogen and stored at -80°C. Cool all buffers and the Homogeniser to 4°C. 10 forbrains were Homogenised with 25 stokes in a dounce Homogeniser in homogenisation buffer (50 mM tris pH 9, 50 mM sodium fluoride, 20 μM zinc chloride, 1 mM sodium ortho-vanadate, 0.5 mM PMSF, 2 μg/ml aprotinin, 2 μg/ml lepteptin) at a ratio of 0.379g tissue/7ml buffer. The extract was subjected to centrifugation at 50,000 x g for 30 min and was concentrated in a Vivaspin 10 kDa MWCO and stored at -80°C.

4.3.2 Protein IMAC of forebrain cytosolic fraction

Fast-flow chelating sepharose with iminodiacetic acid (IDA) (Amersham Biosciences) chelating groups was charged with GaCl3. Cytosolic protein (50 mg) was brought to 6 M urea and incubated with 2 ml of the metal charged resin with mixing for 1 hour at room temperature (RT). The unbound protein was washed with buffer A (6 M urea/50 mM tris-acetate) to baseline and the phosphoproteins were specifically eluted with buffer B (6 M urea, 50 mM tris-acetate, 100 mM EDTA, 100 mM EGTA). Two of these purifications were carried out, eluted phosphoproteins pooled, concentrated and washed with buffer B in a Vivaspin 6 PES membrane spin column (Vivascience), resulting in a final yield of 3 mg of purified phosphoproteins.

4.3.3 Double IMAC of forebrain cytosolic fraction

3 mg of the protein IMAC purified sample was digested with modified porcine trypsin (Promega) in a ratio of 1:20 in digestion buffer (pH 8) (1 M urea and 25 mM NH4HCO3) at 37°C for 4 hours. The resultant digest was desalted and dried and methyl esterification was performed with 2 M methanolic HCl (10). Self Pack POROS® 20 MC Media (Applied Biosystems) for phosphopeptide purification was

charged with GaCl3 as described above for the IDA/NTA resins. Peptide digests were reconstituted in loading buffer (equal volumes of acetonitrile, methanol and water pH 2.5-3). 1 ml of this peptide mixture was incubated with 200 µl of POROS-Ga slurry for 1 hour at RT. The resin was then loaded into a spin-column and washed with 10 volumes of buffer E. Phosphopeptides were eluted with 2 x 100 µl of 200 mM Na2HPO4.

4.3.4 On-line nano LC-MS/MS

Mass spectrometry was performed by Lu Yu (The Wellcome Trust Sanger Institute)

A nanoflow HPLC system, Ultimate™ (LC Packings) was coupled to a, Q-Tof Ultima (Waters/Micromass). Tryptic peptides from the phosphoprotein digest were loaded in 0.1% aqueous formic acid (FA) and desalted on PepMap C18 trapping cartridge (LC Packings) or a BetaMax Neutral (Thermo Hypersil-Keystone) was used to trap the IMAC enriched phosphopeptides in 0.5% aqueous FA. Peptides on the trap were back-flushed to, and separated on the analytical column (PepMap C18, 75 µm id x 15 cm, LC Packings). The Q-Tof Ultima was operated in automated Data Dependent Acquisition (DDA) mode. Each cycle had a 1 sec 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 sec. The collision energy in MS/MS was varied according to the m/z and the charge state of the precursor ion.

4.3.5 Data Analysis

4.3.5.1 Database searching

Raw data was processed to give a peak list file and submitted to a local Mascot V2.0 (Matrix Science) server for iterative searching on a custom, non-identical, combined human and mouse IPI database (EBI). Assignment of phosphorylation sites was verified manually with the aid of PEAK Studio (Bioinformatics Solutions) software.

4.3.5.2 Sequence-based analysis

ProtParam (http://us.expasy.org/cgi-bin/protparam) was used to compute pI values and GRAVY (Grand average of hydropathicity index) scores for all phosphopeptides. All phosphoproteins proteins detected in this study were classified according to Swiss-Prot keywords (http://ca.expasy.org/sprot/). Known phosphorylation sites in Supplementary Table 2 were annotated by extensive literature mining of PubMed. Scansite (http://scansite.mit.edu/) and was used for predicting the most likely kinases responsible for the phosphorylation sites characterised in this study. In addition for ambiguously defined phosphorylation sites Scansite was used to predict the most likely site of phosphorylation when a number of possibilities were present on a phosphopeptide. Scansite was also used to check the presence of phosphorylation sites in Pfam protein domains and to predict if phosphorylation sites were localised in phospho-dependent interaction domains (Supplementary Table 2). DisEMBL, Intrinsic Protein Disorder Prediction 1.4 (http://dis.embl.de/) was used to predict phosphoprotein sequence disorder and in these predictions, the "hotloop" definition was used. PeptideCutter (http://us.expasy.org/tools/peptidecutter/) and Phospho-pepsort7 (http://acrux.igh.cnrs.fr/gs/phospepsort7.html) were used for *in silico* tryptic digestion of phosphoproteins. BCM search launcher in combination with BOXSHADE 3.21 (http://www.ch.embnet.org/software/BOX_form.html) were used for construction of multiple sequence alignments. Members of protein kinases families were downloaded from The Protein Kinase Resource (http://www.kinasenet.org/pkr/). Three-dimensional protein structure data was visualised with the Deep View Swiss-PdbViewer 3.7 (http://swissmodel.expasy.org/spdbv/).

4.4 RESULTS

4.4.1 Distribution of characterised phosphorylation sites

The phosphoproteomic strategy employed here to study cytosolic phosphoproteins in the mouse forebrain resulted in the identification of 485 phosphorylation events on 218 phosphopeptides, corresponding to 321 unique phosphorylation sites (Table 4.1, Figure 4.3). Eighty-six percent of these phosphorylation events were unambiguously assigned exact sites in peptide sequences with the remaining 44 (14%) phosphorylation events mapped to a few possible serine, threonine or tyrosine residues in peptide sequences. The distribution of phosphorylation sites was 81, 9.5 and 0.6% for phosphoserine, phosphothreonine and phosphotyrosine respectively. These numbers include some ambiguously assigned sites for which the type of phosphorylation site could be deduced, for example, the only possibility of phosphorylation in a peptide sequence was on serine residues. Of the ambiguously assigned phosphorylation sites with ambiguous phosphorylation type, 24 were phosphoserine or phosphothreonine and 4 were phosphoserine, phosphothreonine or phosphotyrosine.

4.4.2 Analytical comparison

We performed two parallel analyses of phosphopeptides derived from a tandem protein and peptide gallium-affinity purification from forebrain cytosolic extract. Equal quantities of these peptides were applied to different on-line reversed-phase (RV) chromatographic separations. In the first experiment a PepMap (details) column was used as the analytical column and in the second a BetaMax neutral column was used. Reverse-phase separation is essential for on-line LC-MS/MS and the type of RV column influences the efficiency of peptide separation and the population of peptides that are retained and eluted from the RV column. Two identical LC-MS/MS experiments with RV separations of 4 hours were performed to

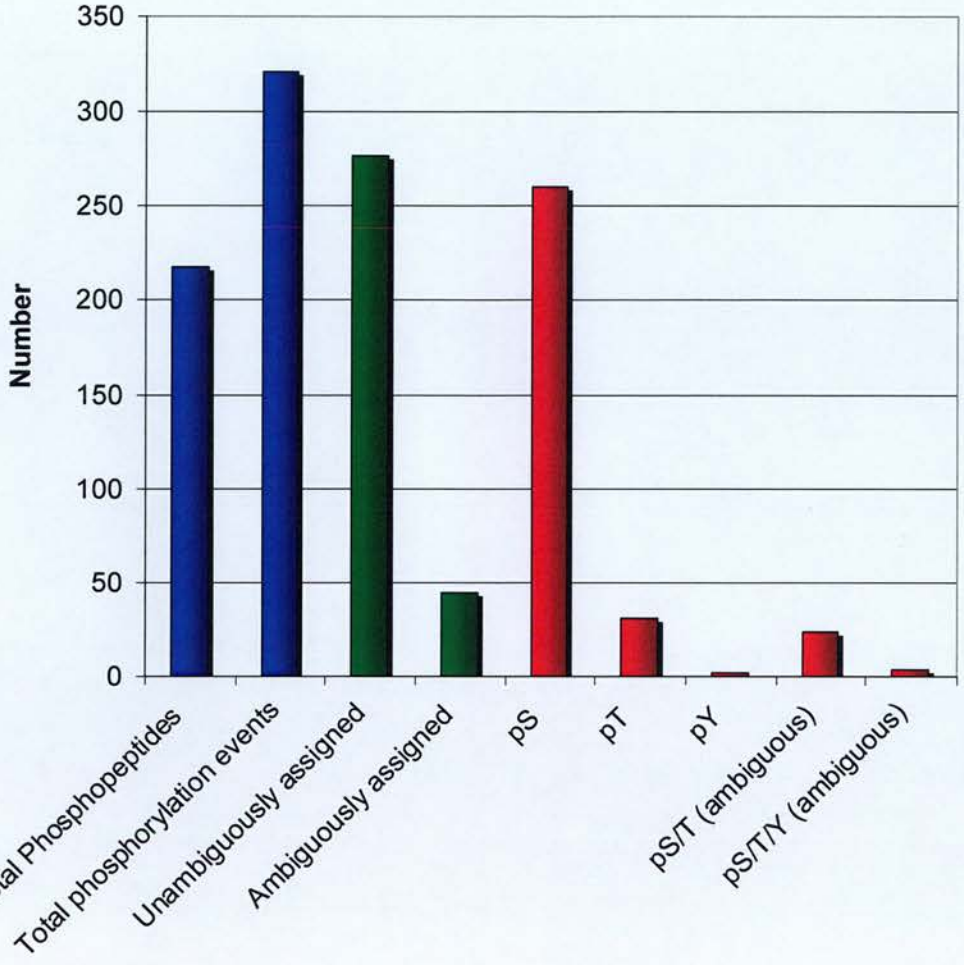


Figure 4.3 Distribution of phosphorylation sites in the cytosolic phosphoproteome. Numbers of phosphopeptides, phosphorylation evens and a breakdown of the types of phosphorylation sites found are shown.

assess the selectivity of these two columns in terms of the phosphopeptides detected in the mass spectrometer.

Phosphopeptides and particularly multi-phosphorylated peptides are more hydrophilic than their un-phosphorylated counterparts. This results in earlier elution of phosphopeptides in a chromatography run and can prevent binding of very hydrophilic phosphopeptides to the RV column and are thus not analysed in the mass spectrometer.

A

RV Column	Number of phosphopeptides	Phosphorylation events	Phosphates per phosphopeptide
PM	173	401	2.31
BM	110	224	2.03

B

	GRAVY	pI	P/phosphopeptide
PM + BM	-0.815	6.89	2.1
PM only	-0.943	7.75	2.55
BM only	-0.762	7.4	1.92

Table 4.1 – Performance of PepMap (PM) and BetaMax neutral (BM) columns

A. The numbers of phosphopeptides, phosphorylation events and stoichiometry of phosphorylation in the datasets from the PM and BM 4 hour gradients are shown.

B. To evaluate the differences in peptide binding between these to column types, GRAVY (Grand average of hydropathicity index) scores, isoelectric points and stoichiometry of phosphorylation of phosphopeptides detected in both analyses (PM + BM), of phosphopeptides specific to the PM analysis and of phosphopeptides specific to the BM analysis are shown.

These LC-MS/MS analyses permitted characterisation of 199 phosphopeptides with 451 phosphorylation events. The use of a PepMap column appears to allow identification of more phosphopeptides than the BetaMax column (173 versus 110, respectively (Table 4.1A). In addition the number of phosphorylation events detected in the PepMap experiment was also greater (401 versus 224, respectively) with a peptide stoichiometry of phosphorylation of 2.31 compared to 2.03 for the BetaMax experiment. These differences likely reflect the underlying binding characteristics of the two RV columns.

Firstly, phosphopeptide distribution in terms of stoichiometry of phosphorylation is different (Figure 4.4A). The PepMap column exhibits a preference for multiply phosphorylated peptides as opposed to the BetaMax column which is more efficient at binding mono and di-phosphorylated peptides. Secondly, the hydrophobicity score and isoelectric point calculated for the peptide sequences highlight a preference for hydrophilic peptides by the PepMap column. It can be seen in Figure 4.4B that peptides specific to the PepMap experiment are generally more hydrophilic than those found in the BetaMax experiment. On average the GRAVY score for PepMap-specific phosphopeptides is -0.943 compared to BetaMax-specific phosphopeptides with an average GRAVY score of -0.762 (Table 4.1B).

The combination of higher gravy scores (hydrophilic) and preference for multi-phosphorylated peptides (phosphate moiety increases the hydrophilicity of a peptide) bound by the PepMap column both indicate that it better suited for capture of hydrophilic peptides. This data would indicate that the use of the PepMap column is preferable for phosphoproteomic analysis as it enabled binding of hydrophilic phosphopeptides that were not bound by the BetaMax column and allowed characterisation of a greater number of phosphopeptides.

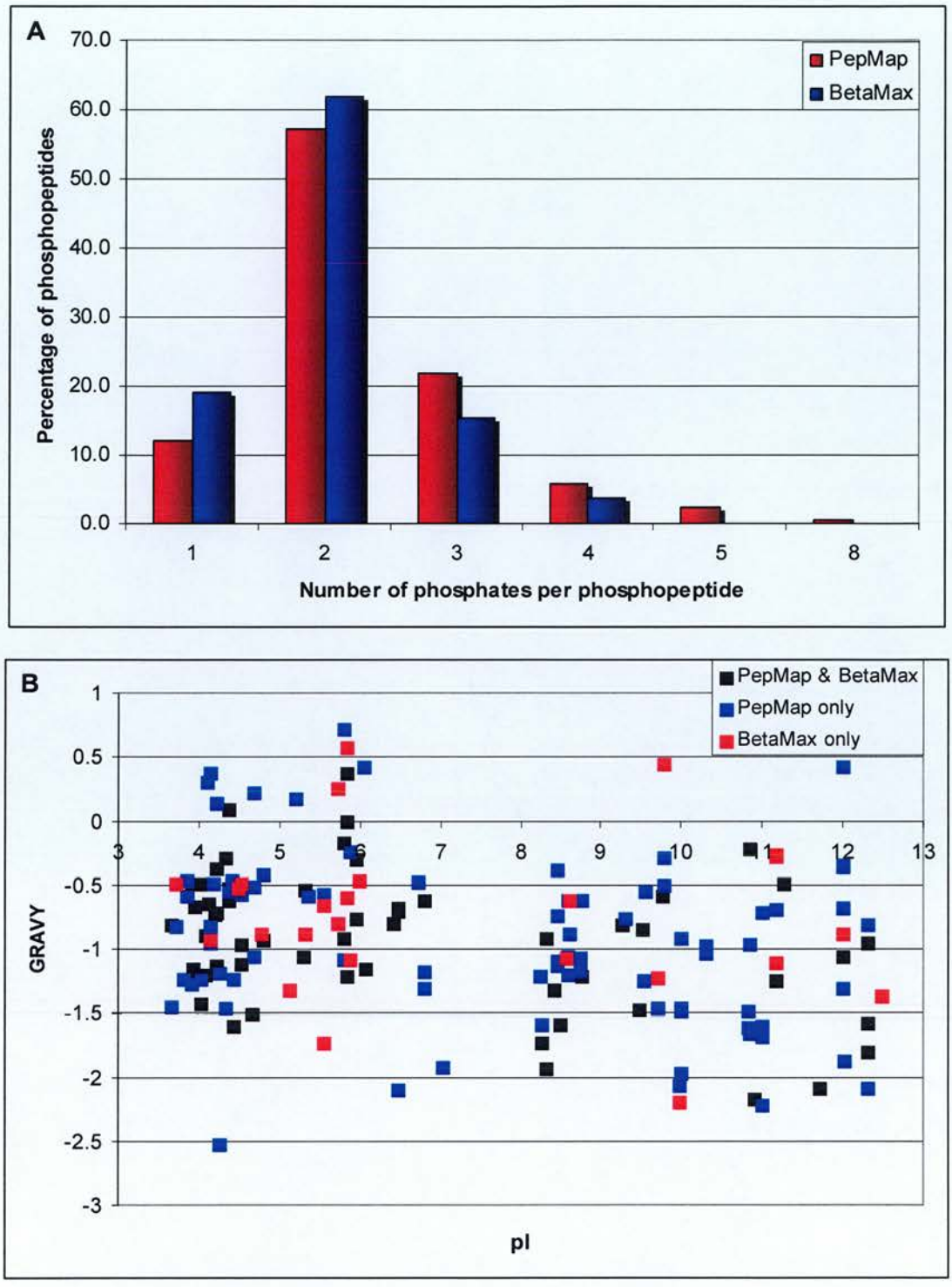


Figure 4.4 – Comparison of PepMap and BetaMax neutral reversed phase columns for analytical separation of phosphopeptides

A. Bar chart comparing the stoichiometry of phosphorylation of phosphopeptides characterized in the PepMap and BetaMax LC-MS/MS experiments.

B. Scatter-plot of GRAVY versus pI of phosphopeptides found in both PepMap and BetaMax experiments and those specific to each experiment.

In addition, to these two 4 hour LC-MS/MS runs that compared the RV columns, a number of 2 hour LC-MS/MS experiments were performed with the PepMap column. The first PepMap 2 hour run resulted in identification of 23 phosphopeptides, of which 18 were found in the 4 hour PepMap run with an additional 5 phosphopeptides not detected previously. The same sample was analysed again but with further tryptic digestion to reduce missed cleavages. This allowed identification of 28 phosphopeptides of which, 16 were found in the previous 2 hour LC-MS/MS experiment, 7 were found in the 4 hour PepMap experiment and 5 were not previously found. Next the endopeptidase Asp-N was used to cleave the same phosphopeptide sample further (cleaves N-terminal to aspartate) and resulted in identification of 17 phosphopeptides, all of which were identified previously in one or more of the 4 or 2 hour LC-MS/MS analyses.

4.4.3 Classification of cytosolic phosphoproteins

The combination of LC-MS/MS approaches described in the previous section resulted in the identification of 321 unique phosphorylation sites on 67 cytosolic phosphoproteins. These proteins were classified according to annotated information in UniProt as well as from literature mining. The largest class of phosphoproteins have no known function and the fact that they represent 20% of the observable cytosolic phosphoproteome in our experiments highlights the lack of comprehensive functional annotation of the mouse proteome. The next most significant class of proteins are involved in splicing and /or are RNA binding proteins and represent 19% of the dataset. Other classes observed include Cytoskeletal proteins (17%), Kinases (13%), Scaffolding proteins (6%), Signalling proteins (6%), Development/neuronal differentiation (4%), DNA binding proteins (4%), Synaptic vesicle proteins (4%) and G-protein modulators (3%). This broad range of protein classes is not surprising as the cytosolic purification protocol is based on protein solubility rather than protein localisation within the cell.

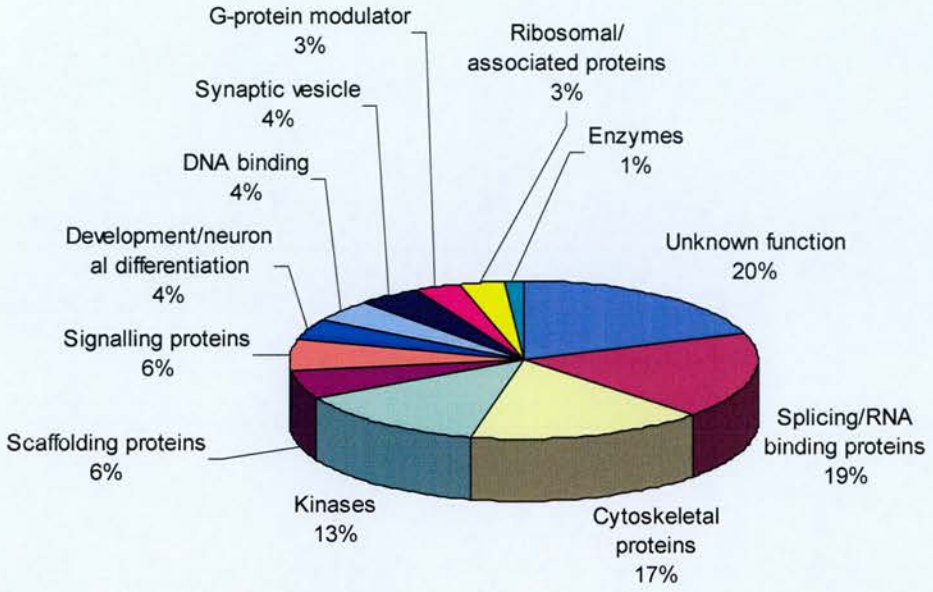


Figure 4.5 – Classes of phosphoproteins characterised from forebrain cytosolic fraction

Protein Class	Accession Number	Protein Name	Number of Phospho Sites
Kinases			
	Q5RJI5	Serine/threonine kinase SADB (BRSK1)	12
	Q6DMN7	BR serine/threonine-protein kinase 2 (BRSK2)	4
	Q68EG2	CaMKII beta	2
	Q5SQZ3	CaMKII gamma	3
	Q9WV60	Glycogen synthase kinase-3 beta	1
	Q9JLM8	Serine/threonine-protein kinase DCAMKL1	7
	Q61136	Serine/threonine-protein kinase PRP4 homolog	4
	Q9R1L5	Syntrophin-associated serine-threonine protein k	2
	Q9UKE5	TRAF2 and NCK interacting kinase	2
Cytoskeletal proteins			
	Q9QYB8	Beta adducin	6
	Q9WV69	Dematin	3
	P78559	Microtubule-associated protein 1A	13
	P14873	Microtubule-associated protein 1B	11
	P20357	Microtubule-associated protein 2	35
	P27546	Microtubule-associated protein 4	1
	P10637	Microtubule-associated protein Tau	19
	Q9QYX7	Piccolo protein	3
	Q9QWI6	SNAP-25-interacting protein (SNIP)	1
	O88935	Synapsin I	4
	Q8CC35	Synaptopodin	1
Scaffolding proteins			
	Q6P9K8	Caskin-1	2
	Q61120	N-SHC	2
	Q9WV48	Shank1	4
	Q80TS1	nArgBP2	2
Signalling proteins			
	Q68FF6	G protein-coupled receptor kinase-interactor 1	6
	Q5RIT9	Similar to PITPNM family member 3	2
	Q8VCD2	SR-related CTD associated factor 6	3
	Q7TQD2	Tubulin polymerization-promoting protein	2
Splicing/RNA binding proteins			
	Q9JIX8	Apoptotic chromatin condensation inducer in the	2
	P62996	Arginine/serine-rich splicing factor 10 (Sfrs10)	7
	O70305	Ataxin-2 (Spinocerebellar ataxia type 2 homolog	2
	Q13427	Cyclophilin G	2
	Q9CYC6	mRNA decapping enzyme 2	3
	Q9R0U0	Neural specific sr protein NSSR 1	3
	O70495	Plenty-of-prolines-101 (Srrm1)	19
	Q8R4E7	Pur-gamma	3
	Q96MU7	Putative splicing factor YT521	1
	Q8BTI8	RNA binding protein homolog (Srrm2)	33
	Q8VE97	Splicing factor, arginine/serine-rich 4	3
	Q99J36	THUMP domain containing protein 1	2
Development/neuronal differentiation			
	P97427	Dihydropyrimidinase related protein-1	4
	O08553	Dihydropyrimidinase related protein-2	7
	Q9QZS3	Numb	3

Protein Class	Accession Number	Protein Name	Number of Phospho Sites
G-protein modulator			
	Q96D85	IQSEC1	2
	XP_290799	Rho GTPase activating protein 23	2
Enzymes			
	Q9D7P6	Nifun	2
DNA binding			
	Q8K019	Bcl-2-associated transcription factor 1 [Bclaf1]	5
	O94894	Prp5-like DEAD-box protein	2
	Q9CTM9	RNA Helicase-related protein homolog Ddx42	2
Synaptic vesicle			
	P17426	Adapter-related protein complex 2 alpha 1 subui	3
	Q5J8K4	Rab3 effector, RIM-1	9
	Q9EQZ7	Rab3-interacting molecule 2 (RIM-2)	2
Ribosomal/associated proteins			
	P62754	40S ribosomal protein S6 (RPS6)	1
	Q7Z417	FMRP (fragile X mental retardation protein) inter	1
Unknown function			
	Q8VDP4	2610301G19Rik protein	3
	Q9CXN5	3110050K21Rik protein	2
	Q7TQF8	Camsap1	2
	Q80TK8	Camsapl1 (MKIAA1078 protein)	1
	Q80UZ7	D430018P08 protein	3
	Q80YT1	Glucocorticoid induced transcript 1 (Glcci1)	1
	Q6NXJ5	Gm1568	2
	Q6PG39	HSHIN1 protein	2
	Q9UQ09	KARP-1-binding protein 1	5
	Q8CIL3	Parcc1	8
	Q6PDH0	Phldb1	2
	Q8K004	Spata2	1
	Q5M7V8	Thyroid hormone receptor associated protein 3	5

Table 4.2 – Phosphoproteins and numbers of phosphorylation sites characterised in the cytosolic phosphoproteome

4.4.4 Intrinsic protein disorder and phosphorylation

4.4.4.1 Putative intrinsically-disordered phosphoproteins in the brain

We investigated the relationship between intrinsic protein disorder and protein phosphorylation in the cytosolic phosphoproteome dataset. The DisEMBL Intrinsic Protein Disorder Prediction server was used to identify intrinsically disordered regions or "hotloops" in each of the cytosolic phosphoproteins characterised in this study. To evaluate this distribution of intrinsic disorder in this dataset, we compared it to a dataset of experimentally verified wholly disordered proteins (52 proteins) (Oldfield et al., 2005) and a dataset of wholly ordered proteins (100 proteins) (Proteins in PDB with X-ray structure data) (Oldfield et al., 2005). Figure 4.6 shows the comparison of these 3 datasets in terms of percentage-disordered sequence for each protein versus log of molecular weight. The percentage intrinsic disorder in cytosolic phosphoproteins ranges from 5.4 to 88.7 % of total sequence lengths, with an average of 46% sequence disorder. The average intrinsic disorder in the experimentally verified wholly disordered dataset predicted by DisEMBL was 63% and 14 cytosolic phosphoproteins have a percentage disorder of 62% or higher. TPPP, Microtubule-associated protein Tau and Beta-adducin, which are known intrinsically disordered proteins, are indicated on the graph in Figure 4.6.

These 14 predicted intrinsically disordered phosphoproteins (including TPPP (which is a known IDP) and 2 other known IDP's (Tau and Beta-adducin) are described in Table 4.3. The majority of these putative IDP's are RNA/DNA binding proteins, a class of proteins that are generally unstructured until bound to their target molecule. The 3 known IDP's (TPPP, Tau and Beta-adducin) are cytoskeletal proteins highly expressed in the brain and are implicated in neurodegenerative diseases where protein aggregation is a prominent pathological feature.

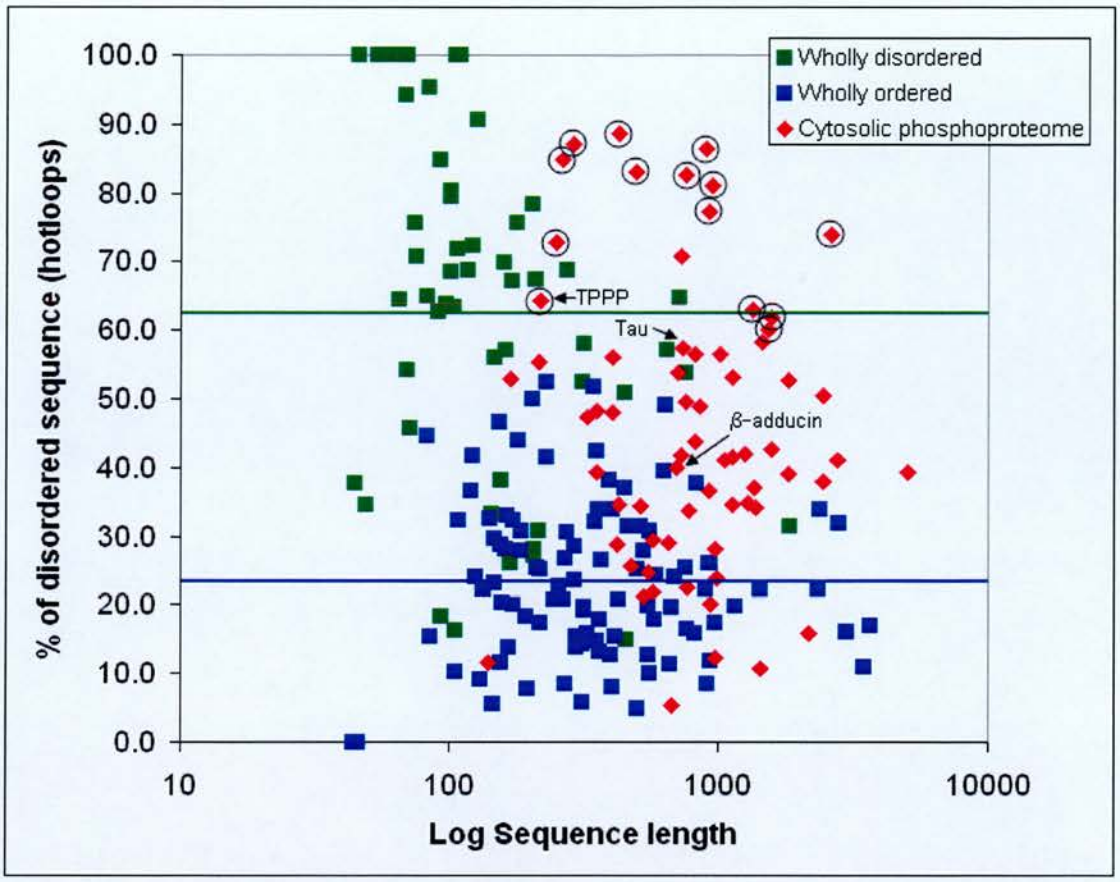


Figure 4.6 – Comparison of the distribution of hotloops in datasets of wholly disordered, wholly disordered and cytosolic phosphoproteins (characterised in this study). The percentage of disordered sequence (hotloops) was calculated for each of these datasets and was plotted against log of protein sequence length. The horizontal blue and green lines represent the average percentage of disorder in the "Wholly disordered" and "Wholly disordered" datasets respectively. Cytosolic phosphoproteins with a percentage disorder equal to or greater than that of the "Wholly disordered" dataset are circled and are therefore putative intrinsically (natively) unstructured phosphoproteins. TPPP, Microtubule-associated protein Tau and Beta-adducin which are known intrinsically disordered proteins are indicated on the graph

UniProt	Name	% disorder	P sites	P sites in disorder
Q9CXN5	3110050K21Rik protein	88.7	2	2
P62996	Arginine/serine-rich splicing factor 10 (Sfrs10)	87.2	7	7
O70495	Plenty-of-prolines-101 (Srrm1)	86.3	19	16
Q9R0U0	Neural specific sr protein NSSR 1	84.7	3	3
Q8VE97	Splicing factor, arginine/serine-rich 4 (Sfrs4)	83.0	3	3
Q13427	Peptidyl-prolyl cis-trans isomerase G (Cyclophilin G)	82.5	2	2
Q5M7V8	Thyroid hormone receptor associated protein 3 [Thrap3]	81.1	5	5
Q8K019	Bcl-2-associated transcription factor 1 [Bclaf1]	77.3	5	3
Q8BTI8	RNA binding protein homolog (Srrm2)	74.0	33	31
P62754	40S ribosomal protein S6 (RPS6)	72.7	1	1
Q96MU7	Putative splicing factor YT521	70.7	1	1
Q7TQD2	Tubulin polymerization-promoting protein (p25 alpha)**	64.2	4	2
Q9JIX8	Apoptotic chromatin condensation inducer in the nucleus (Acin1)	62.9	2	1
Q9UQ09	KARP-1-binding protein 1	61.8	5	4
P10637	Microtubule-associated protein Tau**	57.4	19	14
Q9QYB8	Beta adducin**	41.7	6	6

Table 4.3 – Putative intrinsically disordered cytosolic phosphoproteins in mouse forebrain. The percentage disorder, number of phosphorylation (P) sites and the number of phosphorylation sites in disordered regions are presented. ** indicates experimentally verified intrinsically disordered proteins. The entries highlighted in blue have % disorder values substantially lower than the average for the control disordered dataset but they have been experimentally verified as natively disordered proteins.

4.4.4.2 Phosphorylation in regions of intrinsic protein disorder.

It has been observed that the attributes surrounding phosphorylation sites such as amino acid composition, hydrophobicity and sequence complexity are quite similar to those of intrinsically disordered regions in proteins. In fact, a phosphorylation prediction algorithm (DISPHOS) uses such attributes of IDP's to predict phosphorylation sites. In the cytosolic phosphoproteome dataset, 231 phosphorylation sites (72%) are predicted to lie in intrinsically disordered regions. In addition, of the remaining 90 phosphorylation sites that are predicted to occur in ordered regions, only 5 (1.5%) can be mapped to Pfam protein domains; a tyrosine phosphorylation site (Y216) in GSK3B is present in its kinase domain, 3 pS/T sites in Numb are located in its Numbf domain and phosphoserine 495 in Similar to PITPNM family member 3 is located in a DDHD domain. As few of the phosphorylation sites predicted to be present in ordered regions are actually located in known protein domains, either the algorithm for predicting hotloops

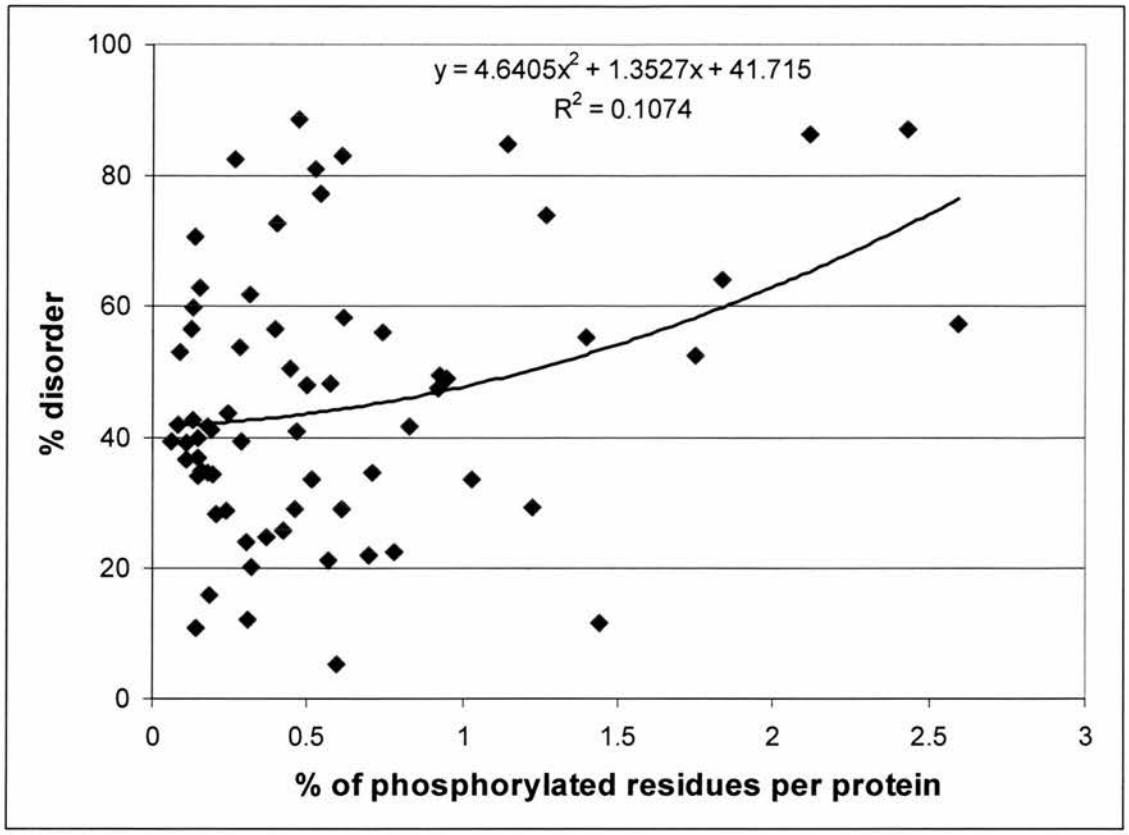


Figure 4.7 – Investigation of the relationship between intrinsic protein disorder and protein phosphorylation. The percentage disorder of each cytosolic phosphoprotein was plotted against the percentage of phosphorylated residues per protein. The best fitting trend line indicates a possible positive relationship between disorder and phosphorylation, although this result is not statistically significant.

underestimates protein disorder or these phosphorylation sites are located in other types of low complexity regions.

Are IDP's phosphorylated to a greater extent than structured proteins? As IDP's contain a greater proportion of protein sequence with attributes similar to that surrounding phosphorylation sites and have more accessible loops for kinases to phosphorylate, we wanted to assess whether IDP's are indeed more likely to be phosphorylated. Figure 4.7 shows the relationship between percentage disorder (of protein sequence) and the percentage of phosphorylated residues per protein. The best fitting trend line to the data points indicates a positive relationship between disorder and extent of phosphorylation. The data used for this analysis is most likely biased towards what appear to be lesser-phosphorylated proteins but this is probably because many of these proteins may be phosphorylated on other sites that were not detected in our experiments. It would seem likely that if this dataset was complete that a stronger correlation would be apparent. Microtubule-associated protein Tau is an IDP is well known to be highly phosphorylated with at least 35 known phosphorylated sites (19 identified in this study). Four phosphorylation sites in Tau are located in tubulin-binding domains, two of which are in disordered sequence (Figure 4.8).

4.4.4.3 Mapping phosphopeptides to regions of intrinsic disorder

Phosphopeptides characterised in this study were mapped to disorder plots and sequences of protein disorder (Figure 4.9). CaMKIIγ is very structured with only 21% disordered sequence and a large protein kinase domain that constitutes more than half of the protein sequence. A phosphopeptide in CaMKIIγ and another in CaMKIIβ (homologous sequence) map to the main disordered region outside the protein kinase domain. The occurrence of these phosphopeptides in a small window of disordered sequence would seem to suggest that there is a link between phosphorylation and disorder.

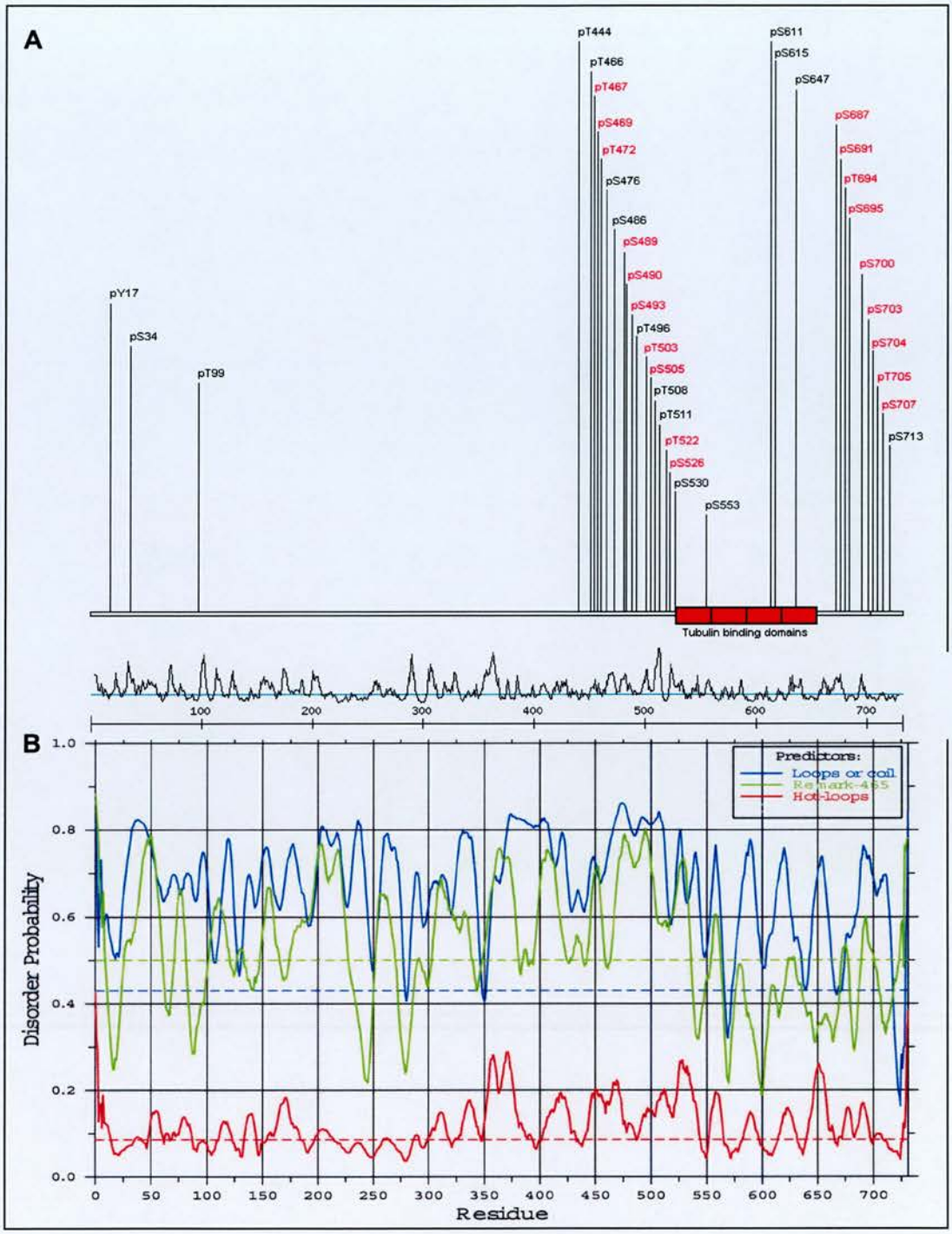


Figure 4.8 – Phosphorylation sites and disorder plot of Microtubule-associated protein Tau. A. Schematic illustration of the domain profile of Tau with all known phosphorylation sites and phosphorylation sites characterized in this study (red). B. Disorder plot of Tau showing three measures of protein disorder. The blue line represents coiled regions, the red line represents coiled regions with high B factors (high degree of mobility as determined from Cα temperature factors) and green line, which represents missing, coordinates in X-ray structure as defined by remark465 entries in PDB. The dashed horizontal lines correspond to the random expectation level for each predictor. 57% of the sequence of this known intrinsically disordered protein is predicted to be disordered. The best measure of disorder is high B factors (red line) and it can be seen that there is more disorder, as measured by high B factors, in the N-terminal part of the sequence, where the majority of phosphorylation occurs.

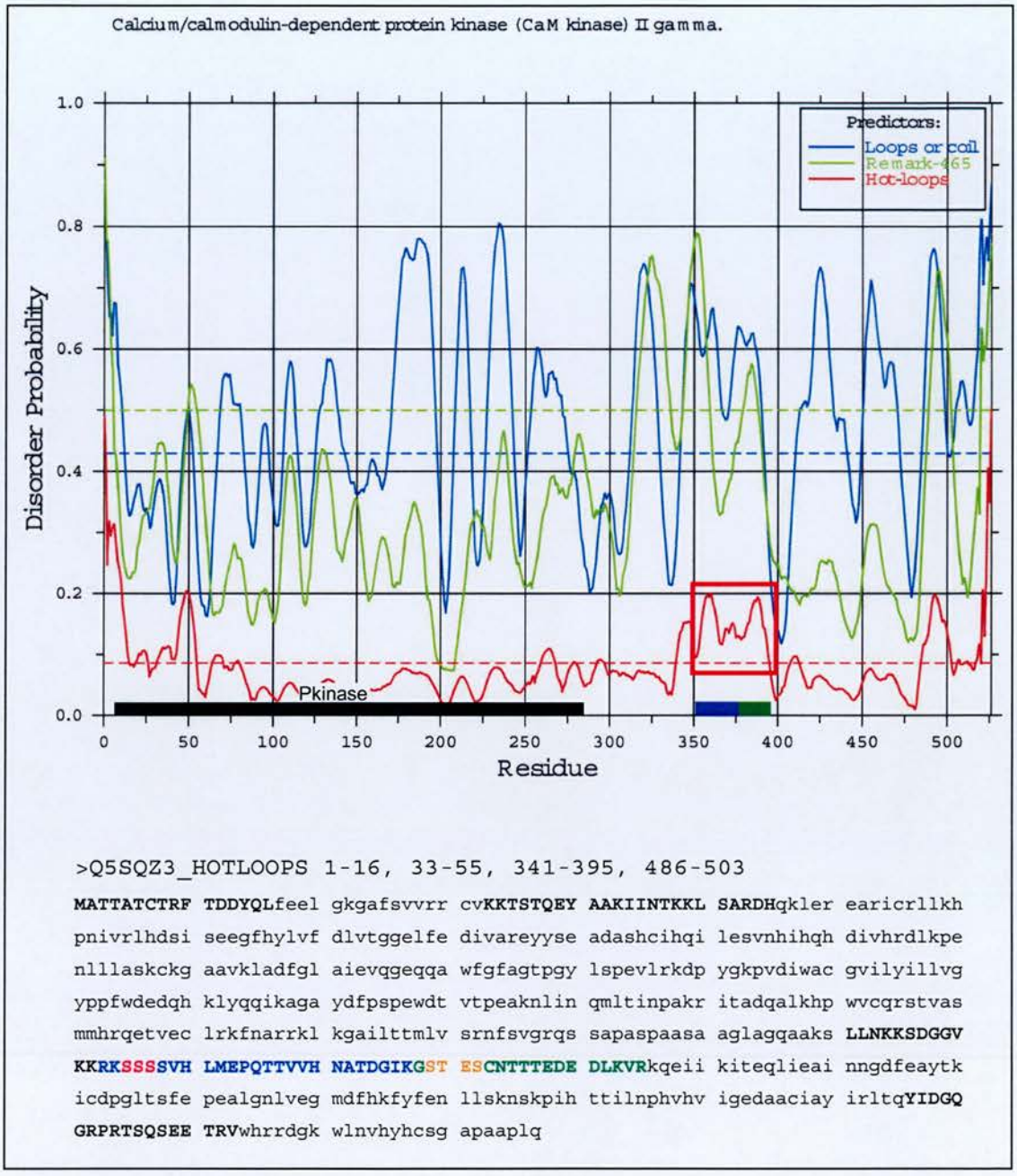


Figure 4.9 – Disorder and phosphorylation of CaMKIIγ
Disorder plot of CaMKIIγ showing three measures of protein disorder. The blue line represents coiled regions, the red line represents coiled regions with high B factors (high degree of mobility as determined from Cα temperature factors) and green line which represents missing coordinates in X-ray structure as defined by remark465 entries in PDB. The dashed horizontal lines correspond to the random expectation level for each predictor. The protein kinase domain is indicated by a black rectangle and a phosphopeptides CaMKIIγ and CaMKIIβ (homologous sequence) in blue and green respectively. Unambiguously assigned phosphorylation sites are coloured in red and ambiguously assigned sites are coloured in gold. Although, CaMKIIγ contains only 21% disordered sequence the phosphopeptides from CaMKIIγ and CaMKIIβ map to a clearly disordered segment of the sequence (red box).

4.4.4.4 Effect of intrinsic disorder on tryptic cleavage and analysis by MS

Analysis of datasets of experimentally verified intrinsically disordered proteins have identified amino acid propensities in disordered sequences that are used by DisEMBL to predict disorder in proteins (Linding et al., 2003). We noticed that lysine and arginine were among the significantly enriched amino acids in disordered sequences and that because trypsin cleaves at these residues, we reasoned that there may be a relationship between the frequency of tryptic cleavage and protein disorder. To investigate this we calculated the percentage intrinsic disorder and number of tryptic peptides (by *in silico* digestion) for all proteins in the cytosolic phosphoproteome dataset (Figure 4.10). This *in silico* digestion did not include the single amino acid peptides R and K. We observed that disordered proteins tend to be cleaved into more tryptic fragments compared to less disordered proteins. The shortest average peptide length was 3.9 amino acids for Splicing factor, Arg/Ser rich 4 (83 % disorder) and the longest peptide length was 11.5 amino acids for SR-related CTD associated factor 6 (15.9% disorder).

If indeed disordered proteins are more likely to be highly phosphorylated then complete mapping of these phosphorylation sites by tryptic digestion and tandem mass spectrometry may prove difficult as many of the resultant tryptic peptides would be outside the mass window for analysis. Next the monoisotopic masses and isoelectric points from an *in silico* tryptic digest of Srrm1 (86.3% disorder, average peptide length 5.4 AA's) and Synapsin I (5.4% disorder, average peptide length of 10.8 AA's) were calculated and plotted with experimentally determined Mr values (and calculated isoelectric points) for phosphopeptides characterised in the cytosolic phosphoproteome dataset (Figure 4.11). We observed that there are substantially

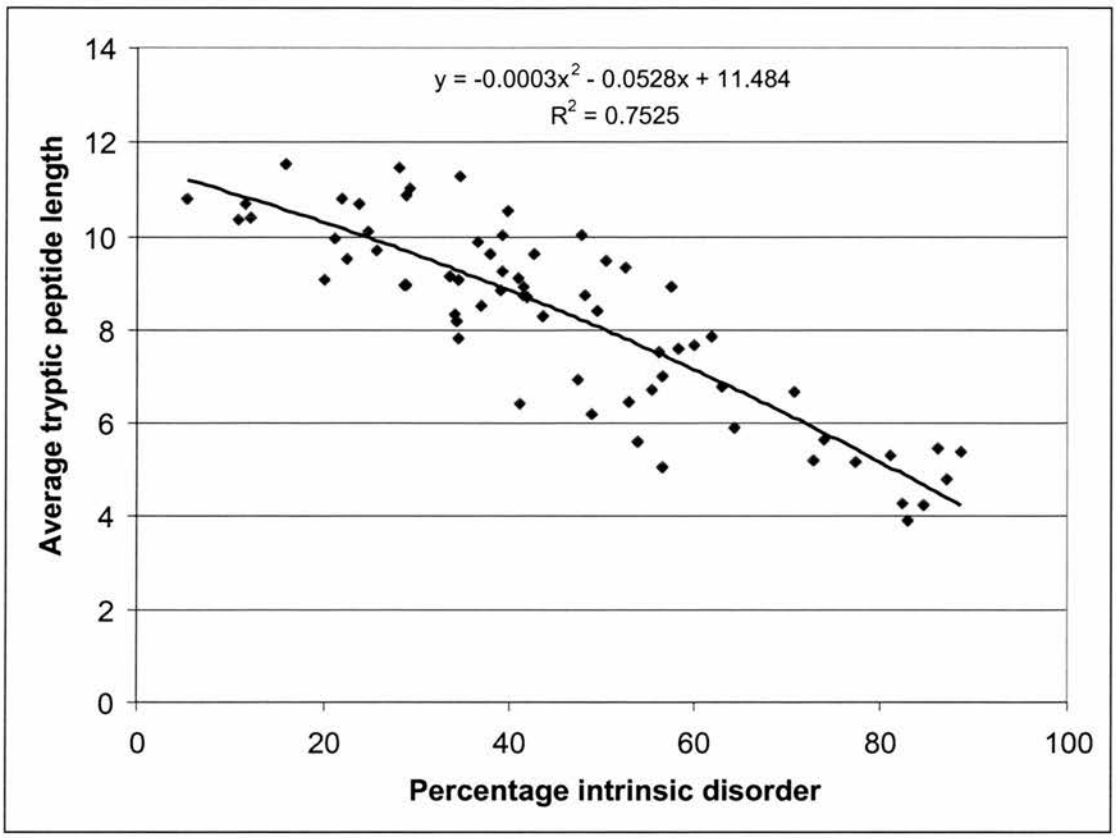


Figure 4.10 – Relationship between intrinsic disorder and frequency of tryptic cleavage. The percentage intrinsic disorder and number of tryptic peptides (in silico digestion) was calculated for all proteins in the cytosolic phosphoproteome dataset. Disordered proteins tend to be cleaved into more tryptic fragments compared to less disordered proteins. Argine and lysine residues are positively enriched in regions of protein disorder and therefore these sequences are more likely to be cleaved by trypsin (cleaves at arg and lys residues). If indeed disordered proteins are more likely to be highly phosphorylated then complete mapping of these phosphorylation sites by tryptic digestion and tandem mass spectrometry may prove difficult as many of the resultant tryptic peptides would outside the mass window for analysis.

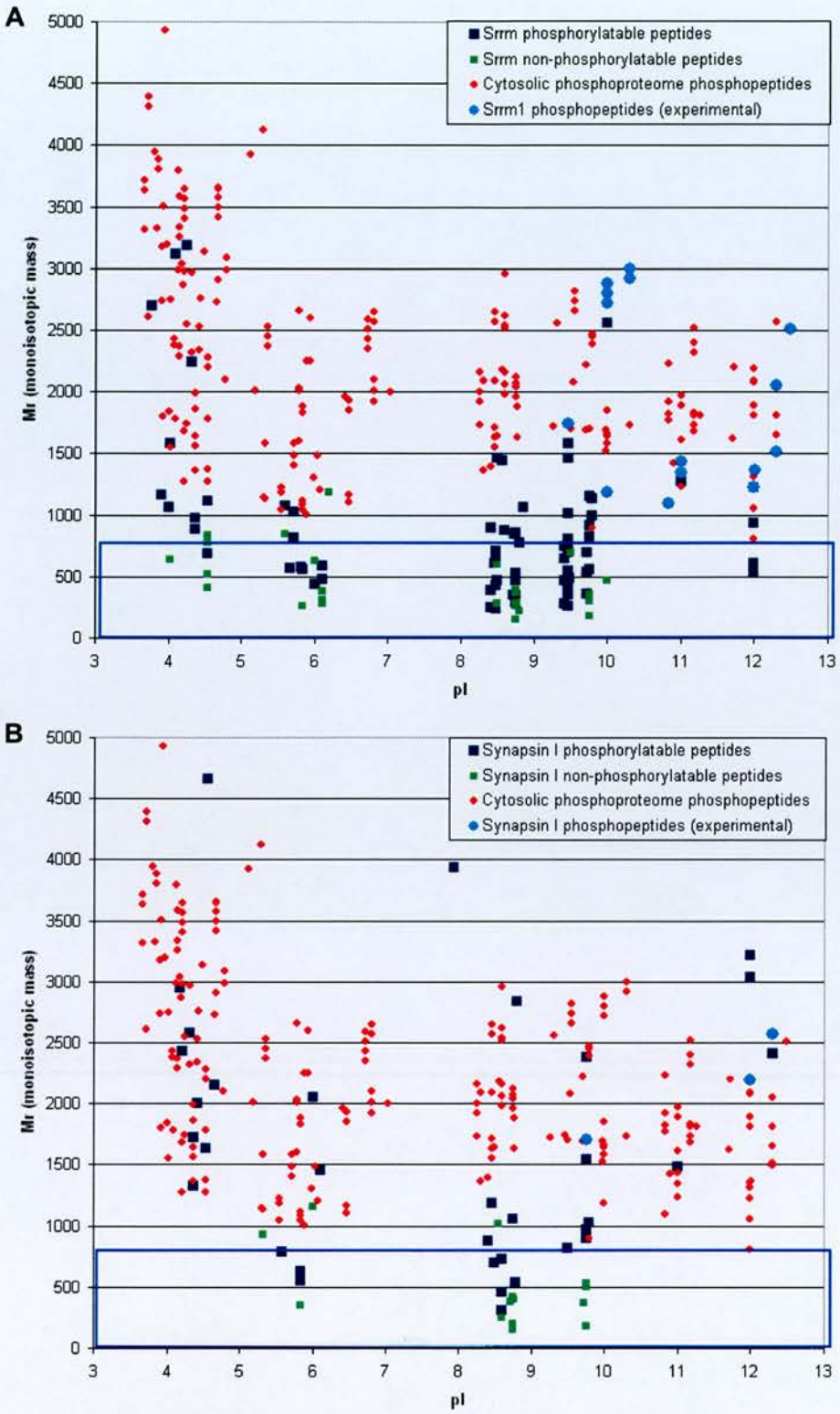


Figure 4.11 – Compatibility of tryptic digestion with phosphoproteomic analysis of disordered and ordered proteins. The monoisotopic masses and isoelectric points from an in silico tryptic digest of Srrm1 (86.3% disorder, average peptide length 5.4 AA's) and Synapsin I (5.4% disorder, average peptide length of 10.8 AA's) were calculated and plotted with experimentally determined Mr values (and calculated pI's) for phosphopeptides characterised in the cytosolic phosphoproteome dataset. Peptides with Ser/Thr or Tyr residues are marked in dark blue to allow comparison with the experimentally verified phosphopeptides. It can be seen that there are substantially more phosphorylatable peptides in Srrm1 outside the mass range of detection (blue box) than for Synapsin I

more phosphorylatable peptides in Srrm1 outside the mass range of detection (blue box) than for Synapsin I. 62 out of a total of 97 phosphorylatable peptides in Srrm1 were below 800 Da meaning that 64% of peptides potentially carrying a phosphorylated serine/threonine or tyrosine were not amenable to analysis by MS. In contrast, only 9 out of 37 phosphorylatable peptides in Synapsin I were below 800 Da meaning that 24% of putative phosphopeptides were not suitable for analysis by MS.

The frequency of tryptic cleavage is not only increased between ordered and disordered proteins but between ordered and disordered regions in the same protein sequence. A tryptic cleavage map of MAP4, which is 53% disordered is shown in Figure 4.12. There are 87 unstructured peptides covering 676 amino acids with an average peptide length of 7.7 compared to 28 structured peptides covering 449 amino acids, resulting in an average peptide length of 16 amino acids. There are more than double the number of peptides arising from disordered compared to ordered sequence when adjusted for relative sequence coverage. This would indicate that even within a single protein, disorder affects the sequence coverage obtained by mass spectrometry when tryptic digestion is used. Also, as phosphorylation is more likely to occur in disordered sequence, the coverage of phosphorylation sites is expected to be influenced by disorder too.

MADLSLVDALTEPPPEIEgeik r dfmaaleaepyddivgetvek tefiPLLDGDEK TGNSESK
K

K PCLDTSQVEGIPSsk ptllangdhgmegnntagsptdfleer

Vdypdyqssqnwp edasfcfqpqqvldtdqaepfnehr ddgladllfvssgptnasafter

dnpsedsygmlpcdsfastavvsqewsvgapnspcsescvspevtietlqpatelsk aaevesvk

eqlpak aletmaeQTTDVVHSPSTDTTPGPdteaalak dieeitk

pdvilanvtqpstesdmflaqdmelltgteaahanniilptepdesstk

dvappmeEEIVPGNDTTSPK ETETTLPIK MDLAPpedvlltk etelapak gmvslseieealak

ndvr saeiPVAQETVVSETevvlaTEVVLPSDPITTLTK DVTLPLEAER

PLVTDMTPSLETEMTLGK ETAPPTetnlgmak dmsplpesevtlgk dvviLPETK

VAEFNNVTPLSEEEVTSVK DMSPsaeteaplak nadlhsgtelivdnsmapasdlalPLETK

VATVPIK DK GTVQTEEK PR EDSQLASMQHK GQSTVPPCtaspepvk

aaeqmstlpidapsplenleqk etPGSQPSEPCSGVSr qeeak aaVGVTGNDITTPPNK

EPPPSPEK K AK PLATTQPAK TSTSK AK TQPTSLPK QPAPTTSGGLNk k

pmslasgsvpaaphk r paAATATAR PSTLPAR DVK PK PITEAK VAEK R TSPSK

PSSAPALK PGPK TTPTVSK ATSPSTLVSTGPSSR SPATTLPK R PTSIK TEGK PADVK R

MTAK SASADLSR SK TTSASSVK R NTTPTGAAPPAGMTSTR VK PMSAPSR SSGALSVDK K

PTSTK PSSSAPR VSR LATTVSAPDLK SVR SK VGSTENIK HQPGGGR AK VEK K

TEAATTAGK PEPNAvtk aagSIASAQK PPAGK vqivsk k vsyshiqsk CVSK DNIK

HVPGCGNVQIQNK K VDISK VSSK CGSK ANIK HK PGGGDVK iesqk lnfk ek aqak

vgsldnvghfpaggavk tegggsealpcpgppageepvipeaapdr

gaptsasglsghtTLSGGGDQR EPQTLDSQIQETSI

Figure 4.12 - In silico digestion of Microtubule-associated protein 4. The predicted tryptic cleavage pattern for MAP4 is shown. This protein contains 53% disordered sequence, which is indicated by capitals and tryptic peptides containing mostly disordered sequence are highlighted in blue. There are 87 unstructured peptides covering 676 amino acids with an average peptide length of 7.7 amino acids and 33 of these peptides contain 4 or less amino acids. The structured sequence is indicated by lowercase sequence and mostly ordered peptides are highlighted in yellow. There are 28 structured peptides covering 449 amino acids, resulting in an average peptide length of 16 amino acids and 7 of these peptides contain 4 or less amino acids. The phosphopeptide found in MAP4 is underlined.

4.4.5 Identification of novel phosphorylation sites in protein kinases

We have identified 37 phosphorylation sites in nine protein kinases from analysis of the forebrain cytosolic phosphoproteome. These kinases are mapped to the human kinome tree in Figure 4.13. The major family of protein kinases in our dataset is the CAMK family with 5 members (CaMKII β, CaMKIIγ, DCAMKL1, BRSK1 and BRSK2), next the CMGC family with 2 members (GSK3β and PRP4) and finally the AGC and STE families with one member each (MAST1 and TNIK respectively). We found 5 pS/T in the variable region of CaMKII β and CaMKIIγ (Figure 4.14).

The phosphopeptide found in CaMKII γ contains 3 consecutive phosphorylation sites which are located in the variable domain in a region present only in CaMKII α and γ. The CaMKII β phosphopeptide is also in the variable domain, located C-terminally and contains 2 pS/T (ambiguously assigned) sites which are in a relatively conserved portion of the variable domain. To help assign the exact sites of phosphorylation in this phosphopeptide sequence we carried out a multiple sequence alignment of sequences from the variable domains of members of the CAMK protein kinase family (Figure 4.15). The phosphopeptide found in the variable domain of CaMKII β is well conserved across all CAMKII family members from different species. The 2 pS/T were located to 4 possible sites in the phosphopeptide sequence. Analysis of this evolutionary alignment of CAMKII family members revealed that the first and third serine residues (S354, S357) are completely conserved and thus are more likely to be the actual sites of phosphorylation.

One of the six known phosphorylation sites out of the 37 characterised on protein kinases was pY216 on GSK3β. This site is completely conserved in many members of the GSK3 kinase family and across 13 species as shown in Figure 4.16. This is an important regulatory site on this kinase and is located in the protein kinase domain, which is also a predicted region of sequence order that has enabled X-ray structure

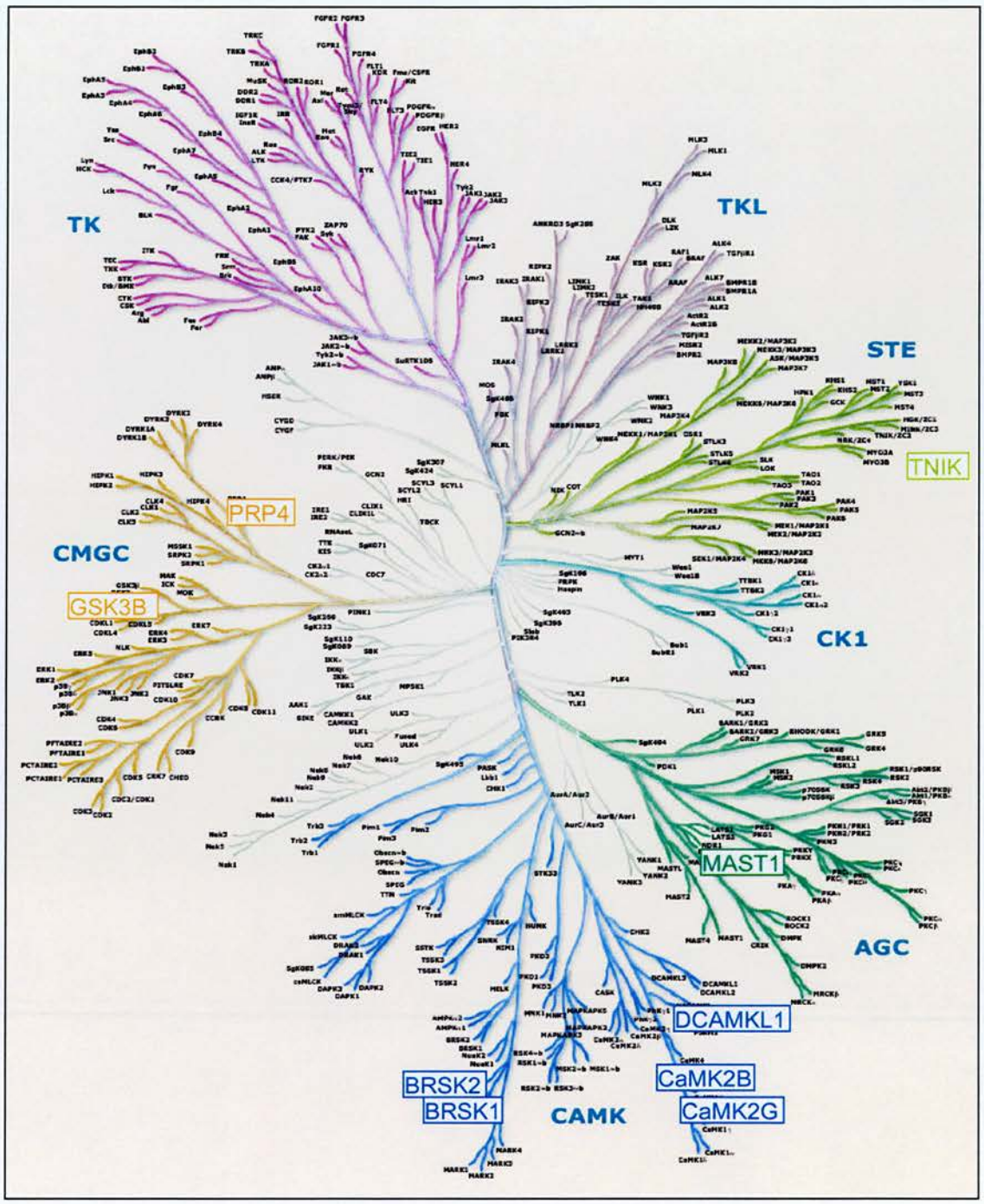


Figure 4.13 – Human kinome tree showing the evolutionary relationship between members of the protein kinase family. Phosphorylation sites on nine protein kinases were characterised in this study and are mapped to the human kinome tree shown.

Adapted from "The protein kinase complement of the human genome." Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. 2002. Science 298:1912-1934.

IPI	UniProt	Protein Name	Phosphopeptide Sequences	Mod ID	Residues	MS defined Phospho-sites	Predicted Phospho-sites	Order	Known kinase	Literature
IPI00326171	Q68EG2	CaMKII beta	GSLPPAALESSDSTNTTIEDEDAK	2 pS/T	345-368		S354	Disord		
							T358	Disord		
IPI00296678	Q5SQZ3	CaMKII gamma	RKpSpSpSSVHLMEPQTTVVHNATDGIK	3 pS	353-377	S355		Disord		
						S356		Disord		
						S357		Disord		
IPI00292228	Q9WV60	Glycogen synthase kinase-3 beta	GEPNVSpYICSR	1 pY	210-220	Y216		Ord	MEK1	15020233
IPI00515701	Q5RJI5	Serine/threonine kinase SADB (BRSK1)	pSpTGLpSpSpSPLpSpSPRpSPVFSFSPEPGAGDEAR	8 pS/T	437-466	S437		Disord		
						T438		Disord		
						S441		Disord		
						S442		Disord		
						S443		Disord		
						S446		Disord		
						S447		Disord		
						S450		Disord		
			SSGGTPLHpSPLHTPR	1 pS + 1 pS	511-525	S519		Ord		
			ASPTGTPGTpTPPPSPGGGVGGAAWR	1 pS/T + 1 pT	526-550		S527	Ord		
						T535		Ord		
IPI00360155	Q6DMN7	BR serine/threonine-protein kinase 2 (BRSK2)	pSMEVLSVTDGGpSPVPAR	2 pS	383-399	S383		Disord		
						S394		Disord		
			SRSISGASSGLSTpSPLSSPR	1 pS + 1 pS	411-430	S424		Ord		
IPI00312081	Q9JLM8	Serine/threonine-protein kinase DCAMKL1	SGKpSPSPSPpTpSPGSLRK	3 pS	327-343	S330		Disord		
						T336		Disord		
						S337		Disord		
			ISQHGGSSTpSLSSTK	1 pS + 1 pS/T	346-360	S355		Disord		
			AQPAPPELNSESEDYpSPSSSETVR	1 pS + 1 pS	727-750		S736	Disord		
							S738	Disord		
						S742		Disord	uk	15345747
IPI00320690	Q61136	Serine/threonine-protein kinase PRP4 homolog	VQSGMGLILQGYEpSGpSEEEGEIHEK	2 pS	130-141	S143		Ord		
						S145		Ord		
			SKDApSPINRWpSPTR	2 pS	427-440	S431		Ord	uk	15302935
						S437		Ord	uk	15302935
IPI00128775	Q9R1L5	Syntrophin-associated serine-threonine-protein kinase	VGpSpSHTRR	2 pS	1226-1233	S1228		Ord		
						S1229		Ord		
IPI00372166	Q9UKE5	TRAF2 and NCK interacting kinase	VRANpSKSEGpSPVLPHEPSK	2 pS	760-778	S764		Disord	uk	15144186
						S769		Disord	uk	15144186

Table 4.4 – Characterisation of phosphorylation sites on cytosolic protein kinases

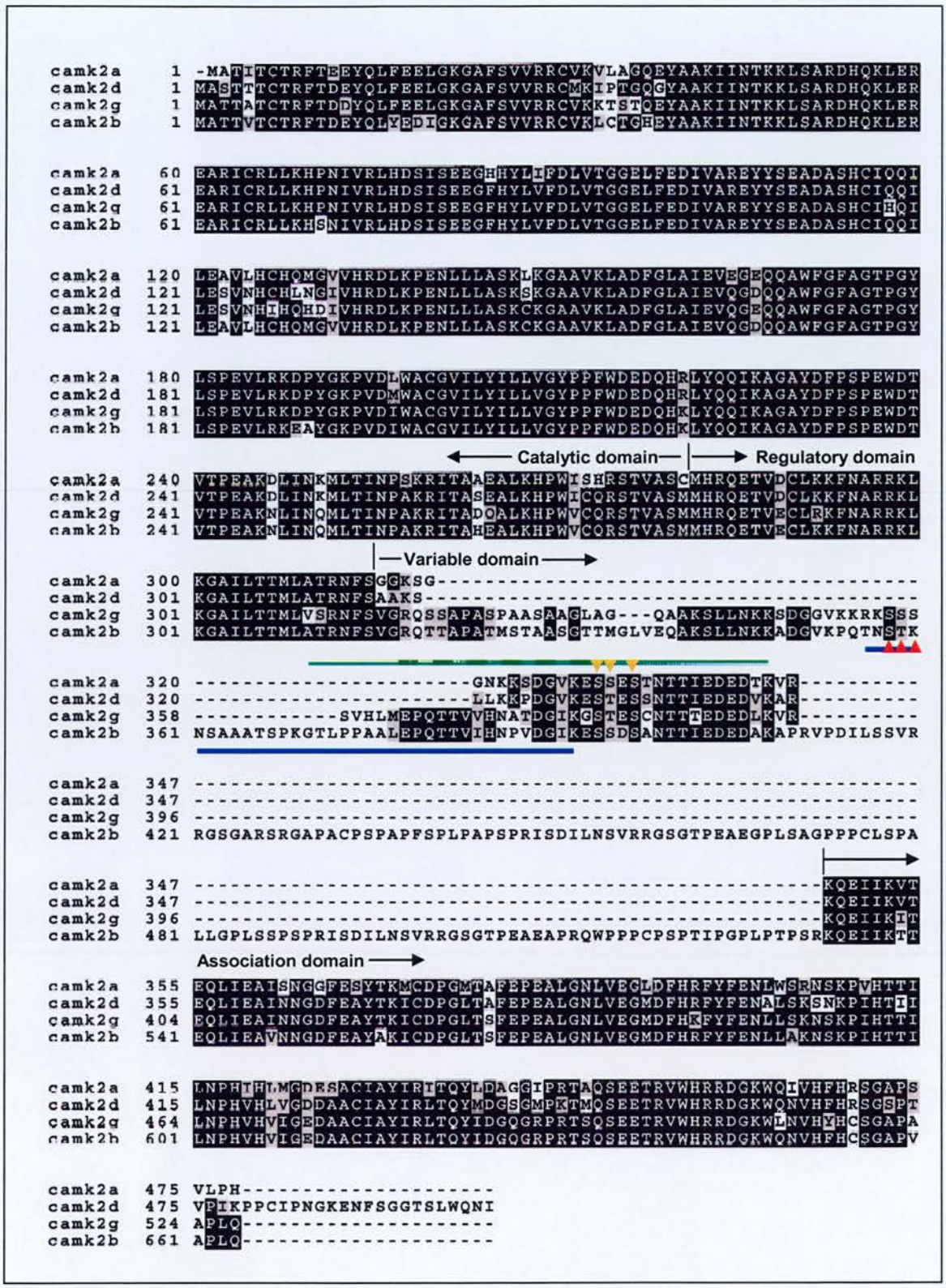
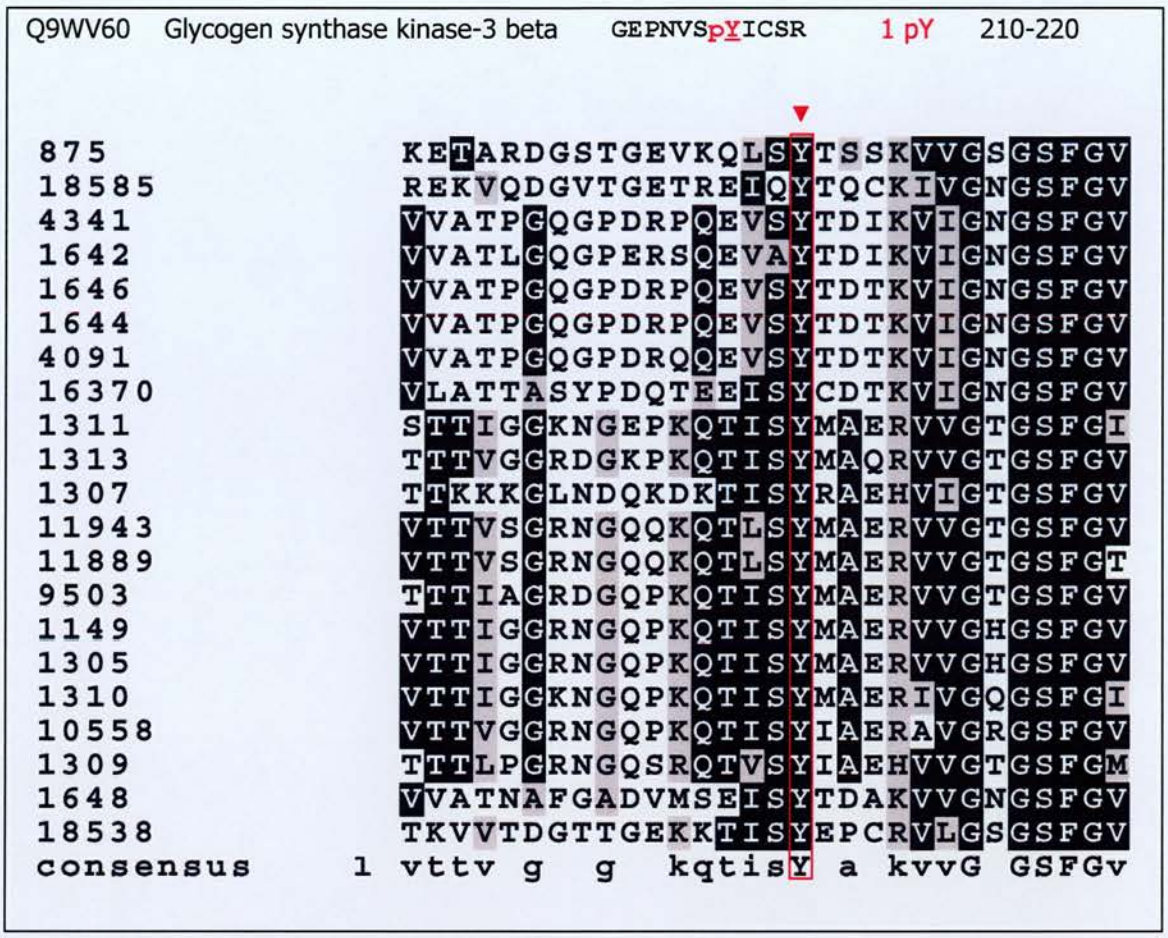


Figure 4.14 – Multiple sequence alignment of CaMKII α, β, γ and δ.
The phosphopeptide found in CaMKII γ (Blue line) contains 3 consecutive phosphorylation sites which are located in the variable domain in a region present only in CaMKII α and γ. The CaMKII β phosphopeptide is also in the variable domain, located C-terminally and contains 2 pS/T (ambiguously assigned) sites which are in a relatively conserved portion of the variable domain.

PKR id	Name	Description	Organism	Class
1430	KCCB_RAT	Calcium/calmodulin-dependent protein kinase type II beta chain	Rattus norvegicus	CAMK
6193	Q8BL41	Calcium/calmodulin-dependent protein kinase type II beta chain	Mus musculus	CAMK
1429	CAMKIIB	Calcium/calmodulin-dependent protein kinase type II beta chain	Mus musculus	CAMK
1428	CaMK2beta	Calcium/calmodulin-dependent protein kinase type II beta chain	Homo sapiens	CAMK
3637	O93560	Calcium/calmodulin-dependent kinase type II beta subunit	Gallus gallus	CAMK
4080	Q91549	Calmodulin dependent protein kinase II beta subunit	Xenopus laevis	CAMK
4088	Q91587	Calcium/calmodulin-dependent kinase type II beta'-subunit	Xenopus laevis	CAMK
1434	KCCG_MOUSE	Calcium/calmodulin-dependent protein kinase type II gamma chain	Mus musculus	CAMK
13647	Q8WMW7	CaM kinase II gamma C-2	Mustela putorius furo	CAMK
12912	O77706	Calmodulin-dependent protein kinase II-gamma dash1	Oryctolagus cuniculus	CAMK
1435	CAMKIIG	Calcium/calmodulin-dependent protein kinase type II gamma chain	Rattus norvegicus	CAMK
16711	Q8NIA4	Calcium/calmodulin-dependent protein kinase II gamma	Homo sapiens	CAMK
1433	CAMK2gamma	Calcium/calmodulin-dependent protein kinase type II gamma chain	Homo sapiens	CAMK
1431	CaMK2delta	Calcium/calmodulin-dependent protein kinase type II delta chain	Homo sapiens	CAMK
1432	CAMKIID	Calcium/calmodulin-dependent protein kinase type II delta chain	Rattus norvegicus	CAMK
12879	Q95266	Calcium/calmodulin-dependent protein kinase II delta 2-subunit	Sus scrofa	CAMK
16765	Q9Y2H4	KIAA0968 protein	Homo sapiens	CAMK
17638	Q8IWE0	Calcium/calmodulin-dependent protein kinase IIA; isoform 2	Homo sapiens	CAMK
1424	CAMK2alpha	Calcium/calmodulin-dependent protein kinase type II alpha chain	Homo sapiens	CAMK
1425	KCCA_MOUSE	Calcium/calmodulin-dependent protein kinase type II alpha chain	Mus musculus	CAMK
1427	CAMKIIA	Calcium/calmodulin-dependent protein kinase type II alpha chain	Rattus norvegicus	CAMK

Figure 4.15 – Multiple sequence alignment of sequences from the variable domains of members of the CAMK protein kinase family. The phosphopeptide found in the variable domain of CaMKII β is well conserved across all CAMKII family members from different species. The 2 pS/T were located to 4 possible sites in the phosphopeptide sequence and analysis of this evolutionary alignment of CAMKII family members, we found that the first and third serine residues (S354, S357) are completely conserved and thus are more likely to be the actual sites of phosphorylation.



PKR id	Name	Description	Organism	Class
875	GSK3_SCHPO	Protein kinase gsk3	Schizosaccharomyces pombe	CMGC
18585	Q9HE92	Probable glycogen synthase kinase 3 alpha	Neurospora crassa	CMGC
4341	Q9YH61	Glycogen synthase kinase 3 alpha	Danio rerio	CMGC
1642	GSK3a	Glycogen synthase kinase-3 alpha	Homo sapiens	CMGC
1646	KG3B_RAT	Glycogen synthase kinase-3 beta	Rattus norvegicus	CMGC
1644	GSK3b	Glycogen synthase kinase-3 beta	Homo sapiens	CMGC
4091	Q91627	Glycogen synthase kinase 3 beta	Xenopus laevis	CMGC
16370	Q9GTK0	Serine/threonine kinase GSK3	Hydra vulgaris	CMGC
1311	KSG6_ARATH	Shaggy-related protein kinase dzeta	Arabidopsis thaliana	CMGC
1313	KSG8_ARATH	Shaggy-related protein kinase theta	Arabidopsis thaliana	CMGC
1307	KSG2_ARATH	Shaggy-related protein kinase beta	Arabidopsis thaliana	CMGC
11943	O82038	Shaggy kinase 6	Petunia x hybrida	CMGC
11889	O24139	Shaggy-like kinase 6	Nicotiana tabacum	CMGC
9503	Q9FSG4	Wound-induced GSK-3-like protein	Medicago sativa	CMGC
1149	MSK1_MEDSA	Glycogen synthase kinase-3 homolog MsK-1	Medicago sativa	CMGC
1305	KSG1_ARATH	Shaggy-related protein kinase alpha	Arabidopsis thaliana	CMGC
1310	KSG5_ARATH	Shaggy-related protein kinase epsilon	Arabidopsis thaliana	CMGC
10558	Q94IZ2	Putative shaggy-related protein kinase	Oryza sativa (japonica cultivar-group)	CMGC
1309	KSG4_ARATH	Shaggy-related protein kinase delta	Arabidopsis thaliana	CMGC
1648	KG3H_DROME	Putative glycogen synthase kinase-3 homolog	Drosophila melanogaster	CMGC
18538	P78772	Schizosaccharomyces pombe	Schizosaccharomyces pombe	CMGC

Figure 4.16 – Multiple sequence alignment of sequences from protein kinase domain of GSK3 family members. Phospho-tyrosine 216 is an important regulatory site in GSK3B and is well conserved across GSK3 family members.

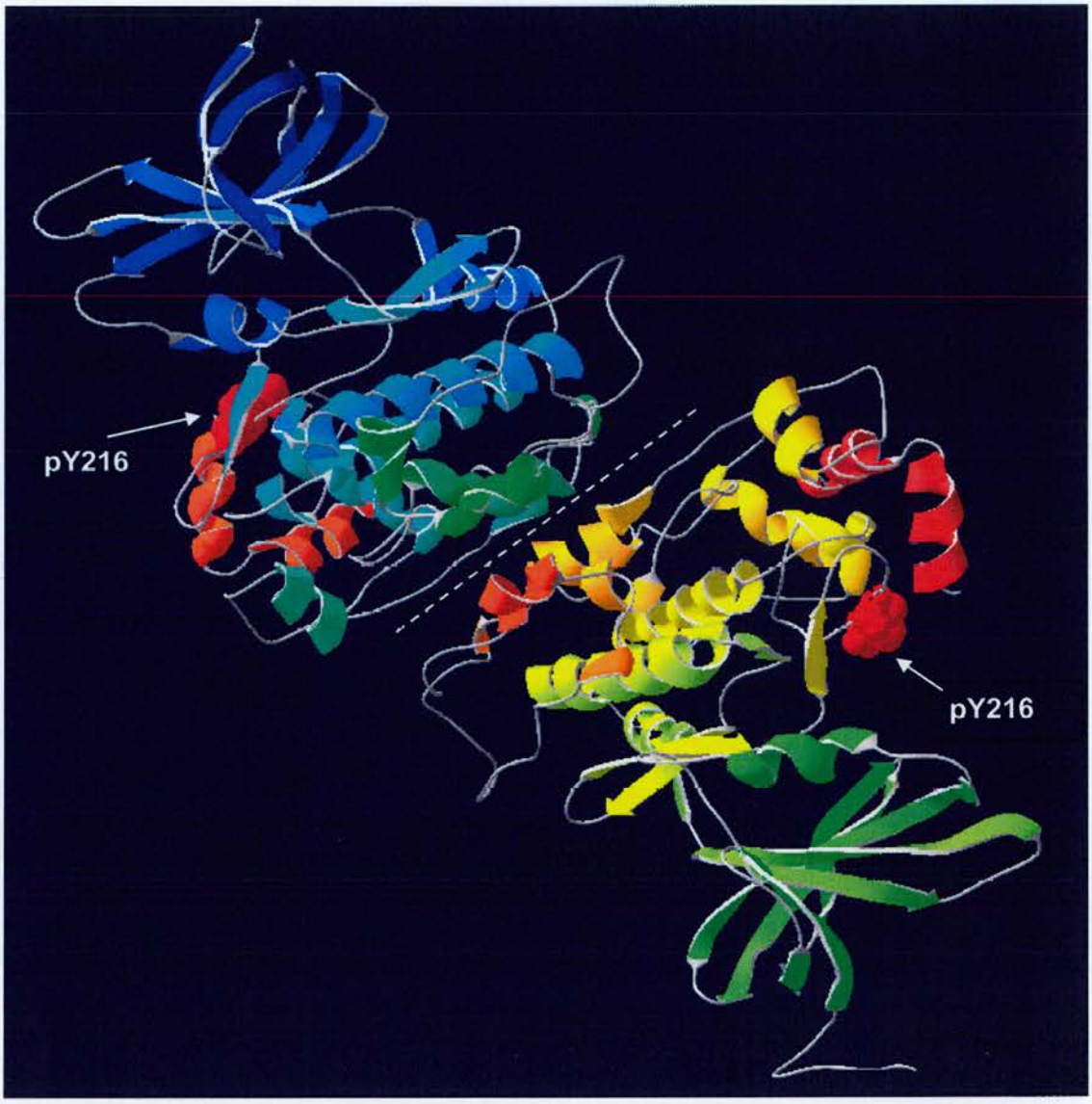


Figure 4.17 – X-ray crystal structure of a Glycogen Synthase Kinase-3 β dimer. This 3D model (1GNG) shows amino acids 36-384 of GSK3β that encompasses the protein kinase domain of the kinase dimer. The dashed white line indicates the dimer interface. Phosphotyrosine 216 which is indicated on both molecules is not essential for kinase activity but does have a regulatory effect on activity. It is proposed to be involved in docking of GSK3β substrates for phosphorylation.

determination (Dajani et al., 2001). (Figure 4.17). GSK3◉ exists as a dimer with pY216 freely accessible on both monomers and is proposed to interact with residues on the substrate thereby facilitating substrate docking prior to phosphorylation (Dajani et al., 2001).

Brain selective kinase 1 and 2 (BRSK1/2) are 2 closely related kinases (Identities = 337/399 (84%), Positives = 365/399 (91%)) in which we have characterised 16 phosphorylation events. Figure 4.18 shows a multiple sequence alignment of BRSK1 and 2 protein sequences from mouse and human and the potential conservation of these phosphorylation sites is summarised in Table 4.5. Of the 12 phosphorylation sites mapped in BRSK1 6 of the unambiguously assigned sites are completely conserved, three are conserved with a serine/threonine substitution. Four potential sites (2 pS/T) were also found in BRSK1 and of these one (S527) was completely conserved (potentially a good candidate for the real phosphorylation site) and the other three possibilities were only present in BRSK1.

Kinase	Phosphopeptide sequence	Mod	Site	Conserved	Conserved S/T	BRSK1 only
BRSK1	pSpTGLpSpSpSPLpSpSPRpSPVFSFSPEPGAGDEAR	8 pS/T	S437	✓		
			T438		✓	
			S441	✓		
			S442		✓	
			S443	✓		
			S446	✓		
			S447	✓		
			S450			✓
BRSK1	SSGGTPLHpSPLHTPR	1 pS + 1 pS	S511			✓
			S512			✓
			S519		✓	
BRSK1	ASPTGTPGTpTPPPSPGGGVGGAAWR	1 pS/T + 1 pT	S527	✓		
			T529			✓
			T535	✓		
BRSK2	pSMEVLSVTDGGpSPVPAR	2 pS	S383	✓		
			S394	✓		
BRSK2	SRSISGASSGLSTpSPLSSPR	1 pS + 1 pS	S419		✓	
			S422	✓		
			S424		✓	

Table 4.5 – Identification and conservation of phosphorylation sites in BRSK 1 and 2. Unambiguously assigned phosphorylation sites are indicated in red and ambiguously assigned sites in blue. Phosphorylation sites conserved in BRSK1 and 2 in the mouse and human proteins are indicated in the "conserved" field. Phosphorylation sites partly conserved (serine or threonine) in BRSK1 and 2 in the mouse and human proteins are indicated in the "conserved S/T" field. Finally, 4 phosphorylation sites were found in BRSK1 only as indicated.

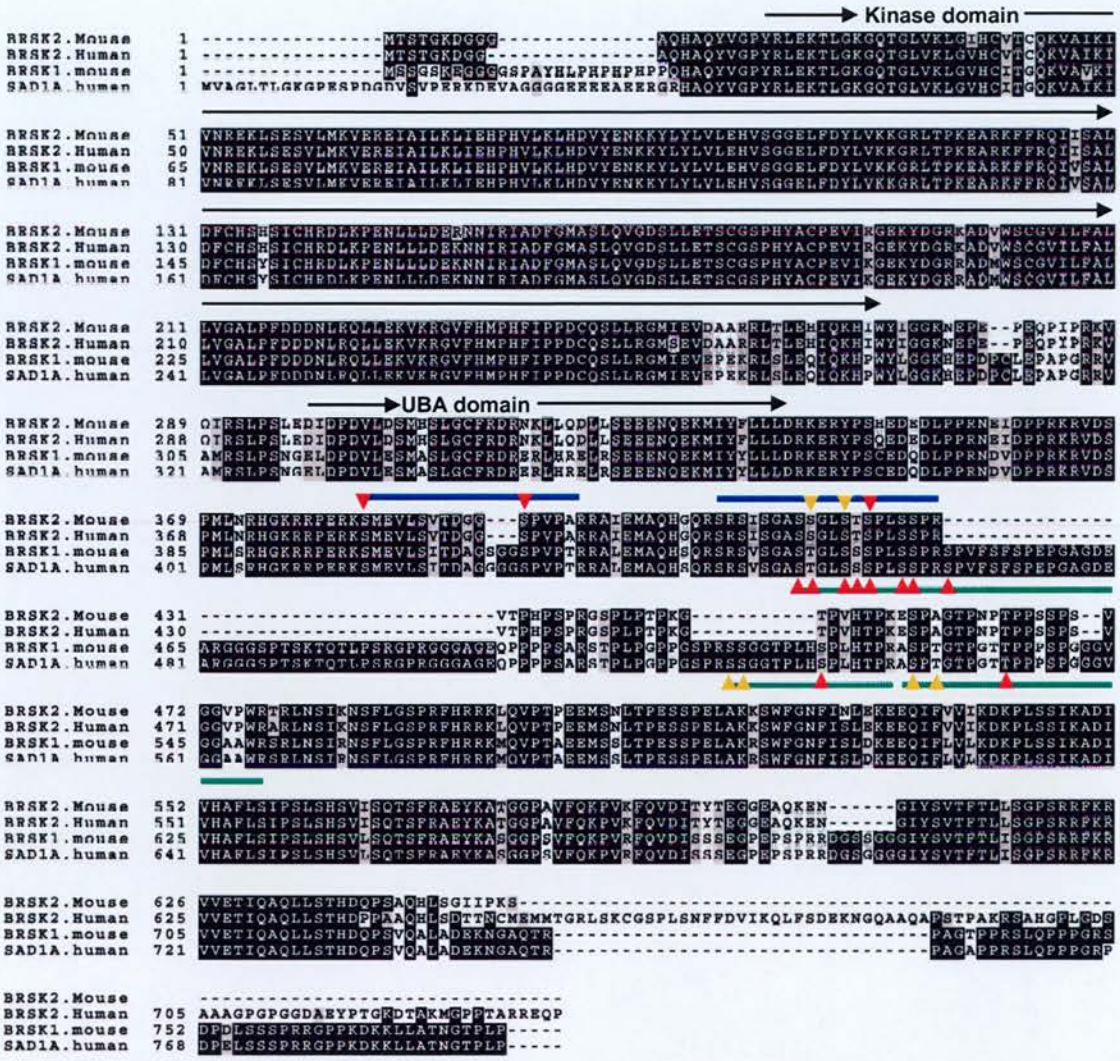


Figure 4.18 – Multiple sequence alignment of BRSK1 and 2 protein sequences from mouse and human. The human ortholog of BRSK1 is SAD1A. Phosphopeptides found in BRSK1 are indicated by the green lines and those found in BRSK2 are indicated by blue lines. Unambiguously assigned phosphorylation sites are indicated in red and ambiguously assigned sites in gold. All of the phosphopeptides are located N-terminal to the kinase and UBA domain.

Four phosphorylation sites were found in BRSK2, three were unambiguously assigned with two being completely conserved and one being an S/T substitution. One out of the two possibilities for the remaining ambiguous site is completely conserved (S442) which makes it the likely phosphorylation site. Only one of the 16 phosphorylation sites found on these two kinases is definitely specific (S450, in BRSK1 only) and may be important in functions specific to BRSK1.

4.4.6 Putative 14-3-3 binding motifs in cytosolic phosphoproteins

One function of phosphorylation is to provide a regulated binding site on proteins. There are many different types of such phosphorylation-dependent protein interactions but the motifs for prediction of most these interactions are not complete. However, 14-3-3 proteins which bind to phosphorylated serine or threonine residues have been intensively studied and their optimal binding motifs have been characterised. We investigated the potential of the phosphorylation sites identified in this study to be sites for 14-3-3 binding by screening each site using Scansite, a short sequence motif prediction server. 40 out of the 321 (12.5%) of the phosphorylation sites on we characterised by mass spectrometry are predicted by Scansite to be sites for 14-3-3 binding. When compared to our previous analysis of the mouse forebrain synapse phosphoproteome, in which 18 out of the 331 phosphorylation sites (5.4%) were predicted by Scansite to be sites for 14-3-3 binding, there is an apparent enrichment for this motif in this dataset. This may be accounted for by the differences in intrinsic disorder between these datasets (46% sequence disorder for cytosolic compared to 34% for synapse phosphoproteome, Figure 4.19)), as such phosphorylation-dependent binding motifs occur more frequently in disordered regions (ref). Indeed, 95% of the 14-3-3 binding sites predicted by Scansite occur in intrinsically disordered sequences.

All 14-3-3 binding sites predicted by scansite fall into three consensus motifs, Rxx[pS/T]xP, Rxx[pS/T]x[G/D/E/S/T] and Kxx[pS/T]x[P/G/D/E/S/T], except for 5

which are omitted from Table 4.6. Seven sites conform to the minimal motif Rxx[pS/T]xP while the majority of sites (66%) conform to the Rxx[pS/T]x[G/D/E/S/T]

A **Rxx[pS/T]xP**

Acc	Name	Binding site	14-3-3 m1	Kinase 1	Kinase 2	Kinase 3
P10637	Tau	RSRTP[S]LPTPP	0.2656	Akt (0.2979)		
O70495	Srrm1	KRRTA[S]PPPPP	0.109	Akt (0.3493)		
Q8BTI8	Srrm2	RARSR[T]PPSAP	0.0631			
Q8CIL3	Parcc1	PSRRS[S]QPSPT	0.3768	PKA (0.4574)		
Q9WV48	Shank1	QGRSM[S]VPDDA	0.2317	Akt (0.4455)		
Q80TS1	nArgBP2	GLRSP[S]PPRS	0.3321	Cdk5 (0.4697)	Cdc2 (0.5182)	
Q8VCD2	CHERP	SSRSP[T]PPSAA	0.3246	Akt (0.4835)		

B **Rxx[pS/T]x[G/D/E/S/T]**

Acc	Name	Binding site	14-3-3 m1	Kinase 1	Kinase 2	Kinase 3
P10637	Tau	GSRSR[T]PSLPT	0.1093	CaMKII (0.3087)		
P62996	SFRS10	YRRRH[S]HSHSP	0.5876	PKA (0.4018)	Clk2 (0.5341)	
Q5SQZ3	CaMK2G	KKRKS[S]SSVHL	0.53	PKA (0.5185)	CaMKII (0.5543)	Akt (0.6614)
Q8VE97	SFRS4	EPRAR[S]RSTSK	0.4017			
Q8VE97	SFRS4	RARSR[S]TSKSK	0.3151			
Q9R0U0	NSSR1	YERRR[S]RSRSF	0.4404	Clk2 (0.4194)	PKA (0.3803)	
Q9R0U0	NSSR1	RRRSR[S]RSFDY	0.3403	Clk2 (0.4233)	Akt (0.4451)	
O70495	Srrm1	ARRRR[S]PSPAP	0.1428	PKA (0.0398)		
O70495	Srrm1	IQRRY[S]PSPPP	0.1158	Akt (0.3761)	Cdc2 (0.2619)	ERK1 (0.415)
O70495	Srrm1	VRRGA[S]ASPQG	0.1193			
P62996	SFRS10	IYRRR[S]PSPYY	0.4961	PKA (0.3333)		
Q6PDH0	Phldb1	GRRTR[S]PSPTL	0.5055	Akt (0.4869)		
Q8BTI8	Srrm2	HSRSR[S]SSPDS	0.3365			
Q8BTI8	Srrm2	RERSS[S]ASPEL	0.0768	Akt (0.19)		
Q8BTI8	Srrm2	RKRSR[S]RSPLA	0.0896			
Q8VCD2	CHERP	RRRSR[S]RSPTP	0.315	Clk2 (0.4208)	Akt (0.4276)	
Q9R0U0	NSSR1	RSRSR[S]FDYNY	0.4602	Akt (0.2830)	Clk2 (0.4154)	PKC zeta (0.4174)
Q9WV69	Dematin	LERHL[S]AEDFS	0.5774	ATM (0.4350)	Clk2 (0.6472)	PKC mu (0.5141)
Q6NXJ5	Gm1568	TARSF[S]LGDLS	0.2867	Akt (0.4244)	PKA (0.5261)	
Q7Z417	82-FIP	LERND[S]WGSFD	0.4725	PKC mu (0.5283)	PKA (0.5788)	Akt (0.5301)
Q80YT1	Glcci1	HQRSA[S]WGSAD	0.3003	Akt (0.5009)	Clk2 (0.6182)	
Q6PDH0	Phldb1	RTRSP[S]PTLGE	0.4794	PKA (0.5115)		
Q8VCD2	CHERP	RSRSR[S]PTPPS	0.3855	Clk2 (0.4799)	Akt (0.4409)	

C **Kxx[pS/T]x[P/G/D/E/S/T]**

Acc	Name	Binding site	14-3-3 m1	Kinase 1	Kinase 2	Kinase 3
O70495	Srrm1	PPKRR[T]ASPPP	0.1477			
P14873	MAP1b	PAKSP[S]LSPSP	0.1446	GSK3 (0.6423)		
Q80UZ7	D430018P0	LKKSH[S]SPSLN	0.0952			
Q9WV69	Dematin	SPKST[S]PPPSP	0.4978	ERK1 (0.4387)	GSK3 (0.5211)	Cdk5 (0.5760)
Q5SQZ3	CaMK2G	VKKRK[S]SSVH	0.5985	PKA (0.4521)	PKC mu (0.5985)	CaMKII (0.5667)

Table 4.6 – Cytosolic phosphoproteome phosphorylation sites predicted to be 14-3-3 binding sites. 35 out of the 40 Scansite predicted 14-3-3 binding sites conformed to the three consensus sequences shown. The prediction score for the presence of 14-3-3 binding sites and phosphorylation of sites by kinases are shown.

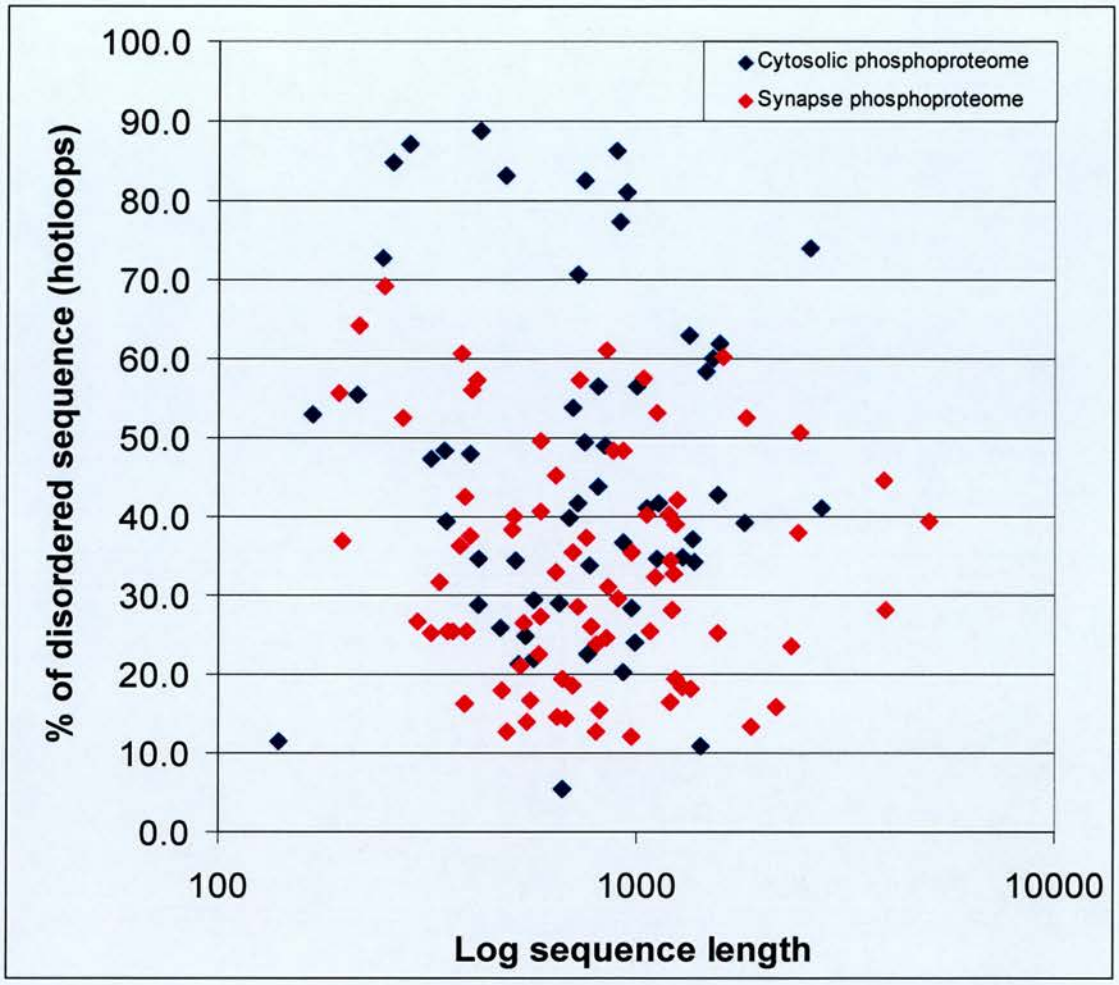


Figure 4.19 – Comparison of intrinsic protein disorder in cytosolic and synapse phosphoproteomes. The percentage of disordered sequence (hotloops) was calculated for each of these datasets and was plotted against log of protein sequence length. The average percentage disorder in the synapse phosphoproteome (34%) is 12% lower than that of the cytosolic phosphoproteome (46) indicating the presence of more structural proteins at the synapse. In addition, there are 11 proteins with more than 70% sequence disorder in the cytosolic phosphoproteome, whereas the synaptic phosphoproteome does not contain any proteins with 70% sequence disorder. The increased disorder in the cytosolic phosphoproteome may account for the 2-fold increase in predicted 14-3-3 binding sites in the cytosolic phosphoproteome compared to synapse phosphoproteome.

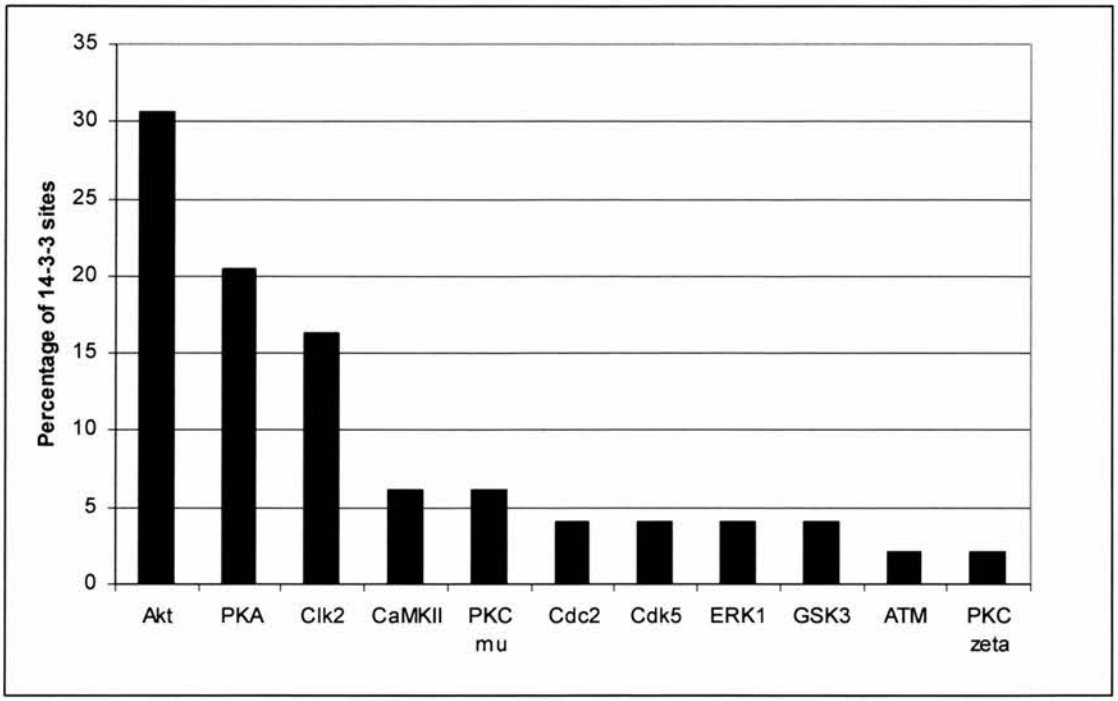


Figure 4.20 – Kinases predicted to phosphorylate putative 14-3-3 binding sites. Analysis of the 35 putative 14-3-3 binding sites in Table 4.6 by Scansite revealed that over 30% of these sites are predicted to be phosphorylated by Akt kinase. The substrate selection of Akt and PKA kinases are similar to the motif recognized by 14-3-3 proteins.

motif. Figure 4.20 shows the kinases predicted (by Scansite) to phosphorylate the putative 14-3-3 binding sites. It is known that basophilic kinases such as Akt, PKA and Clk2 have similar consensus motifs for phosphorylation as 14-3-3 proteins have for phosphorylation-dependent binding.

4.5 DISCUSSION

4.5.1 Analytical comparisons

We have applied our previously developed tandem metal affinity purification approach (Collins et al., 2005), (Collins et al., 2005) in combination with two different reversed-phase separations and analysis by tandem mass spectrometry to study phosphorylation in the brain. We have identified 485 phosphorylation events corresponding to 321 unique phosphorylation sites on 67 cytosolic phosphoproteins in the mouse forebrain. In our previous analyses of the synapse phosphoproteome we used a number of analytical approaches including single protein and single peptide IMAC as well as tandem IMAC, to identify phosphorylation sites. The most efficient approach was the combination of protein and peptide IMAC (tandem IMAC) and this resulted in identification of 308 phosphorylation events on 41 synaptic phosphoproteins.

To further refine this tandem IMAC methodology we sought to determine the optimal analytical reversed-phase separation for phosphopeptides. Efficient capture and separation of phosphopeptides is crucial for efficient tandem MS analysis and high phosphoproteome coverage. The chromatographic conditions used for reversed-phase (RP) separation of phosphopeptides are quite different from those used for cell-lysate digests for basic protein identification. The majority of phosphopeptides elute at 3.5% organic solvent compared to 32% for standard digests because phosphopeptides are more hydrophilic than non-phosphorylated peptides (Figure 2.5, (Collins et al., 2005)). Therefore a long shallow gradient of organic solvent is used for efficient phosphopeptide separation. However, the

hydrophilicity of phosphopeptides can cause them to elute from the RV too early for analysis by the mass spectrometer and therefore the binding properties of the RP column media may be important in such phosphoproteomic analyses.

We compared two different RP columns; PepMap C18 3 mm 100Å, 300 μm i.d. x 5mm and a BetaMax neutral 5μm, 180 μm i.d. x 30mm. The phosphopeptides purified from tandem IMAC purification were subjected to two identical experiments. This involved two 4-hour RP separations of the phosphopeptides on a BetaMax neutral (BM) column and a PepMap (PM) column with a hypercarb column as a trapping column in both experiments. The phosphopeptides separated in these two experiments were analysed on a Q-Tof Ultima with identical acquisition parameters. The PM experiment yielded greater coverage of the sample, with identification 401 phosphorylation events on 173 phosphopeptides compared to the BM experiment in which we identified 224 phosphorylation events on 110 phosphopeptides (Table 4.1). The identification of so many phosphorylation events in a single experiment indicated that the Tandem IMAC purification and MS acquisition settings were performing very well but that the subtle differences in the RP media were having a significant effect on the overall performance of the approach.

To investigate these RP media differences we compared the physiochemical properties of phosphopeptides characterised in both analyses. We found that the hydrophilicity of the phosphopeptides differed between the analyses in two ways. Firstly, GRAVY scores (hydrophobicity score) calculated for phosphopeptide sequences (effect of phosphate not calculated) indicated that the PM column was capturing more hydrophilic peptides compared to the BM column (GRAVY score of -0.943 compared to –0.762 for PM and BM specific phosphopeptides, respectively). Secondly, the PM column captured more highly phosphorylated peptides (and therefore, more hydrophilic) than the BM column, with an average stoichiometry of phosphorylation of 2.31 compared to 2.03 phosphates per phosphopeptides for the

PM and BM experiments, respectively (Table 4.1, Figure 4.4). The combination of these two calculations clearly indicate that the PM column is more applicable to separation of complex mixtures of phosphopeptides because it can capture many more phosphopeptides with a higher sequence hydrophilicity as well as more highly phosphorylated peptides.

4.5.2 Intrinsic protein disorder and its relevance to protein phosphorylation.

We investigated the relationship between protein phosphorylation and protein disorder in the cytosolic phosphoproteome by using DisEMBL to predict the percentage disordered sequence (hotloops) of each cytosolic phosphoprotein. These were then compared with a dataset of experimentally verified ordered proteins and disordered proteins (Figure 4.6). This resulted in the identification of 16 putative natively disordered proteins, which included three experimentally verified natively disordered proteins. The cytosolic phosphoproteome contained a range of proteins: some with very low sequence disorder (structured proteins), those that had regions of sequence disorder (linked-folded domains) to those which had high sequence disorder and are thus putatively unstructured proteins.

Next, we looked at the location of phosphorylation sites in terms of local regions of intrinsic disorder. 72% (231 phosphorylation sites) of the phosphorylation sites characterised in the cytosolic phosphoproteome are located in regions of sequence disorder. Only 5 out of the remaining 90 phosphorylation sites could be mapped to Pfam domains (structured regions), indicating that either DisEMBL under predicts unstructured sequence or there are tertiary structural types which have not yet been mapped to Pfam domains. Since serine and threonine residues are enriched in disordered sequence we asked whether intrinsically disordered proteins are phosphorylated to a greater extent than structured proteins. We compared the extent of phosphorylation with the extent of protein disorder in the cytosolic phosphoproteome dataset. We found a potentially positive relationship between

stoichiometry of phosphorylation and protein disorder; however this dataset is probably biased toward lower phosphoprotein stoichiometry due to difficulties in mapping all phosphorylation sites present in a given protein. A phosphoproteomic experiment targeted to the study of a small number of proteins with differing amounts of sequence disorder would be necessary to define this relationship further. Top-down mass spectrometry-based approaches using for example an LTQ-FTICR (linear ion trap fourier transform ion cyclotron resonance) mass spectrometer would allow measurement of the total number of phosphorylated residues in an intact protein, thereby permitting direct comparison of the total phosphorylation state of a number of proteins with varying intrinsic disorder.

We observed that in proteins with low sequence disorder that phosphopeptides tend to be located in the short regions of disorder that are present. For example, CamKIIγ has only 21% disordered sequence but the phosphopeptide we found for this kinase and another phosphopeptide in a homologous kinase are located consecutively in a region of intrinsic disorder and constitute 43 out of the 55 amino acids of this disordered region (Figure 4.9). The identification of 2 phosphopeptides from this discrete region of a mostly structured protein led us to ask if there were was something specific to disordered sequences which make them more amenable to our analytical approaches.

4.5.3 Enzymatic accessibility of intrinsically disordered regions.

In addition to the enrichment of phosphorylatable residues in disordered sequence, we noticed that arginine and lysine residues are also enriched. As trypsin is the enzyme of choice to digest protein samples for analysis by mass spectrometry, we investigated the frequencies of proteolytic cleavage by trypsin (*in silico* digestion) of proteins in the cytosolic phosphoproteome dataset. Figure 4.10 demonstrates that proteins with a higher percentage of intrinsic disorder produce shorter tryptic fragments. As disordered sequences are enriched for argine and lysine there are

more potential sites of cleavage for trypsin than in structured regions which have lower frequencies of these amino acids. It can be seen in Figure 4.11 that structured peptides in an in silico digest of MAP4 tend to be longer than unstructured peptides. In fact, a structured tryptic peptide of 57 amino acids long could be produced from tryptic digestion of MAP4 and thus mapping of potential phosphorylation sites on this peptide is very unlikely using current shotgun proteomic approaches.

On one hand we have observed and it is generally accepted, that the majority of phosphorylation sites occur in regions of intrinsic disorder and not very often in structural domains. However, on the other hand, proteolytic cleavage of structural domains is more difficult and produces longer tryptic fragments that may be too large for analysis by traditional MS approaches. Is it possible that due to this uneven distribution of cleavage sites in proteins, that we are under-sampling the extent of phosphorylation of structural domains? A systematic study of sequence motifs associated with phosphorylation sites has suggested that they occur predominantly within intrinsically disordered regions (Iakoucheva et al., 2004). This analysis led to predictions that only 1% of ordered serine and threonine residues compared to 6% of ordered tyrosine residues may be phosphorylated. Thus it would seem that tyrosine residues may be phosphorylated in intrinsically disordered and surface exposed ordered states. Of the 2 unambiguously assigned tyrosine phosphorylation sites found in the cytosolic phosphoproteome, one was located in a disordered region in SFRS10 and the other was in the ordered kinase domain of GSK3β. It is known that regions of intrinsic disorder are strongly favoured as sites of protease digestion (Iakoucheva et al., 2004) as ordered regions (in non denaturing conditions) cannot fit into the active site of proteases without disordering at lease 12 residues. In a similar manner intrinsic disorder seems to be necessary for most kinase substrates as hydrogen bonding with kinase partners, which is necessary for phosphorylation, is not possible if the substrate is in a state of disorder (McDonald and Thornton, 1994). Furthermore, the lack of enrichment of tyrosine residues in disordered

sequence (Linding et al., 2003) and the increased likelihood of ordered tyrosine phosphorylation may be explained by the decreased requirement of tyrosine kinases for docking interactions with the substrate that are separate to the site of phosphorylation (Holland and Cooper, 1999). The majority of such docking interactions are associated with serine/threonine kinases, which according to intrinsic disorder predictions, phosphorylate mostly disordered sequences that are in a conformation suitable for docking interactions. Thus, it is possible that tyrosine phosphorylation may occur on partly ordered surface-exposed loops or regions that undergo an order to disorder transition just prior to association with the kinase. In these situations, backbone hydrogen bonding potential is sufficient for interaction with the active site but a secondary docking interaction with the substrate would not be required.

As regions of intrinsic disorder are strongly favoured as sites of protease digestion in native conditions because of accessibility and under denaturing conditions because of more cleavage sites for trypsin, these regions will produce more proteolytic fragments than ordered regions. Production of smaller tryptic fragments is problematic for analysis by traditional MS approaches and it appears that this may occur in disordered sequence. We compared the molecular weight of tryptic peptides from an in silico digest of Srrm1 (86.3% disorder) to that of Synapsin I (5.4% disorder) (Figure 4.11). We found that there are many more phosphorylatable (containing serine/threonine/tyrosine) peptides in Srrm1 outside the mass range of detection than for Synapsin I. 64% of phosphorylatable peptides from Srrm1 compared to only 24% of phosphorylatable peptides from Synapsin I were below 800 Da and thus not amenable to analysis by traditional shotgun MS approaches. It appears that intrinsic protein disorder may have a significant influence on protein sequence coverage and mapping of post-translational modifications using current tryptic digestion and subsequent shotgun MS-based analysis. We suggest that this should be investigated by using proteases with cleavage sites that differ to trypsin

or by analysis of whole proteins or protein domains by the top-down MS-based approaches described above.

4.5.4 Intrinsic protein disorder and phosphoprotein function

Intrinsically disordered proteins or regions of intrinsic disorder in proteins are associated with many cellular functions, including regulation of transcription and translation, signal transduction and protein phosphorylation and assemble of multi-protein complexes such as the ribosome (Dyson and Wright, 2005). Intrinsic protein disorder has evolved from relatively low levels in bacteria (2%) and viruses (7%) to between 18 and 32% in C. *elegans* and D. *melanogaster*, respectively (Iakoucheva et al., 2004)). It has been suggested that intrinsically unstructured proteins evolve by repeat expansion and two of the putatively unstructured phosphoproteins identified in this study (MAP2 and Tau) have statistically significant satellite regions (Tompa, 2003). The reason for positive evolutionary selection appears to be that unstructured proteins are capable of more versatile molecular functions compared to structured proteins. Their unstructured or flexible conformation may permit interaction with multiple binding partners at once and have many more accessible sites for post-translational modifications (Linding et al., 2003). Unstructured proteins have little or no hydrophobic core and as a consequence more of the sequence can form binding interfaces compared to structured proteins, which are limited by their reduced surface area. It has been suggested that they allow large intermolecular interfaces while keeping the sizes of proteins, cells and genomes small (Gunasekaran et al., 2003) (Figure 4.21).

4.5.4.1 Disordered phosphorylation in the splicesome.

The majority of putative phosphorylated IDP's found in this study (Table 4.3) are involved in RNA binding and RNA splicing. Alternative mRNA splicing plays an important role in development and differentiation allowing many polypeptide

products, with potentially different functions, to be produced from a single gene. Alternative splicing occurs on spliceosomes, which are complex particles composed of small nuclear ribonucleoproteins (snRNPs) and non-snRNP proteins (Figure 4.22). Five of the most disordered phosphoproteins found in the cytosolic phosphoproteome belong to the SR family of non-snRNP splicing factors (SRRRM1, SRRM2, NSSR1, SFRS4 and SFRS10) and we have mapped 65 phosphorylation sites on these proteins. These proteins are characterised by the presence of arginine- and serine-rich (RS) domain and an RNA recognition motif. The RS domain acts as a splicing-activation domain and it can enhance splicing when tethers to a pre-mRNA via a heterologous RNA-binding domain (Graveley and Maniatis, 1998).

In addition, to a protein binding function, the RS domain appears to be capable of binding directly to double-stranded RNA (Valcarcel et al., 1996). It has also been suggested that the positively charged arginine residues in the RS domain interact with, and neutralise the negatively charged phosphate backbone of mRNA (Valcarcel et al., 1996), enabling interaction with a wide variety of mRNA sequences. A characteristic feature of SR proteins is extensive serine phosphorylation , which is essential to their function in early stages of spliceosome assembly (Cao and Garcia-Blanco, 1998), (Xiao and Manley, 1997). Structural analysis of an SR protein kinase (Sky1p) revealed that it phosphorylates SR proteins constitutively and in particular phosphorylates Arg-Ser repeats sequentially, resulting in high stoichiometry of phosphorylation (Nolen et al., 2001). Such widespread phosphorylation of serine residues in SR proteins is necessary to complete spliceosome assembly and RS-domain phosphorylation can regulate the activity of SR proteins (Cao et al., 1997), (Graveley, 2000).

The exact mechanism of this regulation is unknown but two hypotheses have been put forward; one that involves regulation of RS domain-protein interactions and another that involves regulation of RS domain-mRNA interactions, both of which are compatible with each other. Hypophosphorylated RS domains are more likely to

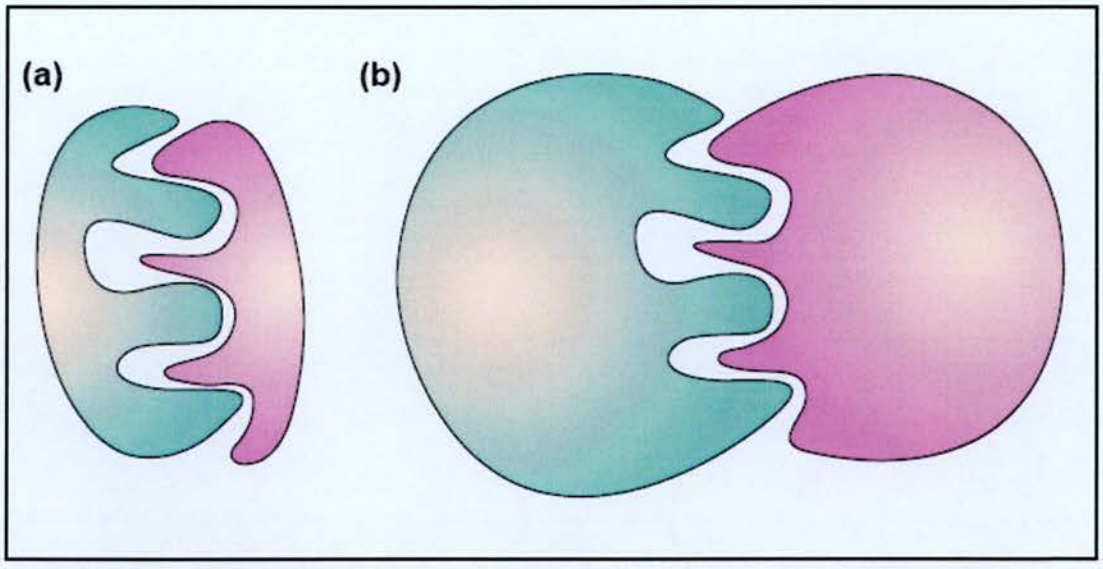


Figure 4.21 – Schematic representation of protein dimers. A represents interaction of disordered monomers and B represents interaction of 2 ordered monomers. The intermolecular interface is the same in both examples but the size of the ordered monomers is much larger than the disordered monomers. This represents the current view of why disordered proteins have been positively selected by evolution.

Adapted from Gunasekaran K, Tsai CJ, Kumar S, Zanuy D, Nussinov R. 2003. Extended disordered proteins: targeting function with less scaffold. Trends Biochem Sci 28:81-85.

interact with RNA, due to charge considerations, whereas phosphorylated RS domains might participate in protein interactions with other RS domains via electrostatic interactions between positively charged argine residues and negatively charged phosopho-serines.

SRRM1 (SRm160) and SRRM2 (SRM300) are two SR proteins and we characterised 19 and 33 phosphorylation sites on these proteins respectively. In addition, we found that they are predicted to be highly disordered, with 86.3 and 74.0% disordered sequence, respectively. The two proteins form the splicing coactivator, which functions in splicing by promoting critical interactions between splicing factors bound to pre-mRNA (Blencowe et al., 2000). SRRM1 is a component of the exon junction complex (EJC) that is involved in RNA export and has an (ATP)–dependent mobility (Wagner et al., 2004).

It has been reported that SRRM1-containing RNA export, but not splicing, complexes have an ATP-dependent mobility and that these RNA export complexes have an ATP-regulated mechanism for release from binding sites at splicing speckled domains (Wagner et al., 2004). It is likely that this ATP-dependent mobility is actually a dependency for extensive phosphorylation of the RS domain, which is proposed to promote proteins interactions, perhaps recruiting proteins associated with RNA export function. Thus, SRRM1 phosphorylation could be responsible for the regulated export of bound mRNA transcripts, as well as functioning in spliceosome assembly. SRRM2 is present in spliceosomes, bound to SRRM1, but is not part of the exon junction complex and was not found in the ATP-dependent RNA export complexes (Wagner et al., 2004). SR-cyp (Cyclophilin G) is another phosphoprotein which we have characterised and it is also highly disordered, with 82.5% predicted sequence disorder.

SR-cyp interacts with Clk1, a SR protein kinase which hyperphosphorylates SR proteins causing their localisation to change from nuclear speckles (zones of

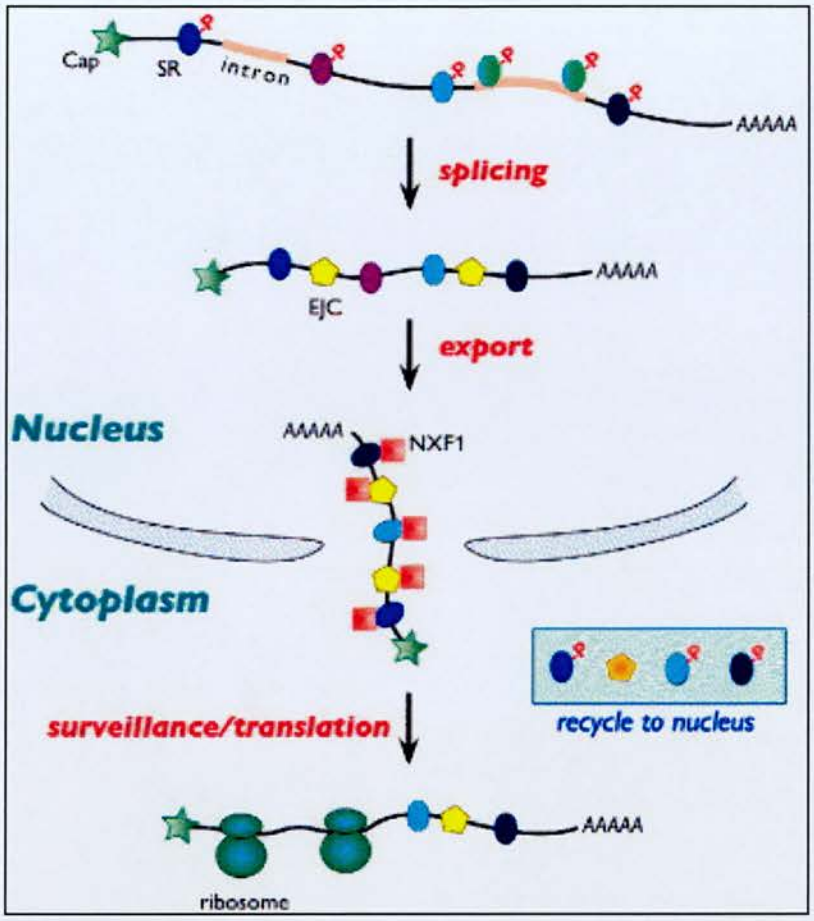


Figure 4.22 - Multiple Roles of SR Proteins in mRNA processing. Hyperphosphorylated SR proteins are recruited to pre-mRNA molecules, where they participate in splicing. SR proteins bind to the branch site and 5′ splice site of the intron during spliceosome assembly and splicing results in the dephosphorylation of bound SR proteins and in the deposition of exon junction complexes (EJC) upstream of exon-exon boundaries in the spliced RNA. Hypophosphorylated forms of the shuttling SR proteins then help to recruit the export adapter NXF1, which orchestrates passage of the mRNP through the nuclear pore. Translation is accompanied by RNP remodeling, leading to the release and reimport of rephosphorylated SR proteins and some EJC components into the nucleus.

Adapted from "Sorting out the complexity of SR protein functions." Graveley BR. RNA. 2000 Sep;6(9):1197-211.

accumulation of transcriptional and mRNA splicing factors) to a diffuse nucleoplasmic localisation (Colwill et al., 1996). Similarly, SR-cyp regulates the localisation of SR proteins, including SRRM2, re-distributing them from nuclear speckles to a diffuse nucleoplasmic localisation (Lin et al., 2004). SR-Cyp contains a PPPIase domain, which could catalyse the conversion of *cis*-proline to *trans*-proline in a polypeptide chain, thereby regulating protein folding and/or the conformational state of SR proteins. Another member of the SR protein family (pinin) interacts with SR-cyp via its RS domain, indicating that this may be a common mechanism of SR-cyp binding to SR proteins.

We have also characterised phosphorylation sites on another SR protein kinase, PRP4 (56.5% disorder). Clk1 has been shown to bind to PRP4 and phosphorylate it at its N-terminus, possibly at N-terminal sites which we have mapped (S143, S145) (Kojima et al., 2001). Overexpression of Clk1 causes re-distribution of PRP4 from nuclear speckles to a diffuse nucleoplasmic localisation, again in a similar manner as that described for other SR proteins (Kojima et al., 2001). There is evidence to suggest that Clk1 is not a direct regulator of SR proteins and that PRP4 is more active than Clk1 in phosphorylating SR-proteins (Kojima et al., 2001).

Neural specific SR protein NSSR 1 (SRp38) was found in the cytosolic phosphoproteome and is also highly disordered with 84.7% sequence disorder. This SR protein appears to be important in neuronal differentiation, possibly via regulating the neural-specific alternative splicing of genes (Liu et al., 2003). Unlike other SR proteins, SRp38 does not copurify with other SR proteins using a classical purification protocol (Zahler et al., 1992). Also, unlike other SR proteins, SRp38 cannot activate splicing in extracts lacking other SR proteins. However, dephosphorylation of SRp38 converts it into a potent, general repressor that inhibits splicing at an early step (Shin and Manley, 2002). Furthermore, this dephosphorylation occurs specifically in mitotic cells, presumably to ensure that inappropriate gene expression does not occur during mitosis (Shin and Manley,

2002). Upon heat shock treatment, SRp38 is rapidly dephosphorylated and therefore activated to repress splicing to ensure that inaccurate spicing does not occur until the heat-damaged splicing apparatus is restored (Shin et al., 2004).

Two other SR proteins, SFRS4 (SRp75) and SFRS10 (Tra2B) are highly disordered phosphoproteins with 83.0 and 87.2% sequence disorder, respectively. SFRS4 interacts with SRRM2 and forms a multiprotein complex which may be involved in pre-mRNA processing (Zimowska et al., 2003). SFRS10 can autoregulate its own protein concentration by autoregulating alternative splicing of its own pre-mRNA (Stoilov et al., 2004). As with other SR proteins, hyperphosphorylation of SFRS10 strongly reduces its binding to RNA and as a consequence is under the regulatory control of SR kinases. SFRS10, like SRp38 is involved in neuronal differentiation and it regulates the splicing of FGF-2R and GluR-B minigenes.

It should be noted that the SR proteins and other splicing-associated phosphoproteins characterised in this study were isolated from the cytosolic fraction, supporting a function in RNA export, where they would shuffle from the nucleus to the cytoplasm. In addition, as these proteins were found in their phosphorylated state, their presence in RNA export complexes or their cytoplasmic shuttling is likely to be regulated by phosphorylation. It has been shown that dephosphorylation of SR proteins is critical for determining the selectivity and efficiency of TAP (an export receptor)-dependent mRNA export (Huang et al., 2004). In fact, RS domain phosphorylation stimulates the nuclear import of mammalian SR proteins (Lai et al., 2000) after release from mRNA export complexes. Thus, the phosphorylation state of SR proteins found in this study can be temporally mapped to just prior to nuclear import, either after their translation or after release from mRNA export complexes.

The current evidence seems to suggest that SR-proteins such as SRRM2 are extensively regulated by phosphorylation by SR kinases such as Clk and PRP4 and

by binding or folding by chaperones such as SR-cyp, both of which function to regulate the function and localisation of splicing factors. SRRM1 is also regulated by phosphorylation, which is necessary for spliceosome assembly and RNA export, and it interacts with RNA as well as proteins in its function as a splicing factor. The demands placed on SR proteins in terms of molecular binding capabilities and extensive regulation by phosphorylation would not be satisfied by ordered proteins. These examples of highly disordered and phosphorylated proteins provide insight into how intrinsic disorder can facilitate both protein and RNA binding and in conjunction with extensive phosphorylation and interaction with other disordered chaperones such as SR-cyp, participate in highly regulated processes such as RNA splicing and export. SR proteins are essential splicing factors present in all metazoans but interestingly are completely absent in lower eukaryotes such as S. *cerevisiae* (Graveley, 2000). It is tempting to speculate that the appearance of SR proteins in higher eukaryotes is related to the parallel positive evolutionary selection for intrinsic protein disorder, a likely requirement for SR protein function.

4.5.4.2 Phosphoprotein disorder in the regulation of the cytoskeleton

Three phosphorylated components of the cytoskeleton which are known to be disordered or contain disordered sequences, were found in the cytosolic phosphoproteome; TPPP (Tubulin polymerisation promoting protein) (64.2% disorder), Tau (Microtubule-associated protein Tau) (57.4% disorder) and Beta-adducin (41.7% disorder) (Figure 4.6, Table 4.3). Beta-adducin binds to spectrin-actin complexes and promotes the association of actin with spectrin (Hughes and Bennett, 1995). The C-terminal tail of beta-adducin is intrinsically disordered and was identified as the site for binding to spectrin-actin complexes (Hughes and Bennett, 1995). We have identified 6 phosphorylation sites on 2 consecutive phosphopeptides spanning 53 amino acids of this unstructured C-terminal tail. This multi-site phosphorylation cluster is well placed to regulate interaction with

spectrin-actin complexes, N-terminal to PKC phosphorylation sites that inhibit beta-adducin's activity in promoting spectrin-actin complexes (Matsuoka et al., 1998).

TPPP was originally identified as a natively unfolded protein (Orosz et al., 2004) but has subsequently been found to be partially ordered with an extended structure (Otzen et al., 2005). TPPP stimulates aberrant tubulin polymerisation which gives rise to microtubule assemblies in inclusion bodies of human pathological brain tissues such as Alzheimer's and Parkinson's disease. We have characterised four phosphorylation sites in TPPP, two of which are located in the unstructured N-terminus, which is missing in shorter forms of TPPP encoded by two separate genes in mammals (Tirian et al., 2003). A phosphorylation sites motif (TPPKSP) within this region is present in Tau and is phosphorylated by Cdk5 (Takahashi et al., 1991). We found both phosphorylation sites in the motif (pTPPKpS) in Tau and also in TPPP and this shared motif may have a similar function in TPPP as it does in Tau.

Tau is a highly phosphorylated protein, with 17 out of a total of 35 phosphorylation sites characterised in this study (Figure 4.8). Phosphorylated Tau has been implicated by many studies in the pathology of Alzheimer's disease (AD), a neurodegenerative disease characterised by deposits of amyloid A-beta peptides in plaques and by Tau deposits in the form of paired helical filaments (PHF's). Most neurodegenerative disorders are disorders of protein folding and the consequences of protein folding, and are therefore classified as foldopathies. Tau interacts with alpha and beta tubulin, as well as other microtubule-associated protein tau isoforms to promote microtubule assembly and stability (Lee et al., 1989). Mutations in exon 10 of the Tau gene are associated with FTDP-17, an autosomal dominant hereditary neurodegenerative disorder (Jiang et al., 2003). These mutations lie in an enhancer region, which SFRS10 (an intrinsically disordered SR protein characterised in this study) binds and regulates the alternative splicing of exon 10 (Jiang et al., 2003). Exon 10 encodes one of four microtubule-binding motifs and aberrant splicing of this exon and has implicated SRFS10 and the increased affinity of Tau for

microtubules in the pathogenesis of tauopathies (Kondo et al., 2004), (Jiang et al., 2003).

The region surrounding the microtubule-binding repeats is highly phosphorylated by multiple kinases and is thought to regulate microtubule binding (Goode et al., 1997). Hyperphosphorylated Tau shows defective microtubule binding and fails to promote microtubule assembly. (Mandelkow, 1999). Tau is natively disordered (Barghorn et al., 2004), especially either side of the microtubule-binding repeats, where most phosphorylation is concentrated (Figure 4.6). Hyperphosphorylated Tau is thought to cause aggregation and the formation of PHF's and it has been shown that Tau displays an increased level of beta-structure in PHF's (Barghorn et al., 2004). It therefore seems likely that the combination of disorder and phosphorylation surrounding the microtubule-binding repeats is important in PHF formation. The folding or conformation of Tau is regulated in a manner analogous to that described for SRRM2, in which a chaperone (SR cyp) regulates SRRM2 conformation by peptidyl-prolyl *cis/trans* isomerisation of the polypeptide chain. Pin1, a PPIase, specifically regulates the conformation of proline-directed phosphorylation sites (Lim and Lu, 2005). Pin1 targets phosphorylated Tau on a Thr321-proline motif and restores its biological function after Tau is inactivated by hyperphosphorylation (Lim and Lu, 2005). Pin1 allows Tau dephosphorylation by PP2A and restoration of normal function. This pathway is characterised by a combination of protein disorder and phosphorylation, which is necessary for normal regulation of Tau function, and is regulated by a specific chaperone which can modify Tau protein conformation and allow restoration of Tau function by dephosphorylation by PP2A. It highlights the importance of regulation of protein disorder and phosphorylation in protein function and how their deregulation may be involved in foldopathies.

4.5.4 Phosphorylation and human brain disorders

The dysregulation of cellular signalling pathways is fundamental to many pathologies and especially in neurodegenerative diseases where aberrant phosphorylation of key proteins such as Tau has been implicated in disease formation and progression. The identification of sites of phosphorylation or mapping of the normal phosphorylation state of a phosphoproteome is necessary to understand how changes in the phosphorylation state of proteins can lead to disease. In addition to mapping many phosphorylation sites on Tau and TPPP, which are implicated in AD, we have also identified phosphorylation sites on other neurodegenerative disorder-associated proteins. Figure 4.23 shows a portion of the AD protein interaction network illustrated in Figure 5.16, highlighting phosphoproteins found in the cytosolic phosphoproteome.

GIT1, a G protein coupled receptor kinase-interacting protein has numerous cellular functions from acting as a scaffold for the kinases ERK1/2 and MEK1 in focal adhesions (Yin et al., 2005) to being part of a signalling complex which includes PIX, Rac and PAK, that plays an essential role in the regulation of dendritic spine and synapse formation (Zhang et al., 2005). Interestingly, when a protein-protein interaction network for Huntington's disease (HD) was constructed from a yeast two-hybrid screen it was discovered that GIT1 directly interacted with huntington protein (HTT) (Goehler et al., 2004). In addition, GIT1 is localised to neuronal inclusions and is selectively cleaved in HD brains, indicating a role in HD pathogenesis (Goehler et al., 2004). GIT1 also interacts with CRMP1 (Collapsin response mediator protein 1) (Goehler et al., 2004), which directly interacts with HTT (Goehler et al., 2004). CRMP1 in turn interacts with CRMP2 (Collapsin

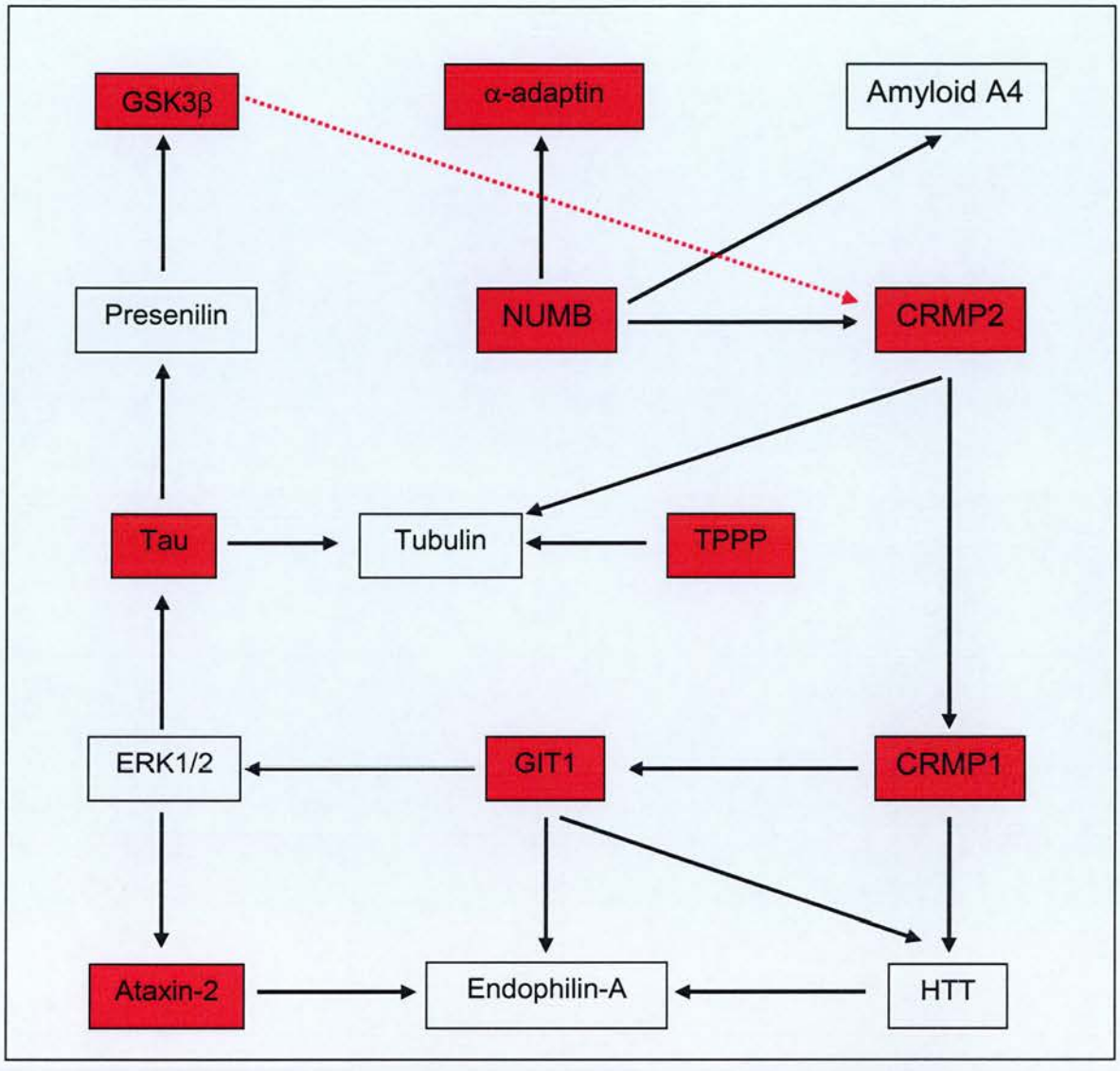


Figure 4.23 – Phosphoproteins involved in neurodegenerative disease. A selected set of proteins involved in AD or HD are shown with arrows representing protein interactions (red arrow indicates phosphorylation interaction). Red-filled boxes indicate phosphoproteins characterised in this study. Protein interaction data was mined from the literature and the following PMIDs describe the interactions shown above: 12942088, 12011466, 12942088, 15383276, 15923189, 10515232, 9689133, 12093283, 15676027

response mediator protein 2), both of which we characterised in the cytosolic phosphoproteome. Increased GSK3◎ activity is associated with AD and CRMP2 has been found to be a physiological substrate for GSK3◎◎(Cole et al., 2004). A hyperphosphorylated region of CRMP2 is an Alzheimer epitope and is physically associated with neurofibrillary tangles (Cole et al., 2004), (Gu et al., 2000). We have characterised this phosphorylated AD epitope and found that it contained 7 phosphorylated residues in a stretch of 16 amino acids, three of which are novel. It appears that phosphorylation of this epitope regulates CRMP2 binding to tubulin and that GSK3◎◎is at least partly responsible for this regulation (Uchida et al., 2005). Also, GSK3◎◎phosphorylation of CRMP2 regulates axon elongation in primary neurons, possibly by promoting microtubule assembly (Cole et al., 2004).

Numb, a phosphotyrosine-binding domain-containing protein was originally identified as a Notch signalling antagonist in Drosophila, where is determines cell fate during nervous system development (Nishimura et al., 2003). Numb interacts with CRMP2 and ◎-adaptin (both in cytosolic phosphoproteome) with a role in endocytosis (Nishimura et al., 2003). CRMP2, in addition to regulating microtubule assembly, regulates Numb-mediated endocytosis for axon growth (Nishimura et al., 2003). In fact, isoforms of Numb have the ability to increase vulnerability of cells to the amyloid beta peptide, possibly by affecting notch-mediated cell fate. Whether the interaction of CRMP2 with Numb has a role in Numb-mediated cell death in response to amyloid beta peptide has yet to be determined, but with CRMP2 also regulating cell fate decisions it would seem likely that they are involved in a common pathway.

The Ataxin-2 gene contains a trinucleotide repeat that encodes a polyglutamine stretch, which when expanded causes spinocerebellar ataxia type 2 (SCA2) (Ralser et al., 2005). SCA2 belongs to the family of polyglutamine disorders, which includes HD, and there is evidence to show that HTT and Ataxin-2 are functionally linked.

They both interact with Endophilin A complexes and are involved in pathways regulation actin organisation (Ralser et al., 2005).

As well as characterisation of phosphoproteins implicated in neurodegenerative disorders we have found phosphorylation sites on 82-FIP, a FMRP (Fragile X Mental Retardation Protein) interacting protein. The absence of FMRP causes fragile-X syndrome, the most common monogenic form of mental retardation. FMRP is an RNA-binding protein and was observed to involved in translational control (mRNA trafficking) at a postsynaptic level (Weiler et al., 1997), (Mazroui et al., 2002). 82-FIP binds specifically to FMRP and is associated with polyribosomes (Bardoni et al., 2003). The distribution of 82-FIP is cell-cycle dependent, being cytoplasmic in the G2/M phase and nuclear in the G1 phase (Bardoni et al., 2003). It has been proposed that this cell-cycle dependent localisation of 82-FIP is regulated by phosphorylation in an analogous manner to dFMRP in drosophila (Bardoni et al., 2003), (Siomi et al., 2002). We have characterised a phosphorylation site on 82-FIP at serine 652 that is predicted to be a site for 14-3-3 binding and may be involved in the observed regulated localisation of this protein.

4.5.6 Conclusions

We have applied our previously developed tandem IMAC purification strategy and analysis by mass spectrometry to study the mouse forebrain cytosolic phosphoproteome. This led to identification of 485 phosphorylation events corresponding to 321 unique phosphorylation sites on 67 cytosolic proteins. We compared the suitability of two reversed-phased columns for the analytical separation of phosphopeptides and found that in a single 4 hour LC run with a PepMap column we could identify 401 phosphorylation events on 173 phosphopeptides. This data yield from a single MS analysis highlighted the efficiency of our approach and application of this approach to a fractionated (e.g.

ion-exchange) phosphoproteome with multiple MS analyses would certainly yield increased coverage.

Analysis of the types of phosphoproteins characterised in this study led to a number of interesting findings. Firstly, as this dataset is derived from a very soluble fraction of the mouse brain proteome, we reasoned that it was likely that these proteins were quite hydrophilic and therefore more likely to be unstructured proteins. Intrinsic disorder predictions confirmed this and allowed identification of a set of putative natively unstructured or intrinsically disordered phosphoproteins. These proteins were mostly RNA or tubulin-binding proteins and further investigation revealed that a combination of disorder and phosphorylation is very likely essential for their function. Secondly, it is known that phosphorylation is more likely to occur in regions of disorder and we investigated the distribution of phosphorylation sites in each of the identified phosphoproteins. We found that proteins with higher sequence disorder seem to be phosphorylated to a greater extent and that in structured proteins and phosphorylation sites usually occur in local regions of disorder. Thirdly, disordered proteins are more susceptible to proteolytic cleavage and we found that a clear relationship between sequence disorder and frequency of tryptic digestion exists, which may impact on the coverage of certain smaller phosphopeptides in phosphoproteomic analyses. Fourthly, we identified 37 phosphorylation sites on nine protein kinases, the majority of which are novel and were not in kinase domains indicating secondary regulatory functions. In addition, 35 phosphorylation sites in the cytosolic phosphoproteome are predicted to be 14-3-3 binding sites, more than a two-fold increase compared to that predicted for the synapse phosphoproteome. This increase in predicted 14-3-3 binding sites was accompanied by increased sequence disorder in the cytosolic phosphoproteome supporting the argument that disorder confers increased molecular regulation. Finally, we have discussed the relationship between phosphorylation and disorder in brain disease and how in a global study of phosphorylation we have

characterised a number of phosphoproteins with disease-associations, which may have related roles in the pathogenesis of neurodegenerative disease.

CHAPTER 5

MOLECULAR ORGANISATION OF THE POSTSYNAPTIC PROTEOME REVEALED BY INTEGRATIVE BIOINFORMATIC ANALYSIS

5. Molecular organisation of the postsynaptic proteome revealed by integrative bioinformatic analysis

5.1 ABSTRACT

Identification of the composition of the synapse proteome provides a framework for understanding the overall organisation, development, evolutionary origins and function of the synapse in normal and pathological conditions. In addition, proteome data allows the derivation of proteome-wide molecular network maps at the level of gene regulation, protein interaction and protein phosphorylation that can be merged with functional or phenotype data from individual molecule studies to derive new models of synapse function. Proteomic analysis has facilitated identification of the components of the postsynaptic density, AMPAR complex and MASC (MAGUK-associated signalling complexes) (Collins MO. J. Neurochem 2005). We have extracted data from six other synapse proteome experiments and combined these with our data to provide a consensus on the composition of the post-synaptic proteome (PSP). In total, 1124 proteins are present in the PSP with 466 of these validated by their detection in two or more studies, forming what we have designated the "Consensus PSD" (cPSD). We have performed bioinformatic analysis of each level of complexity in the PSP; gene evolution, mRNA expression, protein annotation, phosphorylation, protein complexes and protein networks. We have linked these data types and have observed molecular organisation at every level of complexity from the evolutionary origins of the PSP to the proposed model of MASC assembly of in developing neurons. These synapse proteome datasets and the analysis described in this study offer a basis for future research in synaptic biology and will provide useful information in brain disease and mental disorder studies.

5.2 INTRODUCTION

An important goal of cognitive science is to identify neural mechanisms for processing information. The synapse both transmits information between cells and processes it by detecting specific patterns of neural activity and converting this electrical activity into intracellular biochemical events changing the properties of the neuron (Greengard, 2001);(Kandel, 2001) . In recent years it has become clear that synapses, like other cell-cell interactions in metazoans, utilise signal transduction complexes and pathways with a high degree of molecular complexity and cross talk (Pawson and Nash, 2003);(Sheng and Kim, 2002). A major challenge for biology is to devise strategies that extract emergent physiological properties and simple biological principles from highly complex genomic and proteomic datasets, and shed light on both mechanisms and disease processes. The synapse proteome is a suitable prototype to explore these general issues because; i) it contains a highly localised set of proteins found in dendritic spines; ii) an important role for signalling complexes and pathways has been established, iii) signalling can be exquisitely studied using patterns of action potentials, iv) genetic and pharmacological perturbations result in behavioural changes. However, the molecular complexity and global organisation is not well understood.

The post-synaptic density (PSD) is a structural component of the synapse that can be observed under the electron microscope, beneath the postsynaptic membrane (Cotman et al., 1974). Recent proteomic efforts have given an insight into the complexity of the PSD and its subcomponents (Husi and Grant, 2001), (Walikonis et al., 2000), (Yamauchi, 2002), (Yoshimura et al., 2004), (Jordan et al., 2004), (Peng et al., 2004), (Li et al., 2004), (Collins et al., 2005). Advances in mass spectrometry techniques have enabled large-scale investigations of the protein content of organelles with similar complexity to the PSD (for reviews see (Jeong et al., 2000); (Taylor et al., 2003)).

The neurotransmitter glutamate activates synaptic plasticity primarily via the ionotropic N-methyl-D-aspartate (NMDA) receptor and metabotropic (mGluR) receptors. This leads to Ca^{2+} elevation in the dendritic spine and signal transduction to AMPA (alpha-amino-3-hydroxy-5-methylisoxazole-4-propionate) receptors and other effector mechanisms. The cytoplasmic C-termini of NMDA receptor subunits (NR2A/ε1, NR2B/ε2) bind PDZ domains of PSD-95, a MAGUK (Membrane Associated Guanylate Kinase) protein. Previous proteomic analysis of NMDAR-PSD-95 complexes identified 77 proteins that include mGluR receptors, whilst AMPA receptors were found in different complexes (Husi et al., 2000). These complexes are embedded in the postsynaptic terminal of excitatory synapses with other postsynaptic proteins.

Here we present a large-scale proteomic analysis of the postsynaptic proteome (PSP) of the mouse brain. We have performed MS-based analysis of MASC and AMPA receptor complexes, isolated postsynaptic densities and validation of a set of MASC and PSD proteins by large-scale immunoblotting. In addition to generating novel data, we have systematically compiled data from publicly available sources to produce a comprehensive dataset of the composition of the PSP. This includes data on the PSD and various sub-components, namely the NMDA receptor complex (NRC), the AMPA receptor complex (AC), and the MAGUK associated signalling complexes MASC). Functional information for these PSP proteins was extracted from gene expression data derived from developing hippocampal primary neurons in culture. Integration of this data with that of the synapse proteome has resulted in a model of synapse assembly in which, the sequential expression of sets of genes describe how a synapse may be formed. In addition, inclusion of electrophysiological data from the same type of developing neuronal culture provides a functional measure of synaptic outputs, the timing of which is in agreement with that predicted by gene expression data. Extensive analysis of these synapse proteome datasets revealed multiple levels of organisation from gene expression, protein domain organisation, formation of multi-protein

complexes to their evolution. This proteomic map can be considered as a first draft of the synapse proteome and will be used to design experiments to dissect the molecular complexity at the synapse

5.3 MATERIALS AND METHODS

5.3.1 Biochemistry

PSD and complex purifications were performed by Holger Husi (University of Edinburgh)

MASCs were isolated using a peptide affinity method modelled on the PDZ binding domain of NR2 subunits and was reported elsewhere (Husi and Grant, 2001). AMPA receptor complexes (AC) were isolated using immunoprecipitation of AMPA subunits with an anti-GluR2 antibody (BD Biosciences). PSD fractions were prepared as described (Carlin et al., 1980). Protein samples were separated by SDS-PAGE and stained by Collodial Coomassie blue. The entire gel lane was excised into 42 individual protein bands, reduced, alkylated and digested with trypsin

5.3.2 Mass spectrometry

Mass spectrometry was performed by Lu Yu (The Wellcome Trust Sanger Institute).

The resultant peptide mixtures were analysed by on-line Liquid Chromatography Tandem Mass spectrometry (LC-MS/MS) to generate peptide sequence information (Link et al., 1999). Chromatographic separations of the peptide mixture were performed on a 180 μm (intensely stained bands) and 75μm (weakly stained bands) PepMap column using an Ultimate LC system (LC Packings, The Netherlands) delivering a gradient of formic acid (0.05%) and acetonitrile. The LC column was connected directly to a micro-electrospray interface on a Q-TOF mass spectrometer (Micromass Ltd., U.K.). The mass spectrometer operated in automated function switching mode, selecting precursor ions based on intensity for peptide sequencing by tandem MS. Several hundred MS/MS spectra could be generated per run

allowing the analysis of complex mixtures without any prior interpretation. The search engine used in this work was Mascot (http://www.matrixscience.com/). Mass spectrometry data were reduced to MassLynx 'PKL' format peak lists prior to searching. Proteins were identified based on matching the MS/MS data with mass values calculated for selected ion series of a peptide. The searches on in-house non-identical protein database were performed without applying any constraints on molecular weight or species of origin. Most proteins were identified with several peptide matches although a few proteins were assigned based on a single peptide provided near complete peptide sequence had been obtained.

5.3.3 Immunoblotting

Immunoblotting was performed by Holger Husi (University of Edinburgh) and by Julia Brandon (The Wellcome Trust Sanger Institute).

Protein samples were subjected to reducing SDS-PAGE and transferred to polyvinyldifluoride (PVDF) membrane (BioRad) at 4°C for 90 min at 75 V in 10% (v/v) Methanol, 10 mm CAPS pH 11.0. Dilution of primary antibodies was between 1 : 100 and 1 : 1000, depending on the quality of the IgGs. Detection of signals was done using peroxidase-linked secondary IgGs and enhanced chemiluminescence.

5.3.4 Primary Neuronal culture

Primary neuronal cultures were performed by Lawrence Humphreys and Paul Charlesworth (The Wellcome Trust Sanger Institute).

Primary hippocampal neurons were prepared from embryonic day 17.5 E17.5) C57 Bl6 (c-/c- coat colour mutant – required?) mice using standard techniques. Embryos were chilled and decapitated and hippocampi dissected in cold PBS. Hippocampi (from 4-8 embryos) were pooled and placed in papain solution (Worthington, 10 Units/ml, reconstituted in Neurobasal) and incubated for 22 min at 37ºC and 5% CO_2. Enzyme solution was then removed and replaced with 2mls DMEM/F12 + 10% FCS in which the tissue was disaggregated by drawing up and down a fire polished Pasteur pipette. Dissociated cells were centrifuged twice (3.5 min, 400g) in this

medium after which the pellet was re-suspended in Neurobasal / B27 medium with 0.5 mM glutamine. Cells were counted, diluted and plated in this medium. For RNA extraction cells were plated at 4-5 x 10^4 cells/cm^2 in 6 well plates coated with Poly-D-Lysine and laminin (1-2 μg/ cm^2). Multielectrode arrays were coated with polyethyleneimine and laminin. Initially, 10^5 cells in 150 μl were placed on the array. After 1 hr this medium was removed and 800 μl of fresh Neurobasal / B27 was placed on the MEA dish. Final cell density was approx. 1500/mm^2. MEA dishes were sealed with Teflon lids (see Potter and DeMarse 2001) to maintain osmolarity and sterility.

For total RNA extraction, typically 1 ml of TRIzol reagent (Invitrogen) per well was used, scraping the cells by pipetting and then following the manufacturer protocol. An additional clean up was made using the RNeasy Mini Kit (Qiagen) and subsequent ethanol and NaOAc precipitation, or alternatively using the MinElute kit (Qiagen). RNA integrity was checked in a 1% agarose-formaldehyde gel.

5.3.5 Gene expression profiling

Gene expression analysis was performed by Luis Valour (The Wellcome Trust Sanger Institute).

1.5 μg of total RNA were retrotranscribed, labeled and hybridised on MG-430 2.0 array (Affymetrix), one array per sample. We followed the manufacturer instructions, including the washing (Fluidics Stations 450, Affymetrix), scanning (GeneChip Scanner, Affymetrix) and image treatment (GCOS software, Affymetrix). Once verified quality and control report after each experiment, GeneSpring software (Silicon Genetics) was used to normalise and analyse the data, including the clustering method based on the K-means algorithms.

5.3.6 Analysis of PSP protein classes

Triplicate data for PSP probe sets were averaged and normalised ((X-m)/δ)). Hierarchical clustering of datasets was performed using average group linkage with the Pearson correlation coefficient distance measure. Normalisation

and clustering were performed using HCE3 (http://www.cs.umd.edu/hcil/multi-cluster/). A minimum similarity threshold of 0.8 was used to filter the normalised data and the resultant clustering was used to remove probe set outliers and facilitated averaging of data from probe sets to unique genes. Time course series for protein classes were plotted using normalised expression values >0 to indicate up-regulation of expression at a given time point. PSP protein classes were defined by Swiss Prot keywords and can be found in supplementary table 2. The functional class of genes termed as "Neurogenesis" was also defined by Swiss Prot keywords and the class "activity dependent genes" was defined by literature searching (supp table act dep genes). PSP targets of the transcription factor creb were extracted from a recent screen of PC12 cells (Impey et al., 2004). Neural specific transcription factors present in the PSP were extracted from a recent genome wide screen (Gray et al., 2004).

5.3.7 Tissue specific expression

The Gene expression atlas (http://expression.gnf.org) was mined to extract expression data for genes in the PSP. 635 probe sets corresponding to 464 unique PSP genes were obtained and expression data were normalised and clustered as described previously for the culture time course experiments. Significantly enriched ($P < 0.01$) Gene Ontology terms were determined for genes in each cluster using GoTree (Zhang et al., 2004).

5.3.8 Electrophysiology with MEAs

Elextrophysiology was performed by Lawrence Humphreys and Paul Charlesworth (The Wellcome Trust Sanger Institute).

Multielectrode arrays (MEAs) (Multi Channel Systems (MCS), Reutlingen) were used to record the electrical activity of developing E17.5 hippocampal neurons, grown directly on a square grid of 60 titanium nitride electrodes (electrode diam 30 μm, spacing 200 μm). Signals were amplified (1100x), sampled at 25 KHz and recorded using MC Rack software (MCS). Spikes were detected on-line using a fixed

level detection of -24 μV, which was approx 8 x SD of baseline noise. Off line analysis of spike data was performed with Neuroexplorer (NEX Technologies).

Prior to recording, culture medium was removed and retained and recording solution (NaCl 140, KCl 3, $CaCl_2$ 2.5, HEPES 15, glucose 10, pH 7.4 with NaOH, Mg free) placed in the array well (1 rinse). 30 min recordings were made in this solution. The original culture medium was then replaced in the well for further culture in the TC incubator.

5.3.9 Integration of proteomic datasets

A number of different sequence databases were used in these studies for MS mining and as a result, we decided to use UniProt as our standard protein identifier. In a few cases protein identifiers could not be mapped to UniProt, so their original identifiers were retained. UniProt accession numbers were mapped to Unigene via EMBL and Locuslink to cluster the merged total PSD dataset to remove redundancy.

5.3.10 Network construction and gene expression analysis

Cytoscape version 2.1 (www.cytoscape.org) was used as a platform to construct networks onto which gene expression data was visualised. Cytoscape allows construction of network graphs, with genes, proteins, and molecules represented as nodes and interactions represented as links, i.e. edges, between nodes. Protein interaction data from PPID (www.PPID.org) was loaded into Cytoscape as a SIF file, which specifies nodes and interactions only. Once loaded, normalised time-course expression levels from the primary neuronal gene expression experiments were loaded as PVAL files and were visually superimposed on the network.

5.3.11 Ortholog mapping and analysis

PSP gene identifiers were mapped to Ensembl predicted genes for Human NCBI Build 34 using the EnsMart data retrieval system. Orthologs of the Human genes

were identified in 8 eukaryote species using the Ensembl Compara database. Species studied and corresponding genome builds were; *Mus musculus* (Mouse, NCBI m32), *Danio rerio* (Zebrafish, WTSI Zv3), *Rattus norvegicus* (Rat, RGSC 3.1), *Gallus gallus* (Chicken, WASHUC1), *Anopheles gambiae* (Mosquito, MOZ 2), *Drosophila melanogaster* (Fruitfly, BGDP 3.1) and *Caenorhabditis elegans* (WS 116). In addition, homologues of the proteins were identified in the bakers yeast *Saccharomyces cerevisiae* using the Inparanoid database (http://inparanoid.cgb.ki.se/). An ortholog grid was constructed where "1" corresponded to an ortholog present and "0" to an absent ortholog. This data was used for hierarchical clustering to identify evolutionary clusters. The distribution of PSP protein classes populated by orthologs in each species was analysed and numbers of the top enriched protein classes in each evolutionary cluster plotted as bar charts.

5.4 RESULTS

5.4.1 Composition of the Postsynaptic Proteome

We isolated MAGUK associated signalling complexes (MASCs) using a previously described peptide affinity method (Husi and Grant, 2001). This involved using a peptide corresponding to the six C-terminal residues of NR2B that binds MAGUK proteins (including PSD-95), as a specific affinity ligand to isolate NMDA-MAGUK containing complexes. Purified complexes were separated by SDS-PAGE gel electrophoresis, bands excised, digested, and analysed by mass spectrometry. Western blotting of peptide-purified complexes for candidate proteins were also performed. These complexes are similar in composition to NMDA receptor complexes and MAGUK complexes isolated with antibodies (Husi and Grant, 2001),(Husi et al., 2000). Out of the previously identified 100 proteins identified in the NRC by analysis of anti-NR1 immuno- purifications, 84 of these were also found by the peptide approach. In addition, 86 new proteins were identified by the peptide approach, many of which were in the molecular weight range obscured by the contaminating antibody chains in the immuno-purification strategy.

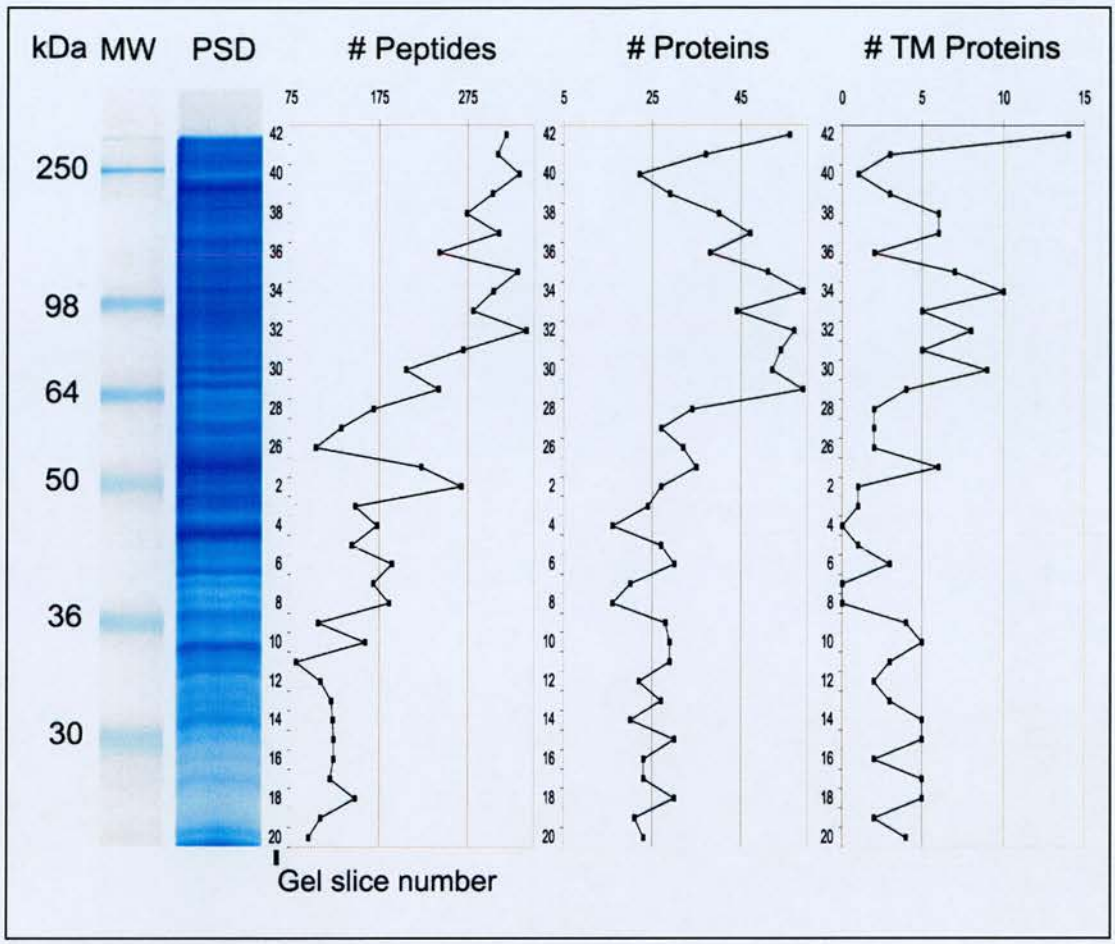


Figure 5.1 - Distribution of peptide and protein identifications of PSD proteins by MS analysis of gel-slices.
Coomassie-stained SDS-PAGE gel lane containing purified mouse forebrain postsynaptic density proteins. The resulting numbers of peptides, proteins and transmembrane proteins identified in each gel-slice by LC-MS/MS analysis are shown. 1D-SDS-PAGE separation offers effective protein separation of complex mixtures with an average of 34 proteins identified per gel-slice and an average of 12 peptides characterised for each protein. The number of peptides detected per gel slice increases proportionally with molecular weight. 1D-SDS-PAGE is very compatible with membrane samples and the distribution of proteins with transmembrane (TM) regions is shown.

We combined datasets of proteins found in both the NRC and the MAGUK-pep under the term MAGUK associated signalling complexes (MASCs) due to the large overlap. 95 MASC proteins were detected by western blotting of purified complexes and 35 of these were also validated by MS data.

The MASC set represents a comprehensive account of NMDAR-MAGUK complexes identified by large-scale western blotting and mass spectrometry based analyses of both antibody and peptide affinity-based purifications. We performed immunoprecipitation of the AMPA receptor complex with an antibody directed to the GluR2 subunit and found a significantly smaller complex of 9 proteins by MS analysis, which included all 4 AMPA subunits. In order to gain an overall view of the postsynaptic proteome and to validate the components of the MASC and AMPA complexes, we performed a systematic analysis of purified postsynaptic densities. PSD proteins were separated on an SDS-PAGE gel (Figure 5.1), 37 gel slices were analysed by LC-MS/MS. 7402 peptides were identified corresponding to 620 non-non-redundant protein identifications, with an average of 16.7 proteins identified per gel slice (Supplementary Table 1).

5.4.2 Integration of postsynaptic proteome datasets

To evaluate the coverage of our PSD data set (now referred to as G2C PSD) we compiled data from seven published proteomic studies of the PSD and 119 individual papers reporting PSD localised proteins. A number of different sequence databases were used in these studies for MS mining and as a result, we decided to use UniProt as our standard protein identifier. In a few cases protein identifiers could not be mapped to UniProt, so their original identifiers were retained. We observed redundancy in datasets, with the same protein ID present numerous times, as well as many IDs corresponding to the same gene product. All 1124 proteins were submitted for blast analysis to the Proteome Analyst (Szafron et al., 2004) to obtain homologous sequences for human, rat and mouse. We then mapped

UniProt accession numbers to Unigene via EMBL and Locuslink to cluster the merged total PSD dataset to remove redundancy. This process enabled grouping of numerous entries into Unigene clusters and resulted in reduction of the dataset to 1124 non-redundant entries (Supplementary Tables 2 and 4). An evaluation the relative coverage of each component PSD dataset is presented in Table 4.1.

	G2C %	Yosh %	Jordan	Peng %	Li %	Literature	Satoh %	Walik %	Method
G2C %	100								1D Gel + LC-MS/MS
Yosh %	45	100							2D-LC-MS/MS
Jordan	54	43	100						1D Gel + LC-MS/MS
Peng %	68	54	53	100					1D Gel + LC-MS/MS
Li %	57	54	50	45	100				2D Gel/ICAT + MALDI-TOF
Literature	66	51	54	50	31	100			Various
Satoh %	85	57	52	57	41	37	100		2D Gel + LC-MS/MS
Walik %	100	97	93	90	90	72	41	100	1D Gel + MALDI-TOF

Table 5.1 - Percentage overlap between PSD datasets. Each dataset is represented as a percent coverage of each other dataset in this grid. The G2C dataset has the best coverage of each other PSD dataset reported indicating that this set is the most complete. Also, it is clear that the G2C PSD dataset is the most complete with 66% coverage of literature defined PSD proteins.

This matrix of percentage coverage shows that the G2C PSD dataset has the highest coverage of each dataset in the group, indicating that this set of proteins is the most representative and complete single dataset.

The analytical approaches used in these 7 studies are varied from using 1D and 2D electrophoresis for protein separation, ICAT for sample simplification to a number of peptide chromatographic methods (RV, 2D-LC) and MS methods (MALDI-ToF, LC-MS/MS) (Table 4.1). We addressed the issue of bias arising from the use of different methodologies by comparing the distribution of proteins identified in; our dataset (1D SDS-PAGE and LC-MS/MS); the dataset of Yoshimura et al (Yoshimura

et al., 2004) (2D-LC-MS/MS) and the dataset of Li et al (Li et al., 2004). (2D gel/ICAT and MALDI-TOF, LC-MS/MS). A 2D scatter plot of Log Mr versus pI was plotted (Figure 5.2) for these 3 datasets. It can be seen that the 2D-LC-MS/MS approach used by Yoshimura et al. has the benefit of identifying many low molecular weight proteins as this approach does not utilise any gel electrophoresis step in which these proteins would be lost. Our dataset is biased towards integral membrane proteins, with this class representing 24 % of our total dataset. Both 2D-LC and 2D gel electrophoresis are poorer approaches for identifying membrane proteins with a representation of 17 and 7 % transmembrane proteins respectively.

The distribution of protein identifications in the Total PSD dataset show that the majority of proteins were detected only once (58%) (Figure 5.2A) and that 198 (18%) proteins were detected twice. The number of proteins detected 5 times or more is just 72 (6%), although this value is limited by the smaller datasets included in the Total PSD dataset. In order to define a set of higher confidence proteins we grouped proteins, which were identified two or more times (466 proteins) and termed it the "Consensus PSD" (cPSD). We reasoned that the majority of this set of proteins are abundant, which are more likely to be detected in multiple analyses and that the set of proteins detected once are perhaps of lower abundance or may reflect differences in sample preparation. The coverage of each individual dataset with respect to the core and Total PSD datasets is illustrated in Figure 5.2B. 363 proteins in our dataset have been validated by their presence in one or more other datasets, representing 78% coverage of the cPSD. In addition, our dataset contains 257 unique proteins representing 39% of proteins detected once.

5.4.3 Verification of Postsynaptic density proteins

We sought to verify the sub-cellular location of a selected set of postsynaptic density and MASC proteins by large-scale immunoblotting of whole extract, synaptosomes and postsynaptic densities from mouse forebrains. 48 commercially available

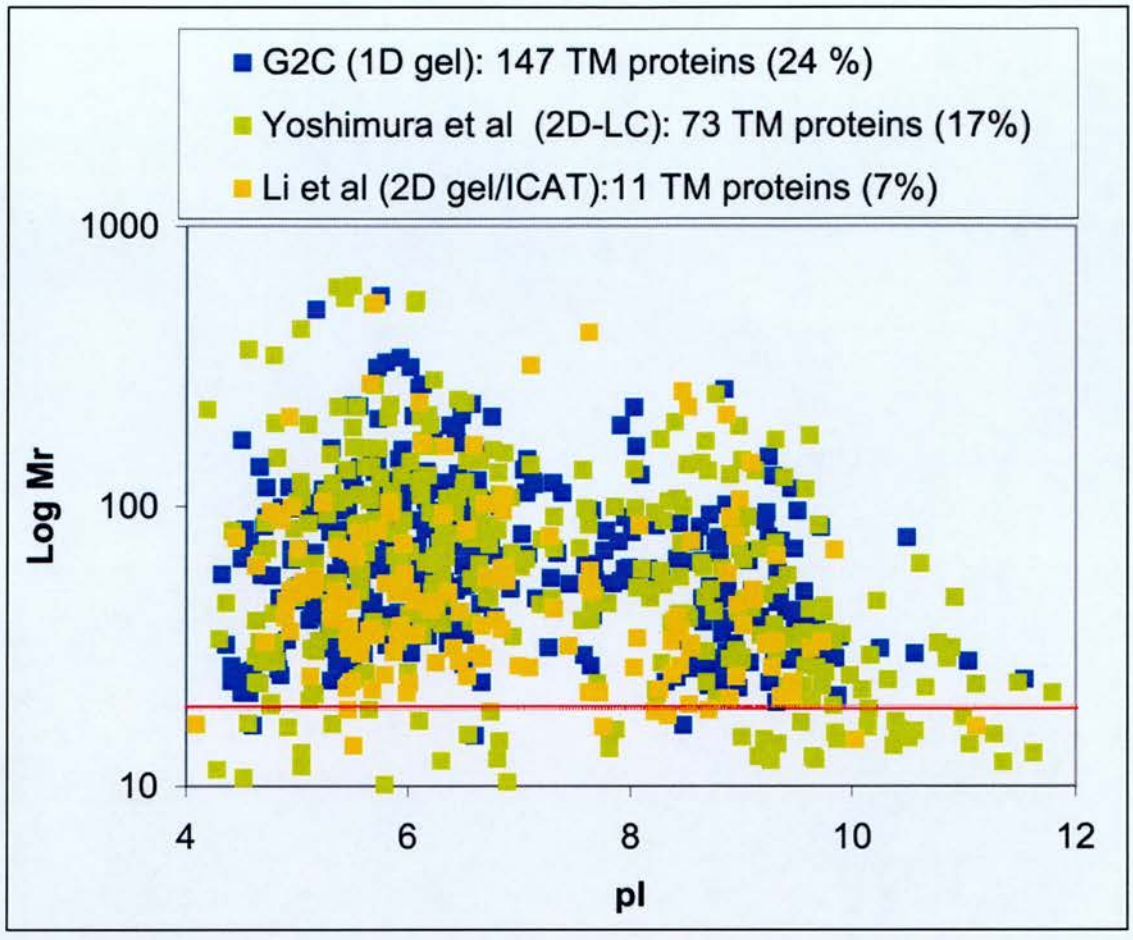


Figure 5.2 - Comparison of analytical approaches by physiochemical properties.
The G2C approach which was based on 1D-SDS-PAGE separation of PSD proteins followed by systematic LC-MS/MS analysis of each gel slice is very well suited to identification of integral membrane proteins and resulted in the identification of 147 transmembrane proteins, 25% of our PSD dataset. Our approach identified less, low molecular proteins than a gel-free MudPIT-type approach (Yoshimura et al.) as these proteins are out of the molecular size range of the 1D gel used. Li et al., used a combination of 2D gel separation as well as ICAT purification which resulted in poorer coverage of transmembrane proteins (7%) and pI restricted protein identifications.

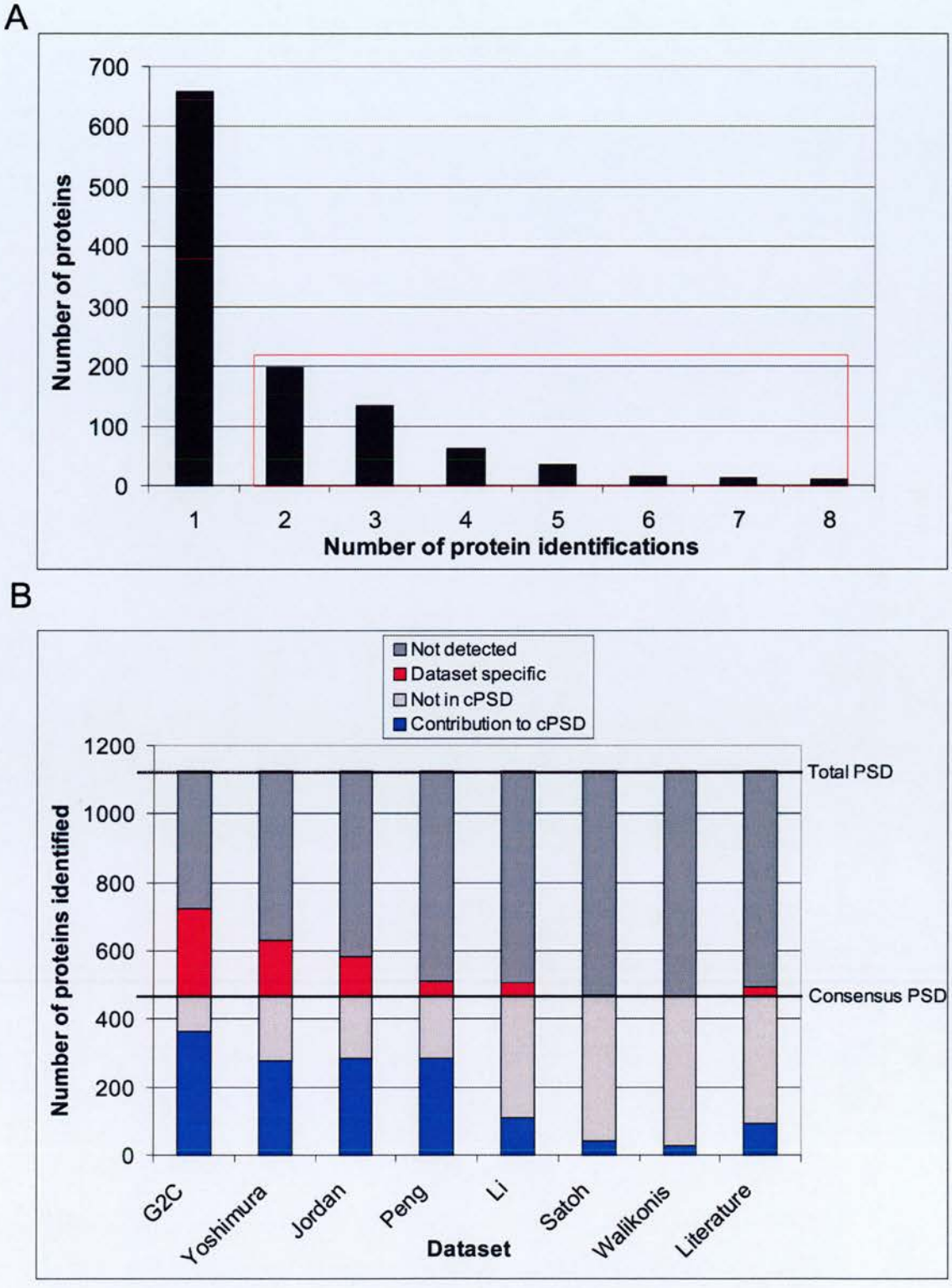


Figure 5.3 - Definition of a "Consensus PSD" set of proteins by integration of multiple synaptic datasets.

(A) Analysis of the distribution of protein identifications reveals that the majority of PSD proteins were identified only once. Proteins identified in two or more studies may represent a higher confidence and most likely, a high abundance cPSD dataset. (B) Contributions of each PSD dataset to the cPSD (Red) and Total PSD (Blue) datasets. It can be seen that the G2C PSD analysis has the highest coverage of the cPSD at 78 % and also contributes the most unique proteins to the Total PSD dataset.

antibodies were tested on these three fractions to demonstrate the presence and enrichment of postsynaptic proteins (Supplementary Table 5). Overall, we found that 21 proteins (44%) were clearly enriched, 8 (17%) showed similar abundance, 14 (29%) were present but not enriched and 5 (10%) were undetectable in isolated postsynaptic densities compared to the other brain fractions (Figures 5.4 and 5.5). 33 of the 48 proteins tested are cPSD proteins and 15 of these displayed enrichment in the PSD fraction. Proteins contained in the MASC dataset were identified in purifications from whole brain extract and although 144 out of 186 have been verified by their identification in the PSD fraction, the remainder of this dataset requires validation. 15 antibodies to MASC proteins were used to confirm their presence in the PSD, of these 8 have not been shown to be PSD proteins and 7 were detected in the PSD by MS just once. We have confirmed the PSD localisation of 14 out of the 15 MASC proteins tested (Figures 5.4 and 5.5), indicating that the majority of MASC proteins are likely *bona fide* postsynaptic density proteins.

In addition to verification by immunoblotting, we performed extensive searching in PubMed for PSD localisation data for these 1124 proteins (Supplementary Table 2). We found that 119 (10.6%) of these proteins have been previously reported to have postsynaptic density localisation. As shown in Supplementary Table 3, we have identified the largest number of known PSD proteins (66% of PSD proteins from the literature) in our analysis of the PSD compared to other similar analyses, indicating the high coverage of our dataset.

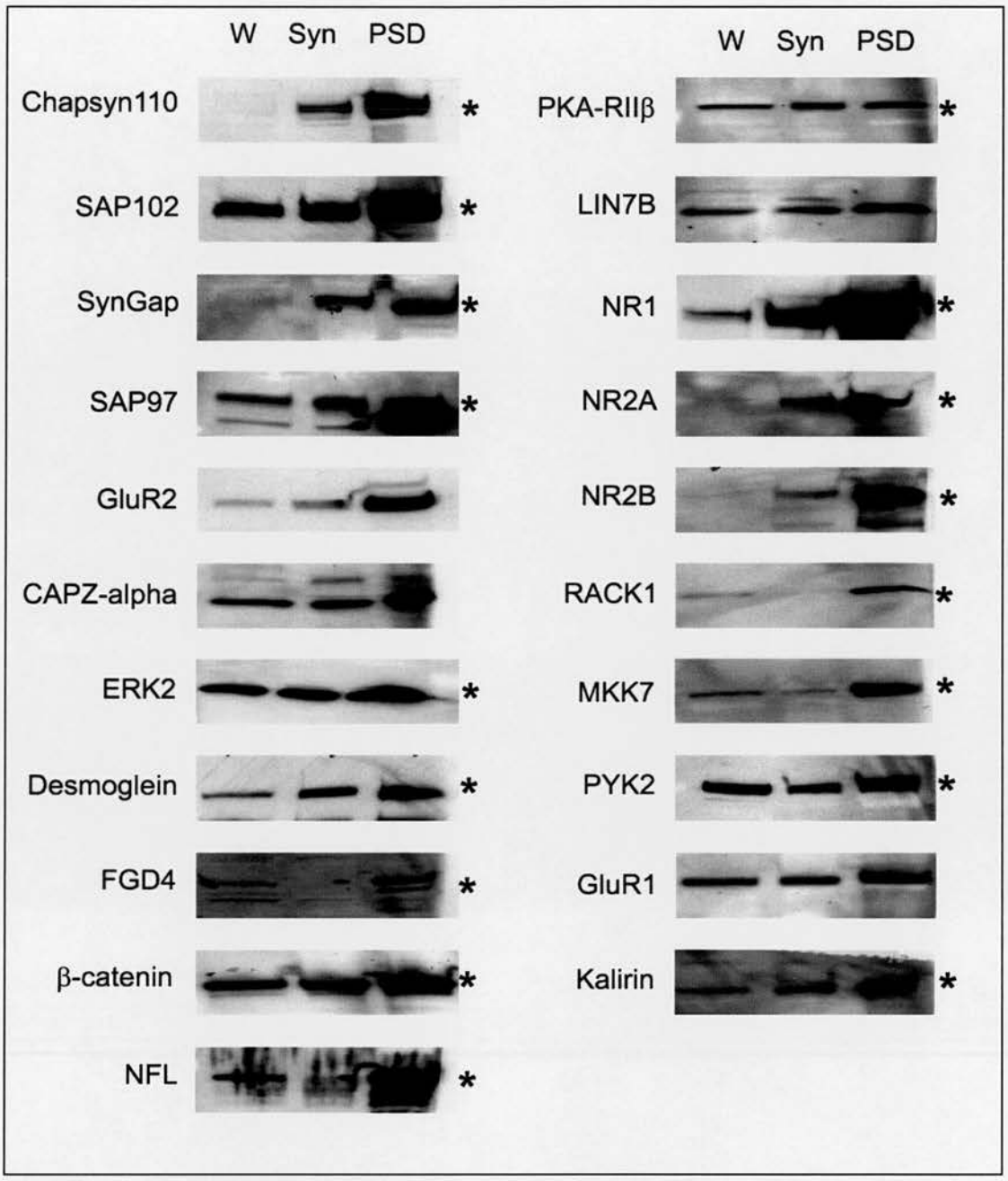


Figure 5.4 - Immunoblotting of proteins enriched in the PSD fraction
Antibodies directed to PSD/MASC proteins were used in immunoblotting experiments to validate their presence in the postsynaptic density fraction. 21 of these proteins shown here exhibited enrichment in the PSD fraction, compared to whole extract and synaptosomal fractions. 3 MASC components (FGD4, MKK7 and RACK1) have not been found in the PSD by MS analyses and 3 (Desmoglein, ERK2 and PYK2) were detected only once by MS analyses of PSD fractions. All six of these MASC components have been validated by immunoblotting which revealed their presence and enrichment in the PSD. MASC proteins are indicated by an asterix (*).

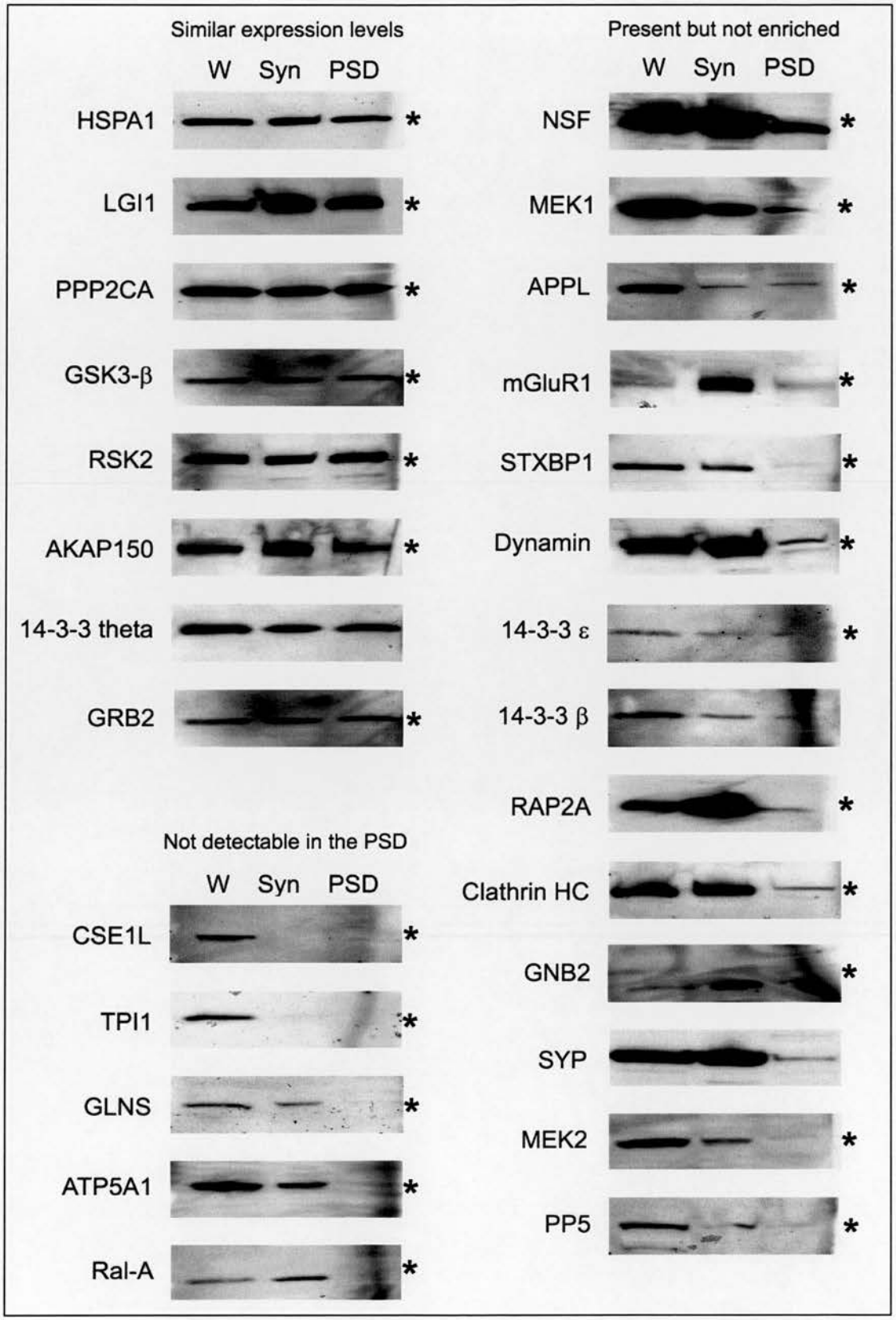


Figure 5.5 - Immunoblotting of proteins present or absent in the PSD fraction
Antibodies directed to PSD/MASC proteins were used in immunoblotting experiments to validate their presence in the postsynaptic density fraction. 8 proteins showed similar expression levels in whole extract, synaptosome fractions and PSD fractions. 14 proteins were present in the PSD but not enriched and 5 proteins were not detectable in the PSD. MASC proteins are indicated by an asterix (*).

5.4.4 Functional classification of PSP components

The initial approach to characterise the organisation of the postsynaptic proteome was to classify the proteins into categories that describe their known functional properties (Table 1). Within the overall postsynaptic proteome, representatives of many classes of proteins representing a broad range of cell biological functions were observed: membrane bound receptors, adhesion proteins and channels; a plethora of signalling proteins and adaptors; and proteins involved with regulation of transport, RNA metabolism, translation and transcription. The cPSD contains similar numbers of proteins in most classes compared to our PSD dataset, with the exception of channels and receptors, signalling proteins and synaptic vesicle proteins.

	Total PSD	G2C PSD	Consensus PSD
Channels and Receptors	80	60	39
MAGUKs / Adaptors / Scaffolders	54	31	33
Ser/Thr Kinases	46	23	20
Tyr Kinase	3	3	2
Protein Phosphatases	18	13	11
G-proteins and Modulators	77	40	34
Signalling molecules and Enzymes	278	167	98
Transcription and Translation	119	36	41
Cytoskeletal and Cell Adhesion Molecules	153	99	97
Synaptic Vesicles and Protein Transport	159	119	76
Novel	107	29	11
Other	30	0	4
Summary	1124	620	466

Table 5.2 - Summary of PSD protein classifications.
Proteins found in the Total PSD dataset, our PSD dataset and the cPSD dataset were functionally classified to show protein class enrichment. The distribution of our PSD dataset and cPSD protein classes are similar except for Signalling molecules and Enzymes, which are particularly enriched in our PSD dataset. In addition, the main contribution of low abundance proteins such as tyrosine kinases and phosphatases to the Total PSD dataset was from our PSD dataset. In fact, our study was the only one to identify tyrosine kinases in the PSD, further supporting the increased sensitivity of our approach. Details for individual proteins and further sub-classifications are found in Supplementary Table 2.

This becomes evident when the cPSD is compared to the Total PSD, indicating that the reduced presence of these protein classes in the cPSD is most probably due to the lower abundance of these proteins, and thus the decreased likelihood of detection. This supports the idea that the cPSD represents a consensus of protein identifications which, on one hand, decreases individual dataset bias for contaminating proteins (nuclear or mitochondrial) but on the other hand, has limited coverage of some protein classes because of abundance issues. It is worth noting that the G2C PSD dataset contributed all three tyrosine kinases found in the PSD, an indication of the sensitivity of our approach as tyrosine kinases are of relatively low abundance. The G2C PSD dataset also has greater coverage of most protein classes compared to individual datasets, which comprise the Total PSD.

5.4.5 Analysis of protein domains in PSP proteins

Do synapse proteomes comprise particular functional subsets of proteins? We examined their protein domain profiles using the InterPro resource (http://www.ebi.ac.uk/interpro). All domains in three PSD datasets (G2C PSD (594 proteins with domains (1709 domains)), the Total PSD (1032 proteins with domains (2786 domains), and cPSD (447 proteins with domains (1298 domains)) were identified and ranked by frequency (percentage of proteins in each PSD dataset with a specific domain) and compared with the UniProt mouse proteome (http://www.ebi.ac.uk/integr8). The enrichment of the top 10 domains found in the cPSD was compared to the Total and our PSD dataset (Figure 5.6). The greatest enrichment in these datasets is for protein interaction domains and in particular, the cPSD has a striking enrichment of PDZ and SH3 domains (7.7 and 7.2 fold respectively). This is in agreement with the notion that the cPSD contains abundant proteins such as synaptic scaffolders containing PDZ and SH3 domains. The most common domains in all three datasets are associated with kinase function and the enrichment profile for these domains is similar across these datasets.

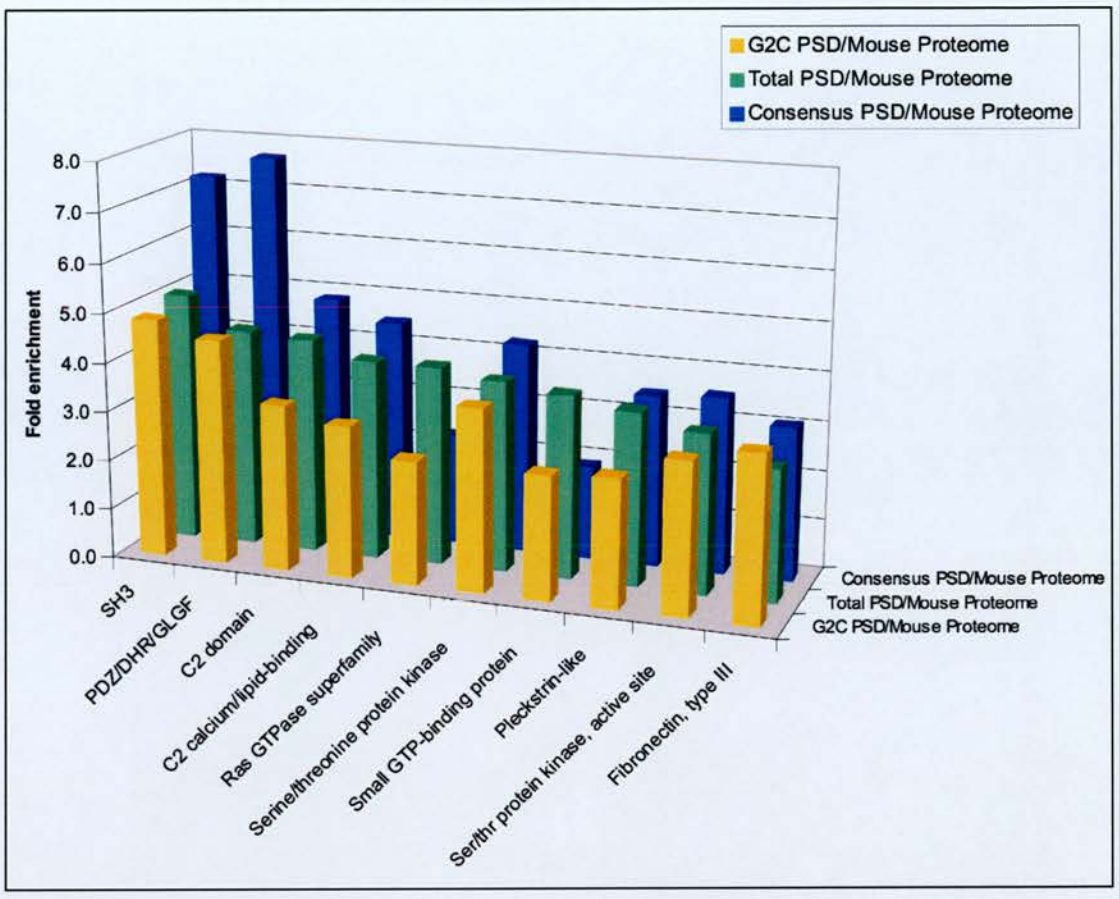


Figure 5.6 - Enrichment of protein domains in PSD datasets. Interpro protein domains were extracted for all PSD datasets and expressed as a ratio of their frequency in PSD datasets compared to the mouse proteome. The top 10 enriched domains in the Consensus PSD are plotted, along with the Total and G2C PSD datasets for comparison. The most highly enriched domains are associated with scaffolding proteins, namely PDZ and SH3 domains which characterise MAGUK proteins, important postsynaptic density scaffolders. Other enriched domains are associated with prominent postsynaptic processes such as calcium signalling, phosphorylation and G-protein signalling.

This enrichment in kinase domains correlates with enrichment in other domains commonly found in protein kinases, which allow them to interact in networks (e.g. SH3, PDZ interaction domains (Manning et al., 2002). Importantly, domains involved with Ca^{2+}-dependent signalling, a major feature of postsynaptic signal transduction, were also highly abundant (C2, C2 calcium/lipid-binding). Similarly, members of the Ras GTPase superfamily were enriched by over 5-fold in the Total PSD dataset. Conversely, some domains that are highly abundant in the mouse proteome were absent from all PSD datasets, such as olfactory receptor homeobox domains (2^{nd} and 19^{th} most abundant domains respectively).

5.4.6 Protein domain organisation

To evaluate the differences in molecular organisation between the MASC (a protein complex dataset) and the cPSD, we compared their domain organisation by hierarchical protein domain clustering. Proteins in these datasets that contained Interpro domains that were present more than once, and were present in more than one protein were clustered and the resultant clusters were annotated according to their protein classes (Supp table). MASC protein domains cluster into 14 discrete sets of proteins (Figure 5.7) reflecting many of the structural and functional aspects known about this set of proteins. This complex is centred on interactions with glutamate receptors and PDZ-containing scaffolders, which provide a framework in which a host of kinases, phosphatases and signalling molecules transduce information to gene expression regulatory pathways and to cytoskeletal and transport networks. A large proportion of the MASC (113 proteins) is represented in the cPSD, so to compare these two datasets; we clustered domains from cPSD proteins, which are not MASC proteins (Figure 5.8). This cPSD (minus MASC) set forms 25 clusters that encompass all MASC clusters except for tyrosine kinases, which are well-known low abundance proteins. There is an increase in the diversity

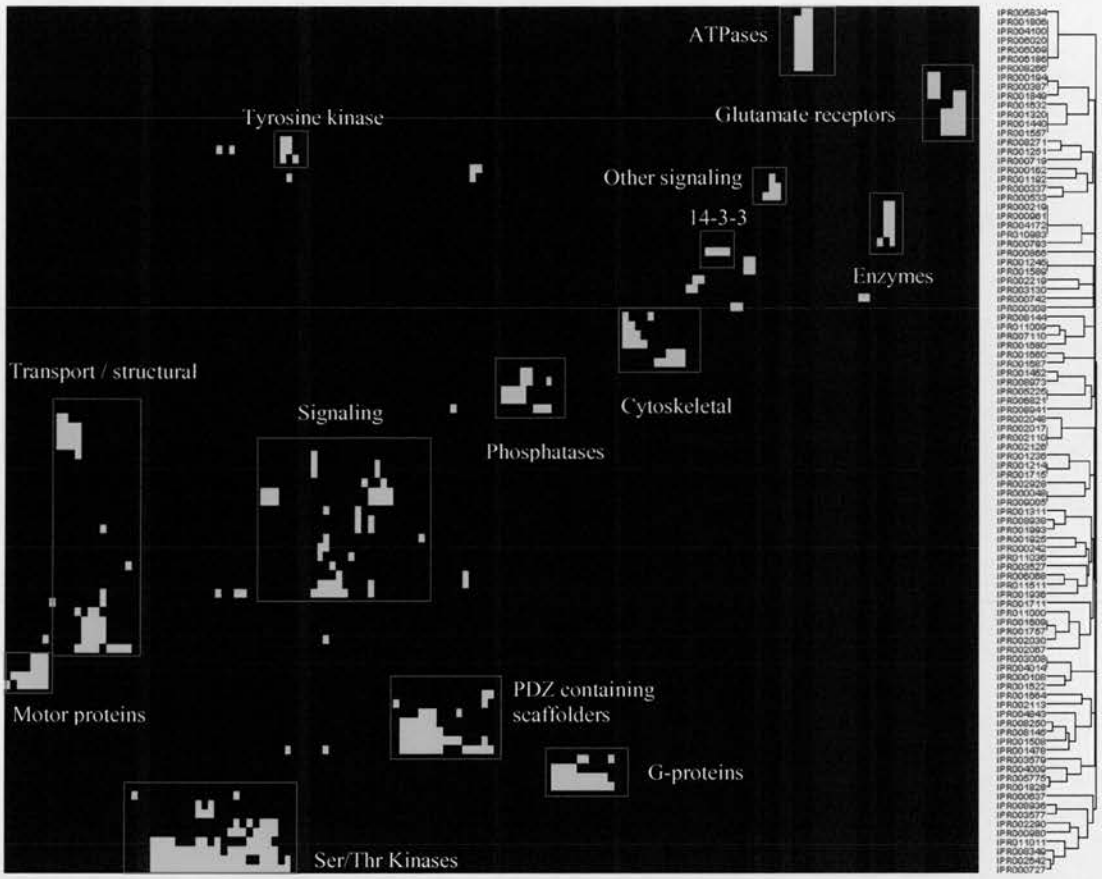


Figure 5.7 - Hierarchical domain clustering of MASC proteins. Interpro domains in MASC proteins occurring more than once in the MASC dataset were clustered. Discrete clusters (red boxes) were annotated with the protein class of which the majority proteins belonged to. Clusters are formed from co-occurrence of 2 or more domains characteristic of a protein class

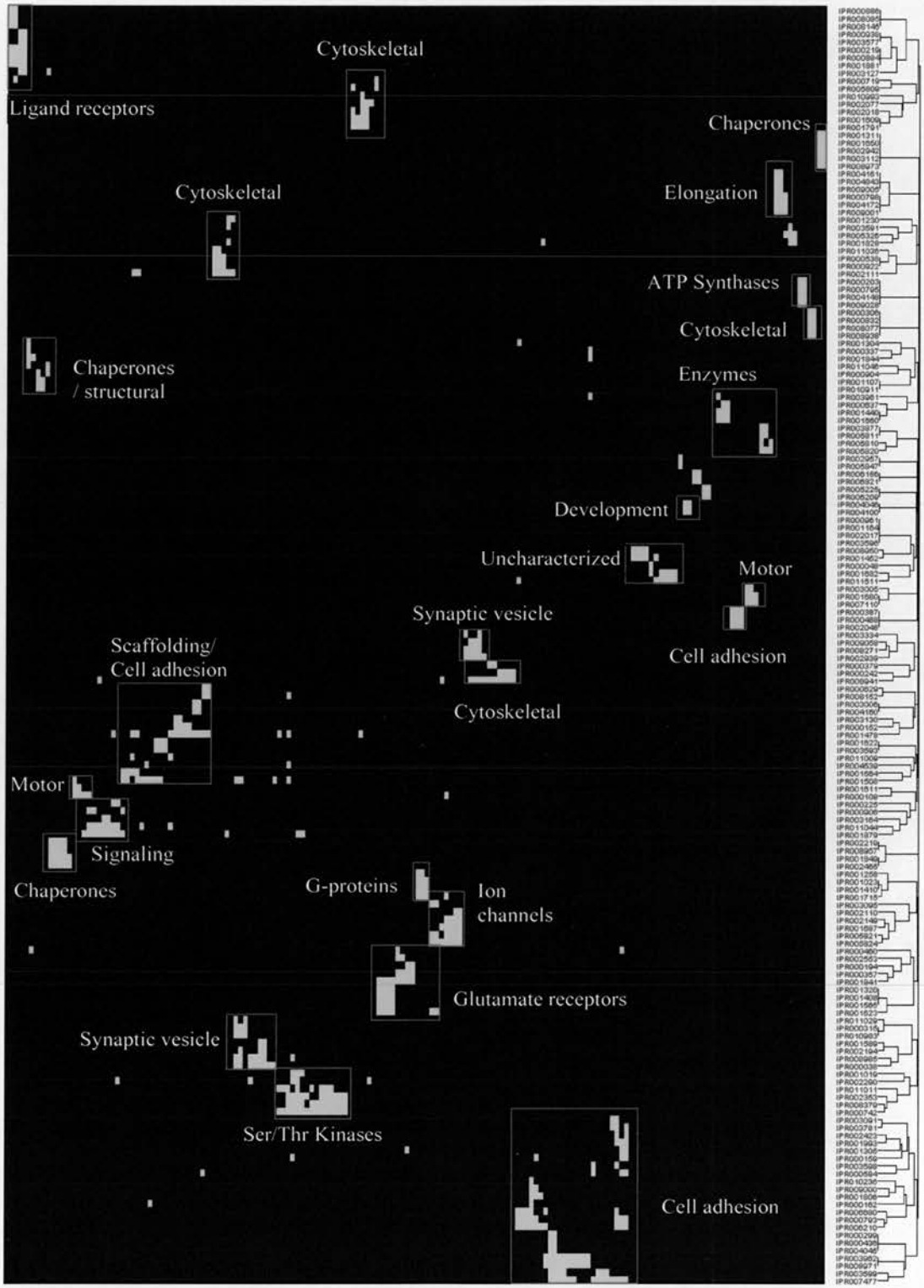


Figure 5.8 - Hierarchical domain clustering of cPSD minus MASC proteins. In order to dissect the molecular differences between the cPSD and MASC, Interpro domains in cPSD minus MASC proteins, occurring more than once in the MASC dataset were clustered. Discrete clusters (red boxes) were annotated with the protein class of which the majority proteins belonged to. Clusters are formed from co-occurrence of 2 or more domains characteristic of a protein class

of protein domains/protein classes represented in this subset of the cPSD and also a clear enrichment of cytoskeletal/cell adhesion proteins representing 7 clusters. This dataset also contains other receptors and channels and may represent abundant parts of other multiprotein complexes yet to be characterised.

5.4.7 Multiprotein complexes in the postsynaptic density

Analysis of organelles and protein sub-fractions based on affinity purifications presented in this study provides complimentary information. It provides preliminary validation of proteins, sub-cellular localisation information on components identified in protein complexes and vice versa. Organelle-like structures such as the PSD, represent complex protein mixtures that contain molecules across a large dynamic range of abundances and fractionation methods are known to enhance protein identification. In most cases, this would have been based on general physical properties of proteins such as molecular weight or charge (Lubec et al., 2003), however we have used methods to isolate protein complexes, the functional models within the organelles. Characterisation of protein complexes from the PSD, such as MASC and AMPA has enabled identification of 78 proteins that were not characterised in our G2C PSD analysis and increased our coverage of the PSP by 12.5%.

Previous proteomic studies showed that NMDAR-PSD95 complexes were 2-3MDa particles which included mGluR receptors, and that AMPA receptors were in different complexes (Husi and Grant, 2001); (Garry et al., 2003). The mGluR5 receptor complex has recently been characterised in a similar fashion to the MASC and was found to comprise of 76 proteins, including the NMDA receptor subunit NR2A (Farr et al., 2004). The majority of proteins in all three glutamate receptor complexes have been independently identified as PSD proteins (Figure 5.9). These complexes also contain proteins, which have not yet been detected in PSD analyses,

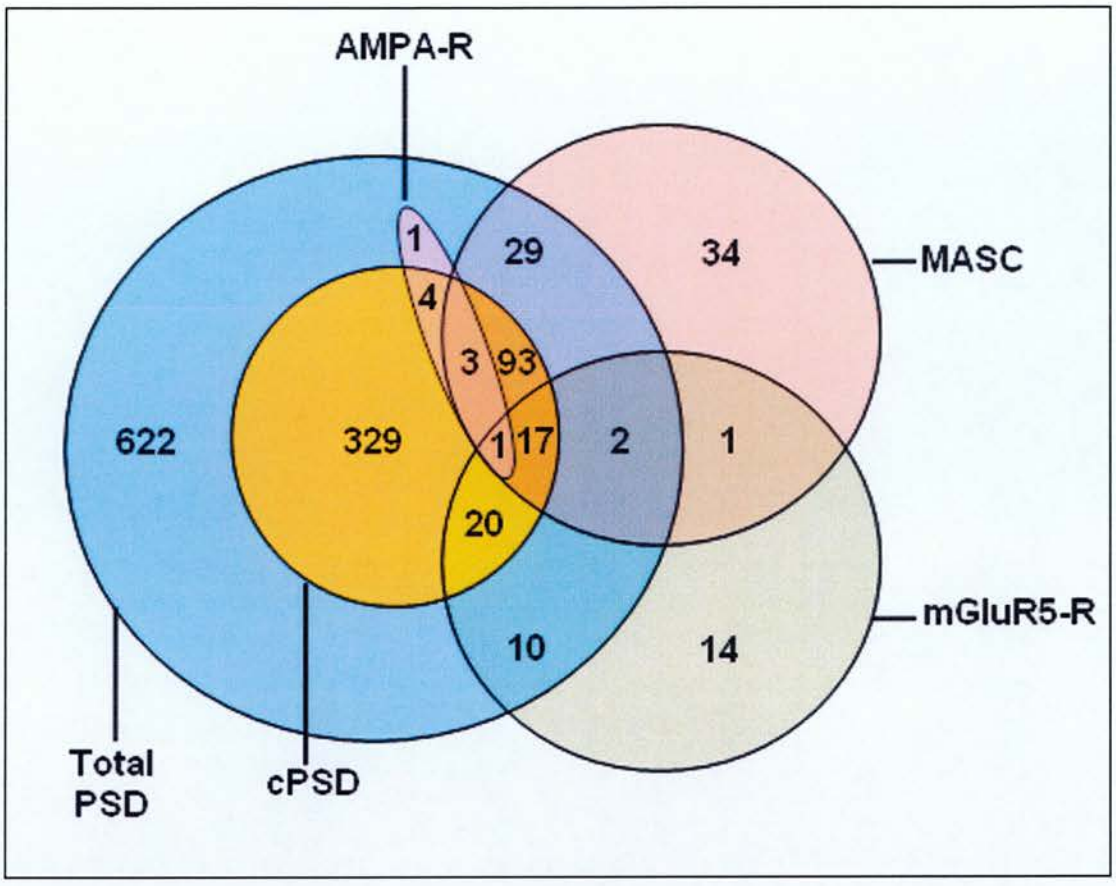


Figure 5.9 - Multiprotein complexes at the postsynaptic density.
Venn diagram illustrating the overlap of three glutamate receptor complexes with the Total and cPSD datasets. It can be seen that the majority of overlap between components of these complexes and the PSD occurs in the high confidence cPSD dataset. Proteins detected in these multiprotein complexes, which were not found in any of the PSD datasets, are generally of low abundance that are enriched in immuno-purifications of complexes compared to whole PSD analyses.

presumably due to abundance issues in analyses of PSD preparations compared to enriched protein complexes. We have validated the presence of seven such proteins in the PSD by immunoblotting and this has been taken into account in the numbers shown in Figure 5.4. It is interesting to note that the majority of the components of these complexes validated as PSD proteins and the overlap between these complexes occur in the cPSD (Figure 5.9), supporting the notion that the cPSD is an important subset of the Total PSD dataset. This is consistent with a postsynaptic organisation where the MASC signalling complex is connected to multiple cell biological effector mechanisms organised into their respective complexes. In addition to MASC and AMPA complexes, components of other complexes such as cell adhesion, growth factor, cytoskeletal, transport and ribosomal complexes were detected.

5.4.8 Developmental assembly of PSP components

5.4.8.1 PSP expression profiles

Having established a map of the protein components of the PSP in adult brain, we sought to investigate how these components are assembled in a temporal fashion using a genome-wide approach provided by the Affymetrix microarray technology. Triplicate mouse primary hippocampal cultures were grown and total RNA extracted at 0.25, 1, 2, 4, 8, 12 and 16 days, corresponding to different stages in the neural morphology of the cells (Figure 5.10A). We did not extend further to avoid possible apoptotic processes that could modify the gene expression profile. Briefly, the morphology of the culture could be described as follows: first, some incipient neurites were observable at 0.25 day but the initial neuronal contacts were taking place around 4 days. Since here, a profusion of branches and prolongations appeared more and more dense, producing mesh and fibre-like structures at the final points of the time course (12 and 16 days).

PSP protein accession numbers (UniProt, Unigene and Entrez Gene) and gene names were used to mine probe set ids on the Mouse 430_2.0 affymetrix array, which was used for gene expression analysis of primary neuronal cultures. An initial hierarchical clustering analysis with a minimum similarity of 0.8 was used to remove outliers and to average data from multiple probe sets to provide single data points corresponding to individual genes. This data was then re-clustered as shown in Figure 5.10B. This analysis resulted in 26 gene expression clusters, of which there were four main clusters with similar expression profiles. Gene ontology (GO) analysis of each cluster was performed and significantly ($P > 0.01$) enriched GO terms (compared to distribution of GO terms for all of the genes on the array) are shown in Figure 5.10B. Cluster 1 represents a steady up-regulation of gene expression with respect to time and corresponds to genes involved in signalling complex formation and synaptic signalling processes. Cluster 2 represents a similar steady up regulation of genes till 288 hours in culture but then a down-regulation after this point. This cluster of genes are involved in the secretory pathway and vesicle-mediated transport processes. Cluster 16 represents a steady down-regulation of initially highly expressed genes, which correspond to genes involved in protein biosynthesis and ribosome function. Finally, cluster 21 represents a complex up regulation and subsequent down-regulation of genes and corresponds to processes associated with cytoplasmic organisation and dendritic morphogenesis.

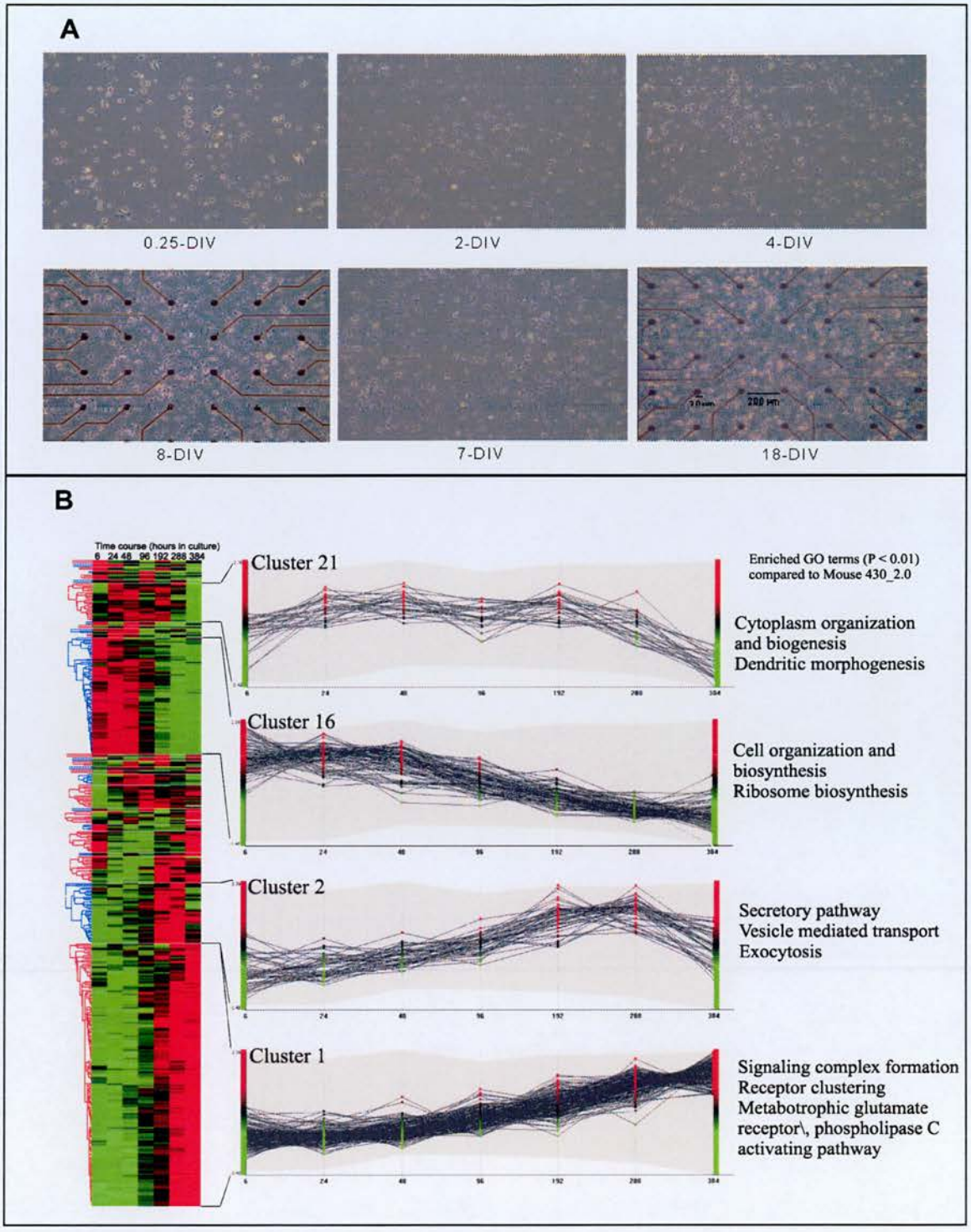
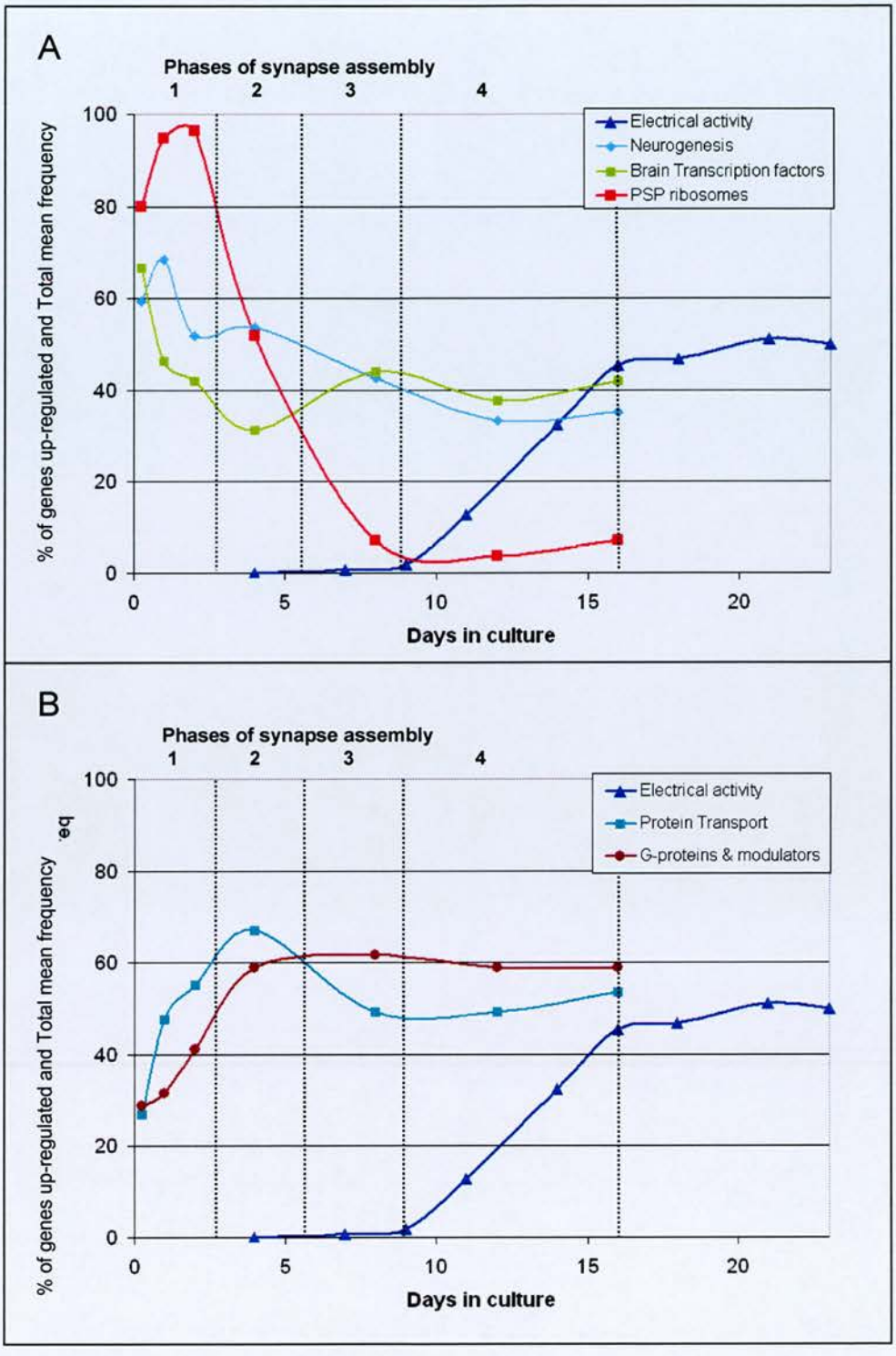
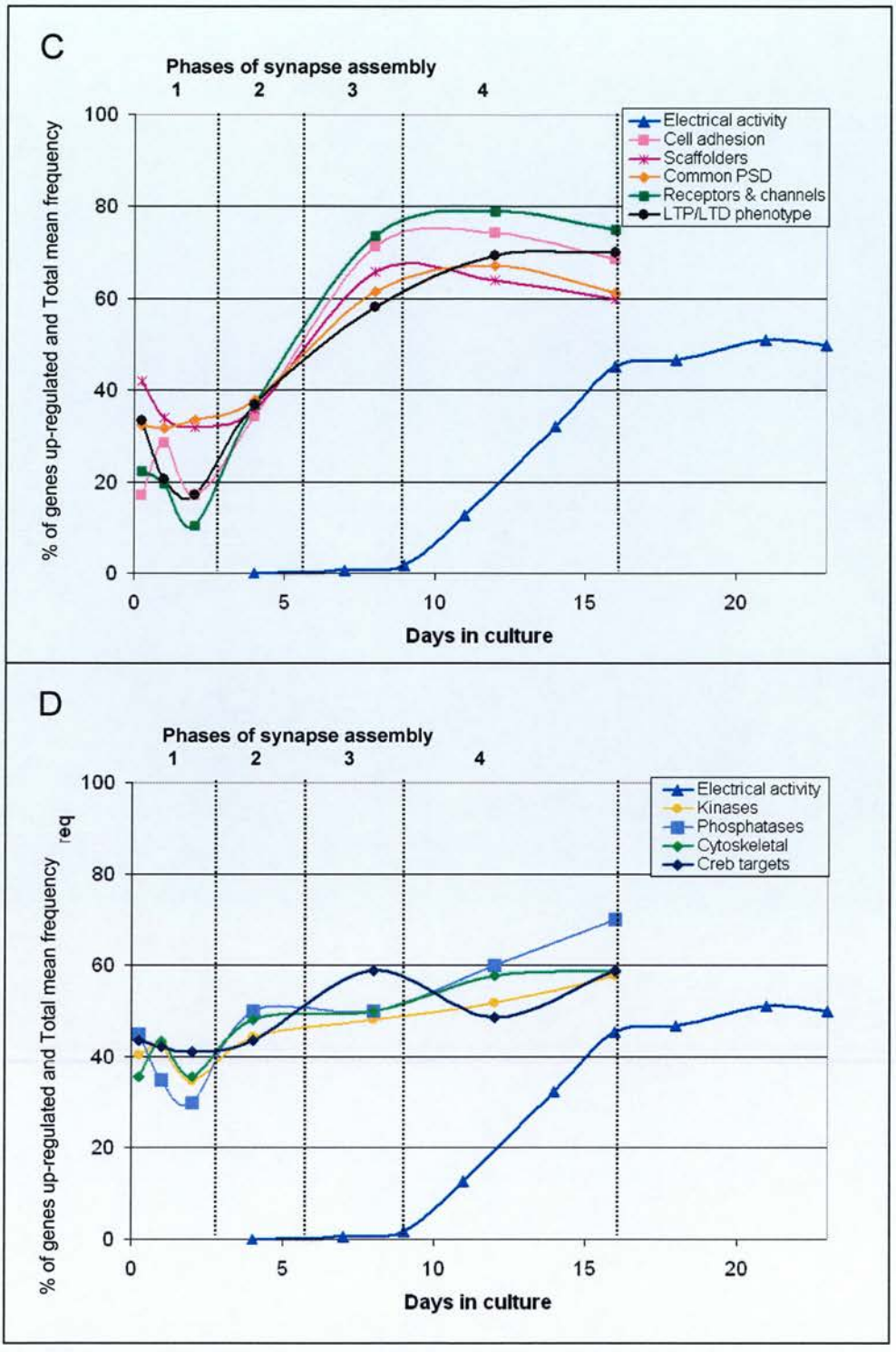


Figure 5.10 – Profiling of developing hippocampal cultures
A - Hierarchical clustering of PSP gene expression time course data from developing hippocampal primary neurons. PSP expression data from 7 time points (0, 24, 48, 96, 192, 286, 384 hrs) was obtained from Mouse 430_2.0 affymetrix array analysis of hippocampal primary neurons. Hierarchical clustering of PSP expression data with a minimum similarity of 0.8 resulted in 4 significant clusters. The significantly (P <0.01) over-represented GO terms associated with proteins (compared to the Mouse 430_2.0 affymetrix array) in each of these clusters is indicated on the right of this figure.
B - Development of hippocampal neurons in culture. Representative contrast phase pictures of neurons at the indicated times (with or without the electrodes of the multielectrode arrays) showing normal development.

5.4.8.2 Time course analysis of synapse assembly

In order to observe detailed temporal profiles, we focused in this case on specific classes of proteins in the PSP. In total 12 PSP classes were analysed, of which (ribosomal (55), protein transport (67), G-proteins and modulators (73), cell adhesion (35), receptors and channels (76), scaffolders (50), kinases (51), phosphatases (20) and cytoskeletal proteins (104)) were defined by Swiss Prot keywords (Supp table 2). In addition, 54 proteins with the Swiss Prot keyword neurogenesis (11 in PSP), Neural specific transcription factors (93), (Gray et al., 2004), Creb targets (78) (Impey et al., 2004) and 29 activity dependent genes (supp table adg) (8 in PSP) were analysed. Genes with a normalised expression value of > 0 at a given time point were considered to be up-regulated and the percentage of genes up-regulated for each class are represented in Figure 5.11. Time course analysis of these defined gene/protein classes revealed four distinct expression phases in developing hippocampal neurons. Phase 1 is characterised by high expression of neural specific transcription factor genes, which decreases to its average expression at the end of this phase (2.5 days). Phase 2 is characterised by the expression of genes involved in protein transport, followed by a plateau of the steadily increasing expression of G-proteins and modulators. Phase 3 represents the main phase of up-regulation of the core synaptic machinery including receptors and channels, scaffolding proteins and cell adhesion molecules. Phase 4 shows that the necessary molecular machinery is in place to produce electrical activity and for the up-regulation of activity-dependent genes.


A
Phases of synapse assembly
1
2
3
4
100
80
60
40
20
0
% of genes up-regulated and Total mean frequency
Electrical activity
Neurogenesis
Brain Transcription factors
PSP ribosomes
0
5
10
15
20
Days in culture
B
Phases of synapse assembly
1
2
3
4
100
80
60
40
20
0
% of genes up-regulated and Total mean frequency
Electrical activity
Protein Transport
G-proteins & modulators
0
5
10
15
20
Days in culture


C
Phases of synapse assembly
1
2
3
4
% of genes up-regulated and Total mean frequency
Days in culture
Electrical activity
Cell adhesion
Scaffolders
Common PSD
Receptors & channels
LTP/LTD phenotype
D
Phases of synapse assembly
1
2
3
4
% of genes up-regulated and Total mean frequency
Days in culture
Electrical activity
Kinases
Phosphatases
Cytoskeletal
Creb targets

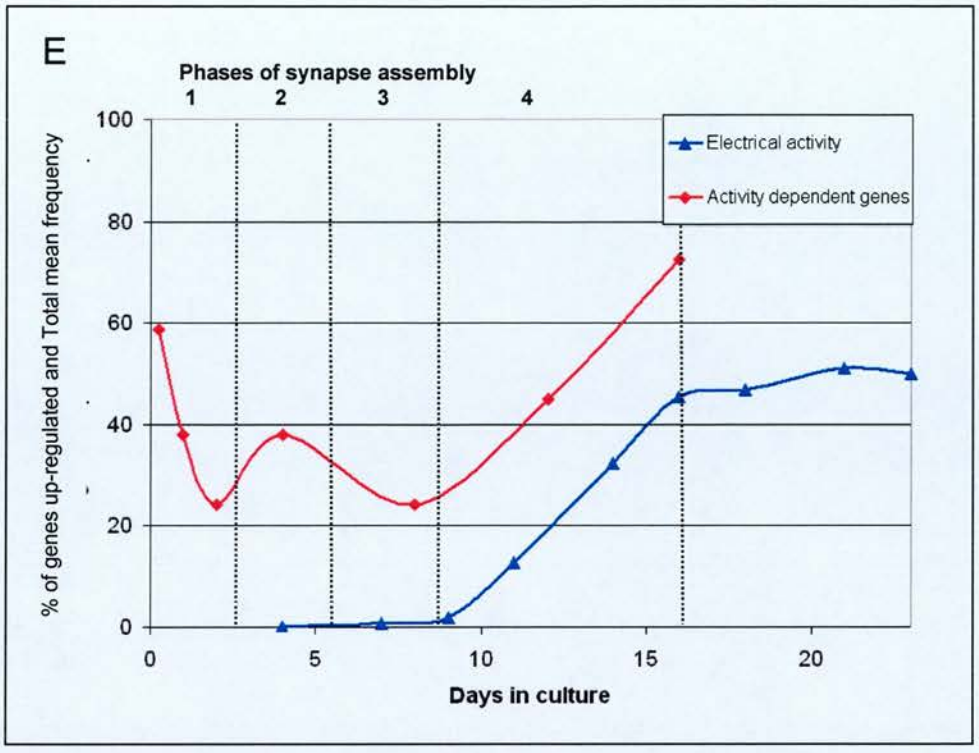
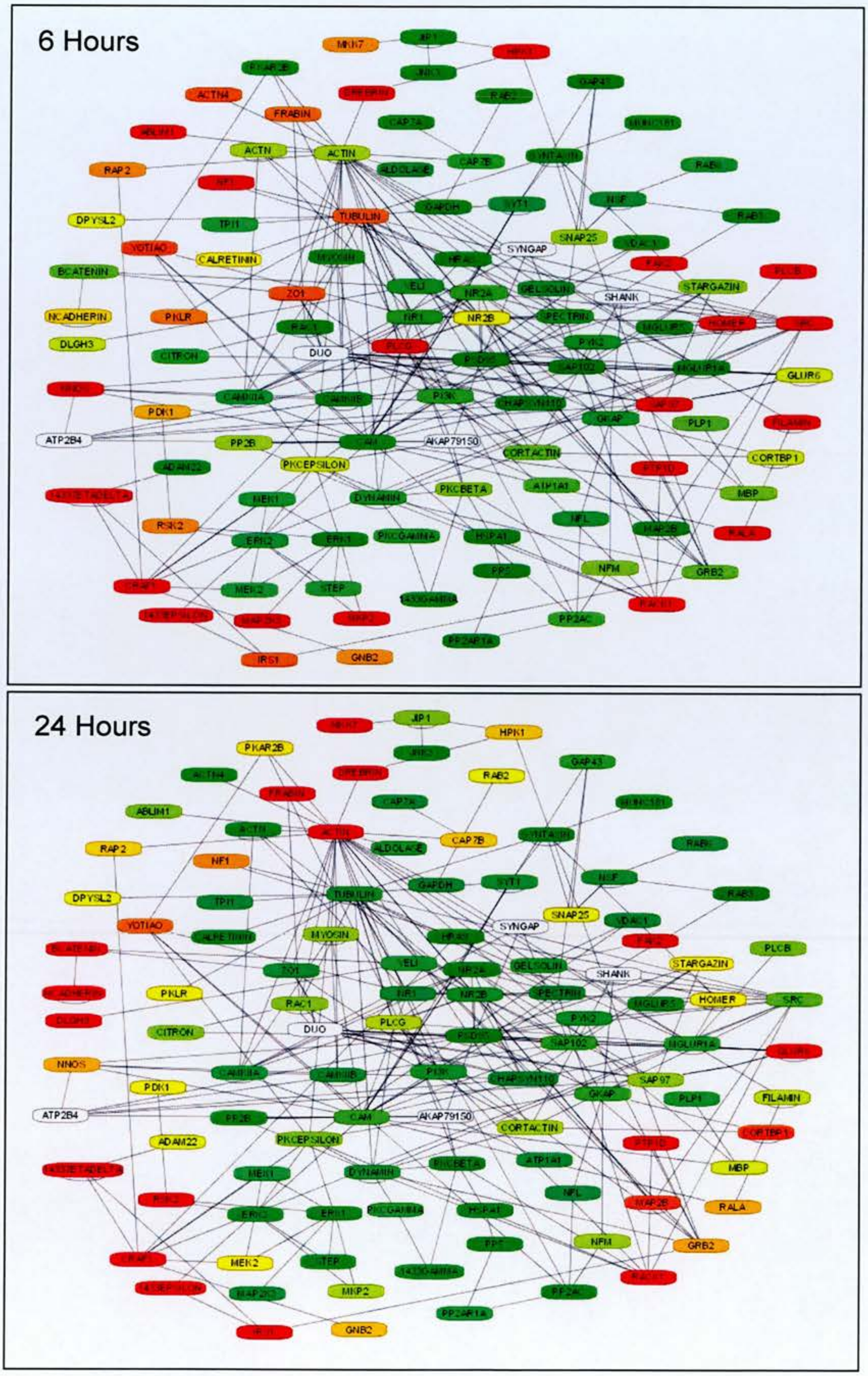


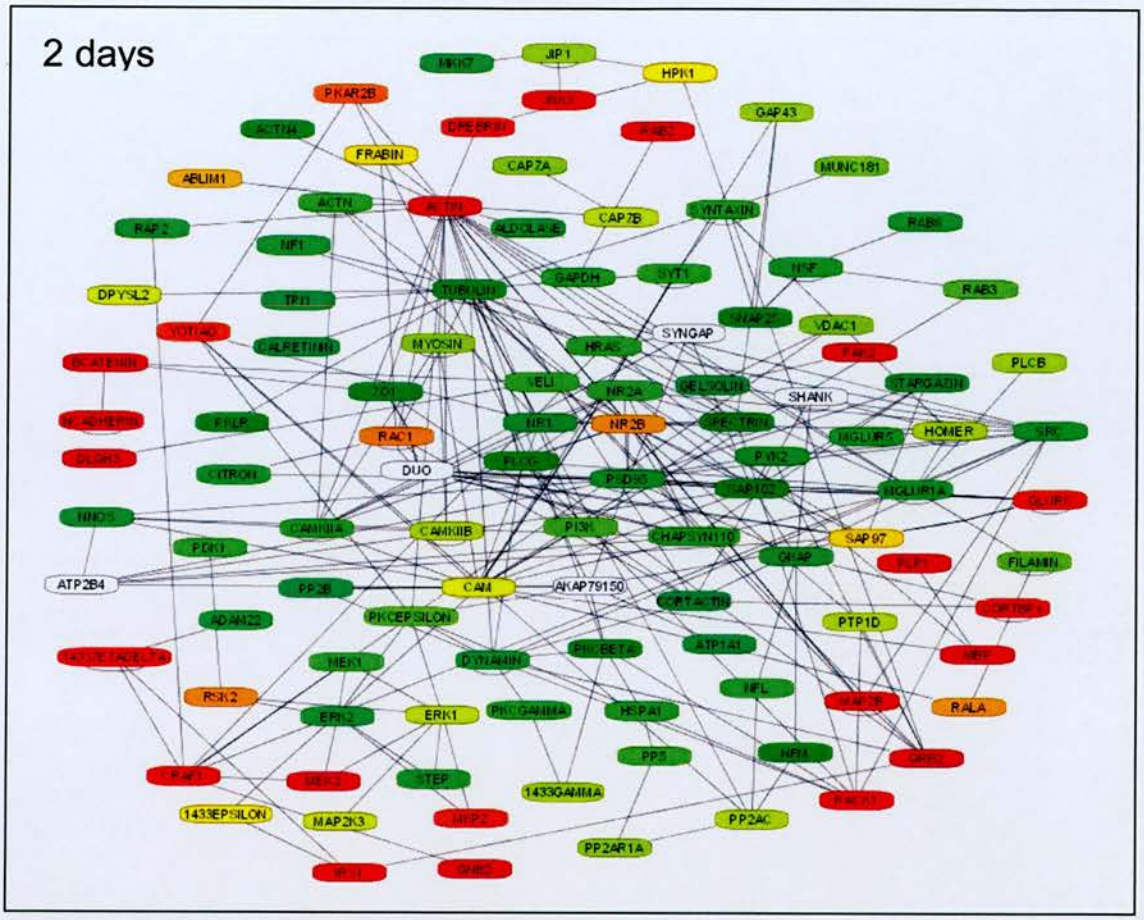
Figure 5.11 - Phases of synapse assembly determined by analysis of time course gene expression data for PSP protein classes. Normalised gene expression data for classes of PSP proteins was used to define potential phases of synapse assembly. The percentage of genes up-regulated at a given time point is the % of genes with a normalised expression value of > 0 and at each time point and represents the average up or down regulation of PSP protein classes. The total mean frequency is the measure of electrical activity in the cultured neurons at each time point. This analysis revealed 4 main time courses of gene expression during neuronal development which correspond to stages of synapse assembly and ultimately to formation of mature neurons with robust electrical activity and up-regulation of activity regulated genes

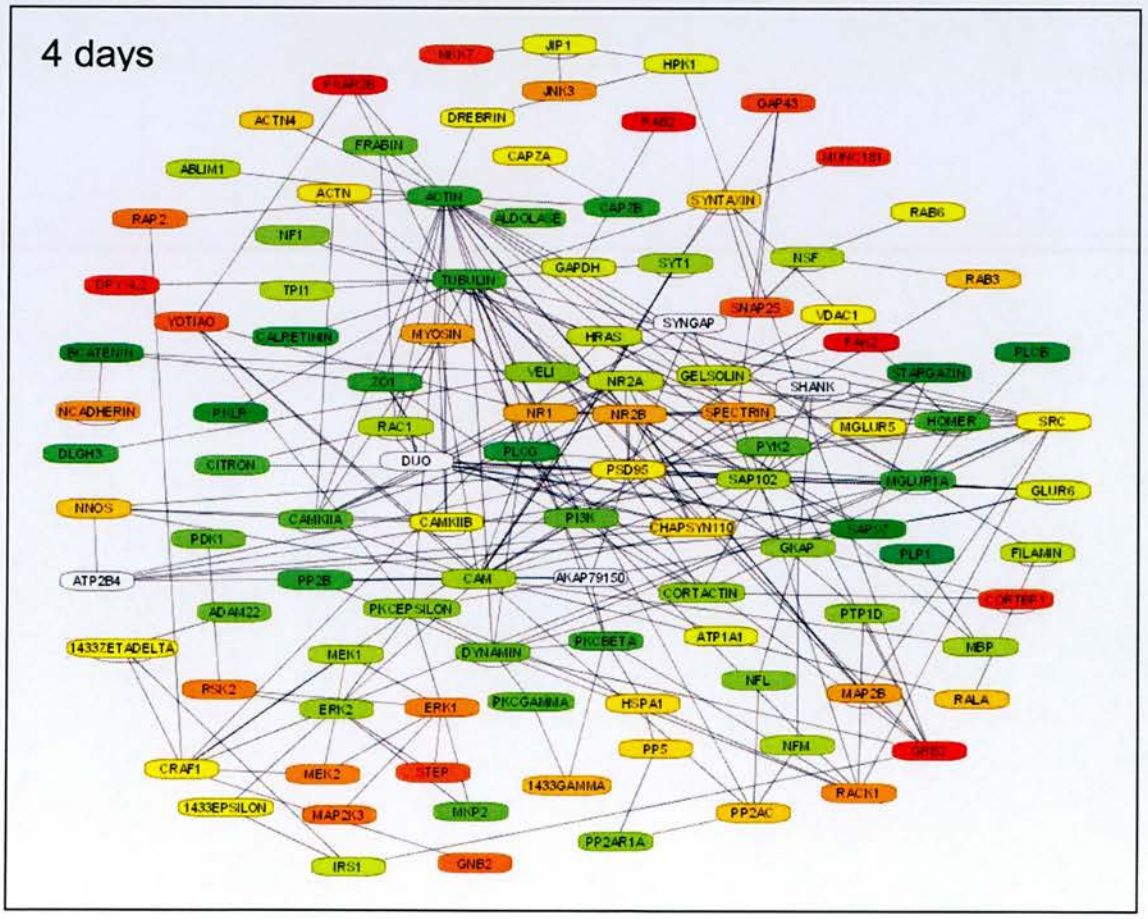
5.4.8.3 Co-expression and protein interactions

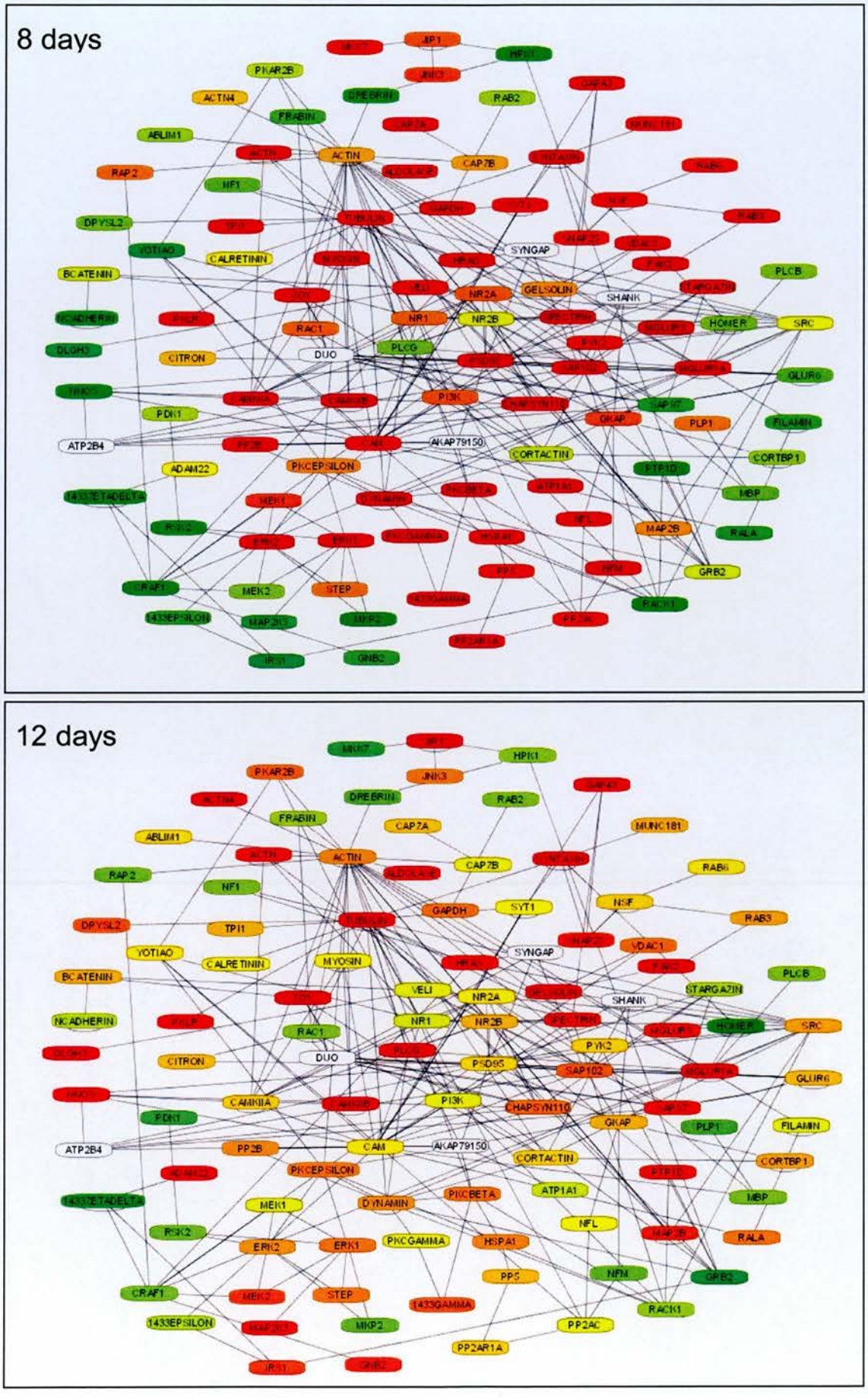
Interacting proteins are more likely to be involved in similar cellular functions and processes and are more likely to be co-expressed and co-regulated by sets of transcription factors. It has been shown that interacting proteins are more likely to be present in the same expression clusters than random sets of proteins in yeast (Ge et al., 2001). To investigate this phenomenon in the PSP, protein interaction data for MASC was integrated with gene expression data extracted for MASC genes from the primary neuronal gene expression experiments. This was achieved by using Cytoscape (Shannon et al., 2003), which enables construction of protein interaction networks and allows integration of processed gene expression data.

Normalised MASC gene expression data was colour-coded from -2 (low expression (green)) to +2 (high expression (red)) and overlaid with the protein interaction network for the entire time course of gene expression (Figure 5.12). The main trend observable from this analysis is that there are two sub-groups of proteins in this expression/interaction network. One group (~ 25 %) is highly expressed early in the primary culture experiment and their expression decreases with neuronal development. This does not preclude a function in MASC in mature neurons because although the expression is low at the end of the time course they are not absent. It is possible that these proteins have a developmental function at non-synaptic sites and then an alternative role in MASC in mature neurons. The majority of MASC genes in this network however, start with a low expression level but increase in a similar profile to high expression levels at the end of the time course. This does not prove that co-expressed synaptic proteins are more likely to interact with one another but when compared to the time course analysis of PSP protein classes (Figure 5.11) it can be seen that the majority of the components of the MASC complex seem to be co-regulated.


6 Hours
FRABIN
ACTN
ACTIN
RAP2
DPYSL2
TUBULIN
YOTIAO
CALRETININ
BCATENIN
NCADHERIN
PKLR
DLGH3
ZO1
DUO
SYNGAP
SNAP25
SHANK
STARGAZIN
NR2B
GLUR6
PDK1
ATP2B4
PP2B
AKAP79150
PKCEPSILON
CORTACTIN
CORTBP1
PKCBETA
MBP
PLP1
NFM
GRB2
PP2AC
GNB2
24 Hours
JIP1
HPK1
PKAR2B
RAB2
ABLIM1
RAP2
CAP7B
NF1
DPYSL2
SNAP25
YOTIAO
MYOSIN
SYNGAP
STARGAZIN
SHANK
PKLR
HOMER
PLCB
CITRON
DUO
PLCG
NNOS
PDK1
SAP97
FILAMIN
ATP2B4
AKAP79150
CORTACTIN
ADAM22
PKCEPSILON
MBP
RALA
GRB2
MEK2
MKP2
GNB2


2 days
JIP1
HPK1
GAP43
FRABIN
CAP7A
MUNC18I
ABLIM1
ACTN
CAP7B
SYNTAXIN
RAP2
RAB6
NF1
ALDOLASE
GAPDH
SYT1
NSF
RAB3
DPYSL2
TPI1
TUBULIN
SNAP25
VDAC1
YDTIAO
CALRETININ
MYOSIN
HRAS
SYNGAP
PLCB
ZO1
VELI
NR2A
GELSOLIN
STARGAZIN
SHANK
NCADHERIN
PHLP
NR1
NR2B
SPECTRIN
MGLUR5
HOMER
SRC
RAC1
DLGH3
CITRON
DUO
PLCG
PYK2
PSD95
SAP102
MGLUR1A
GLUR6
NNOS
CAMKIIA
CAMKIIB
PI3K
CHAPSYN110
SAP97
PDK1
GKAP
PLP1
FILAMIN
ATP2B4
PP2B
CAM
AKAP79150
CORTACTIN
ADAM22
PKCEPSILON
PTP1D
ATP1A1
MEK1
DYNAMIN
PKCBETA
NFL
RSK2
RALA
ERK2
ERK1
PKCGAMMA
HSPA1
PP5
CRAF1
MEK2
STEP
1433GAMMA
1433EPSILON
MAP2K3
MNP2
PP2AC
PP2AR1A


4 days
JIP1
HPK1
JNK3
GAP43
DREBRIN
ACTN4
FRABIN
CAP7A
MUNC18I
ABLIM1
ACTN
CAP2B
SYNTAXIN
RAP2
RAB6
NF1
ALDOLASE
GAPDH
SYT1
NSF
RAB3
TPI1
TUBULIN
SNAP25
VDAC1
YDTIAO
CALRETININ
MYOSIN
HRAS
SYNGAP
PLCB
ZO1
VELI
NR2A
GELSOLIN
STARGAZIN
SHANK
NCADHERIN
PHLP
NR1
NR2B
SPECTRIN
MGLUR5
HOMER
SRC
RAC1
DLGH3
CITRON
DUO
PLCG
PYK2
PSD95
SAP102
MGLUR1A
GLUR6
NNOS
CAMKIIA
CAMKIIB
PI3K
CHAPSYN110
PDK1
GKAP
PLP1
FILAMIN
ATP2B4
PP2B
CAM
AKAP79150
CORTACTIN
ADAM22
PKCEPSILON
PTP1D
14337ETADELTA
MEK1
DYNAMIN
PKCBETA
ATP1A1
MBP
RSK2
NFL
MAP2B
RALA
ERK2
ERK1
PKCGAMMA
HSPA1
PP5
NFM
CRAF1
MEK2
STEP
1433GAMMA
RACK1
1433EPSILON
MAP2K3
MNP2
PP2AC
PP2AR1A
IRS1
GNB2


8 days
JIP1
PKAR2B
JNK3
ACTN4
DREBRIN
RAB2
FRABIN
ABLIM1
ACTIN
CAP2B
RAP2
NF1
DPYSL2
TPI1
YOTIAO
CALRETININ
SYNGAP
BCATENIN
PLCB
NR2A
GELSOLIN
SHANK
NCADHERIN
NR1
NR2B
HOMER
SRC
RAC1
DLGH3
CITRON
DUO
PLCG
GLUR6
PI3K
PDK1
GKAP
PLP1
FILAMIN
ATP2B4
AKAP79150
CORTACTIN
CORTBP1
ADAM22
PKCEPSILON
MBP
MAP2B
RALA
GRB2
CRAF1
MEK2
STEP
1433EPSILON
IRS1
12 days
JIP1
HPK1
PKAR2B
JNK3
DREBRIN
RAB2
FRABIN
CAP2A
MUNC18I
ABLIM1
ACTIN
CAP2B
RAP2
NF1
GAPDH
SYT1
NSF
RAB3
DPYSL2
TPI1
VDAC1
YOTIAO
CALRETININ
MYOSIN
SYNGAP
PLCB
BCATENIN
VELI
NR2A
STARGAZIN
SHANK
NCADHERIN
NR1
NR2B
HOMER
SRC
RAC1
PYK2
DUO
PSD95
SAP102
GLUR6
CAMKIIA
PI3K
CHAPSYN110
PDK1
GKAP
PLP1
FILAMIN
ATP2B4
PP2B
CAM
AKAP79150
CORTACTIN
CORTBP1
PKCEPSILON
ATP1A1
MBP
1433ETADELTA
MEK1
DYNAMIN
PKCBETA
NFL
RSK2
MAP2B
RALA
ERK2
ERK1
PKCGAMMA
HSPA1
PP5
NFN
GRB2
CRAF1
MEK2
STEP
1433GAMMA
RACK1
1433EPSILON
MKP2
PP2AC
PP2AR1A
IRS1

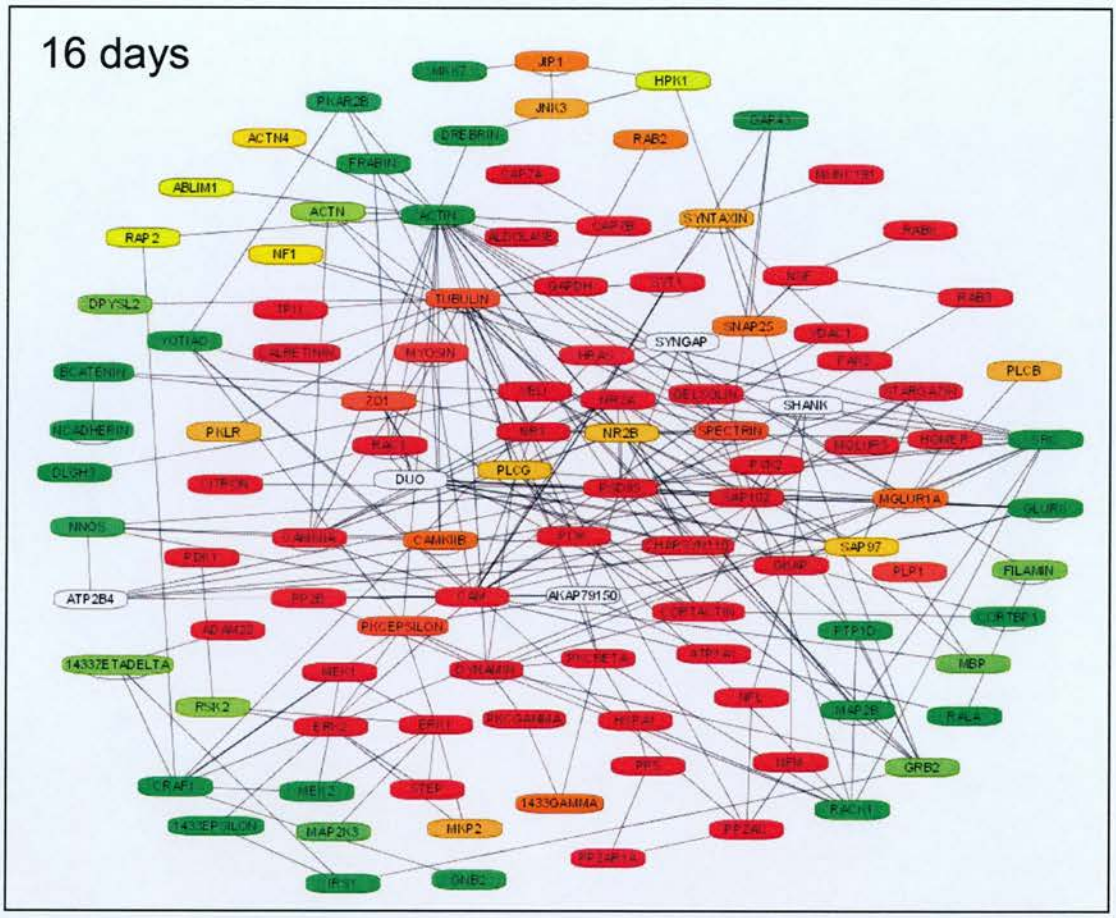


Figure 5.12 - Network of protein interactions and dynamic gene expression in MASC. Protein-protein interaction data (www.PPID.org) was used to construct this network onto which gene expression data from hippocampal primary neuronal cultures was overlaid. Normalised gene expression data was used and represented in colour from green (low expression) to red (high expression) and white (no expression data). It can be seen in the time course that the expression of most proteins is up-regulated over time, although the expression profile of some proteins is inverse to this indicating possible developmental requirements for these proteins.

5.4.9 Tissue expression transcriptome profiling of synapse proteomes

To extend our gene expression analysis of the PSP we mined the gene expression atlas and retrieved data for 464 unique PSP genes across 45 mouse tissues (13 brain/neuronal tissues). Hierarchical clustering analysis with a minimum similarity of 0.5 produced three broad clusters (Figure 5.13). Significant enriched ($p > 0.01$) GO terms associated with genes in each cluster were determined and are summarised in Figure 5.13. Cluster 1 contains genes whose expression is lower in the 13 neuronal tissues than the other tissues included in this analysis. The majority of GO terms enriched in this cluster correspond to biosynthesis processes especially translation with a number of GO terms referring to ribosome or protein biosynthesis. Cluster 3 also contains genes whose expression is lower in the 13 neuronal tissues than the other tissues included in this analysis and the majority of GO terms enriched in this cluster correspond to catabolic and metabolic processes. In contrast to cluster 1 and 3, cluster 2 contains genes highly enriched in neuronal tissue and in particular a sub-cluster (marked by a red box, Figure 5.13) in which the gene expression is not only high in neuronal tissues but appears to be nearly exclusive to neuronal tissues. The majority of GO terms enriched in this cluster correspond to classical synaptic functions, from ion transport to a variety of signalling processes. The distribution of cPSD and MASC genes were mapped to these gene expression clusters and it can be seen that both the cPSD and MASC are more prominent in the highly expressed/neuronal enriched cluster 2.

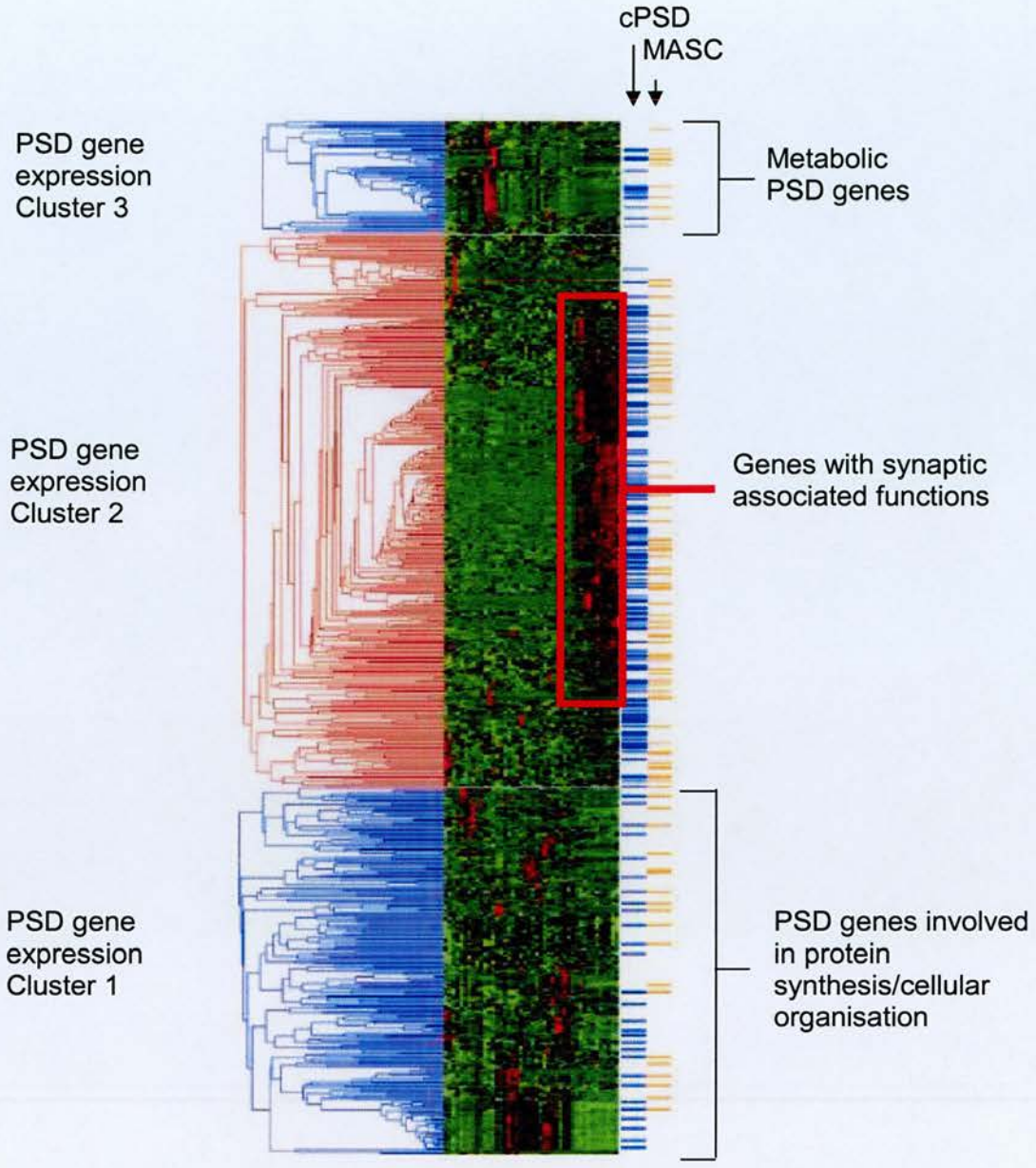


Figure 5.13 – Hierarchical clustering of PSP gene expression from 45 tissues. Gene expression data for PSP proteins from the Gene expression atlas was clustered to identify neuronal tissue enriched gene expression. This resulted in identification of 3 clusters of expression and GO terms significantly enriched in these clusters are summarized on the right of this figure. A sub-cluster of high gene expression over 13 nervous tissues (red box) can be seen in cluster 2 which is associated with a higher density of MASC and cPSD (orange and blue horizontal lines, respectively) genes.

5.4.10 Evolutionary of analysis of PSP genes

The enrichment of signalling domains in PSP proteins compared to the whole proteome (section 5.4.5) correlates with the expansion of these domains in the genome, with the evolutionary step from single to multi-cellular organisms (Manning et al., 2002). In addition, in the tissue expression transcriptome analysis (section 5.4.9) we observed that one quarter of the PSP genes analysed appear to be specific to nervous tissue. These two observations support a model where proteome diversification was necessary for the evolution of neuronal function and synapses in particular. In order to investigate this further, ortholog analysis for PSP genes was performed using data available for 696 PSP genes, which included the G2C PSD and MASC datasets. Hierarchical clustering of PSP gene orthologs from C. *elegans*, A. *gambiae*, D. *melanogaster*, D. *rerio*, S. *cerevisiae*, G. *gallus*, R. *norvegicus*, M. *musculus* and H. *sapiens* was performed to identify taxon specific genes (Figure 5.14).

This ortholog analysis resulted in the identification of five major clusters of PSP genes. Cluster 1 is the largest group of genes and corresponds to PSP genes present in vertebrates only and represents 34% (234 genes) of the total dataset. Cluster 2 corresponds to higher vertebrate specific PSP genes and represents 12% (85 genes) of the total dataset. Cluster 3 corresponds to genes absent in C. *elegans* and S. *cerevisiae*, which represents 8% (56 genes) of the total dataset. Cluster 4 corresponds to genes absent in S. *cerevisiae* only and is a significant proportion (25% (174 genes)) of the total dataset. Interestingly cluster 5 corresponds to PSP genes conserved in all phylogenetic classes analysed, which means that 9% (64 genes) of the PSP proteome is composed of genes with yeast orthologs.

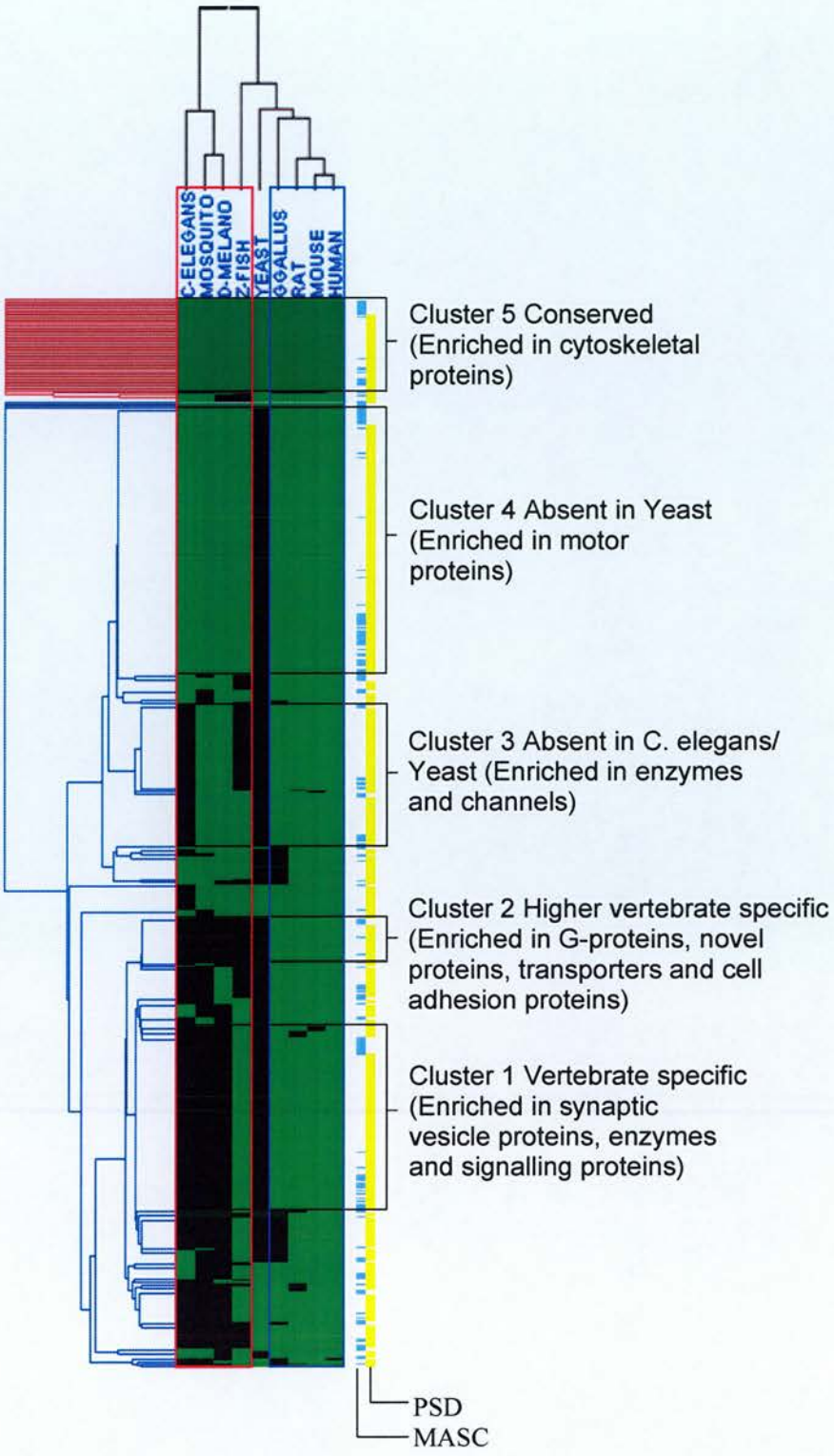
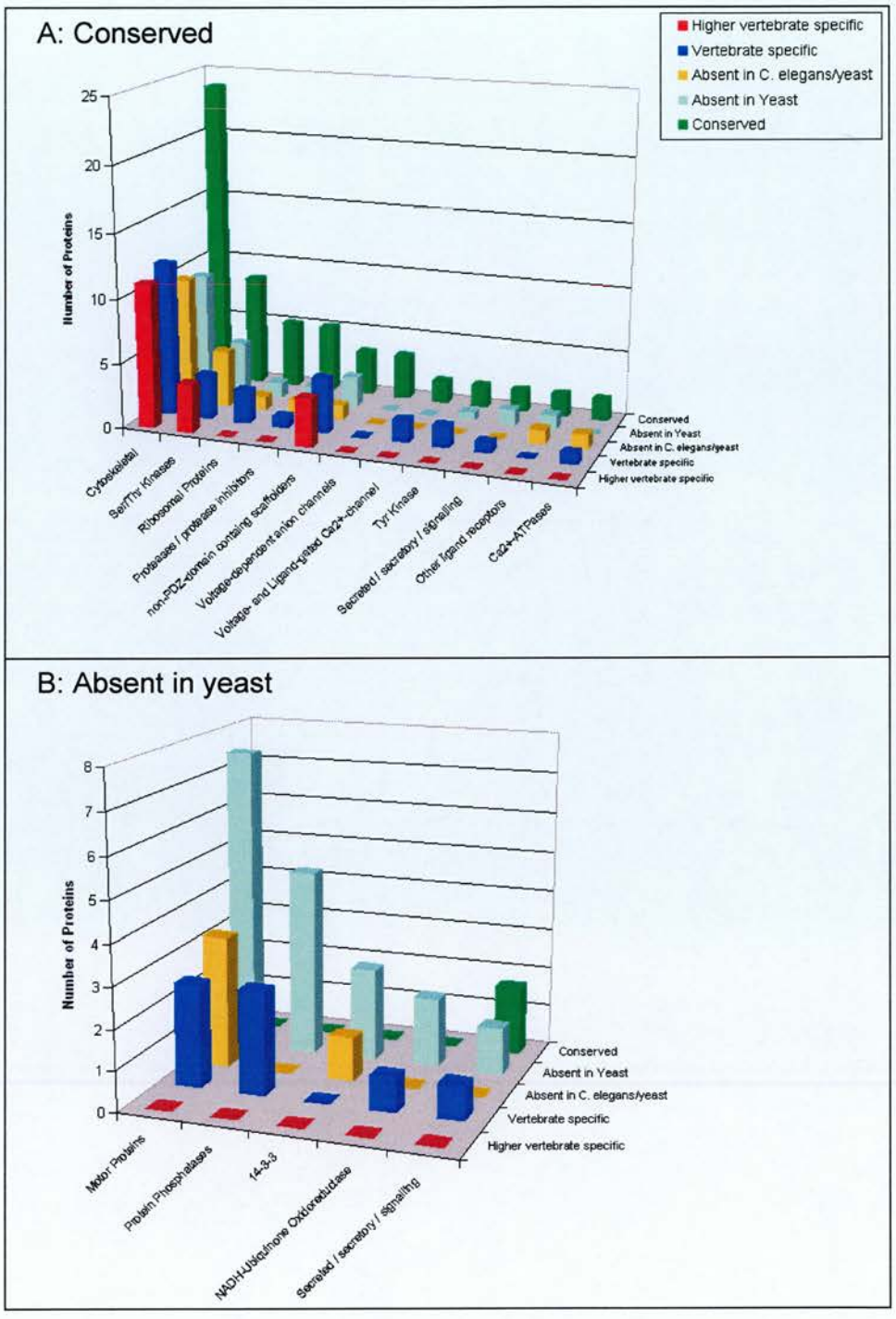
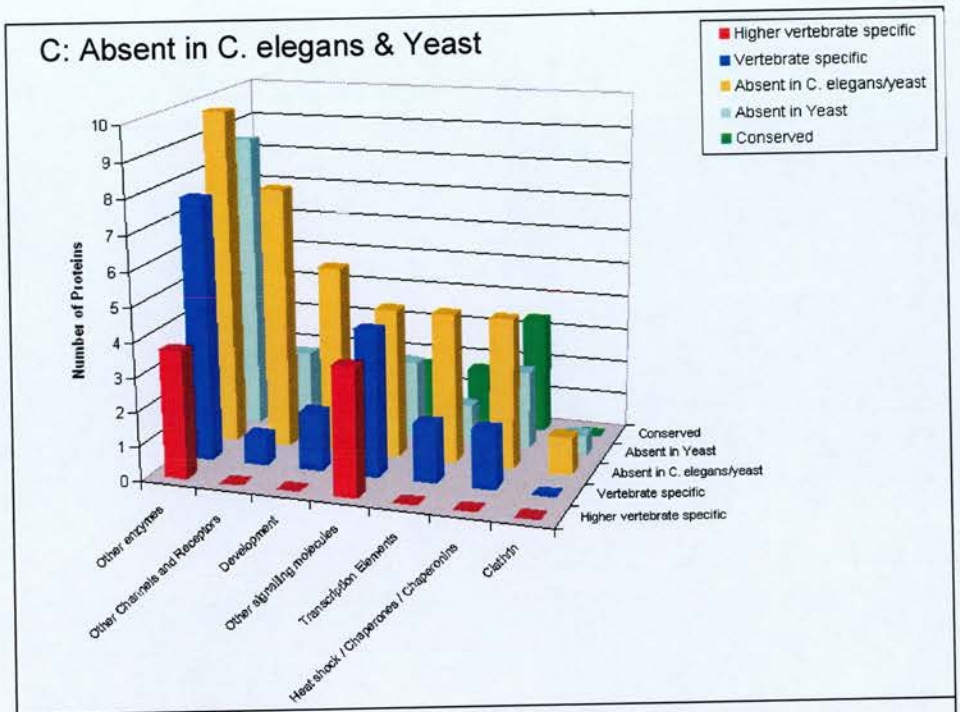


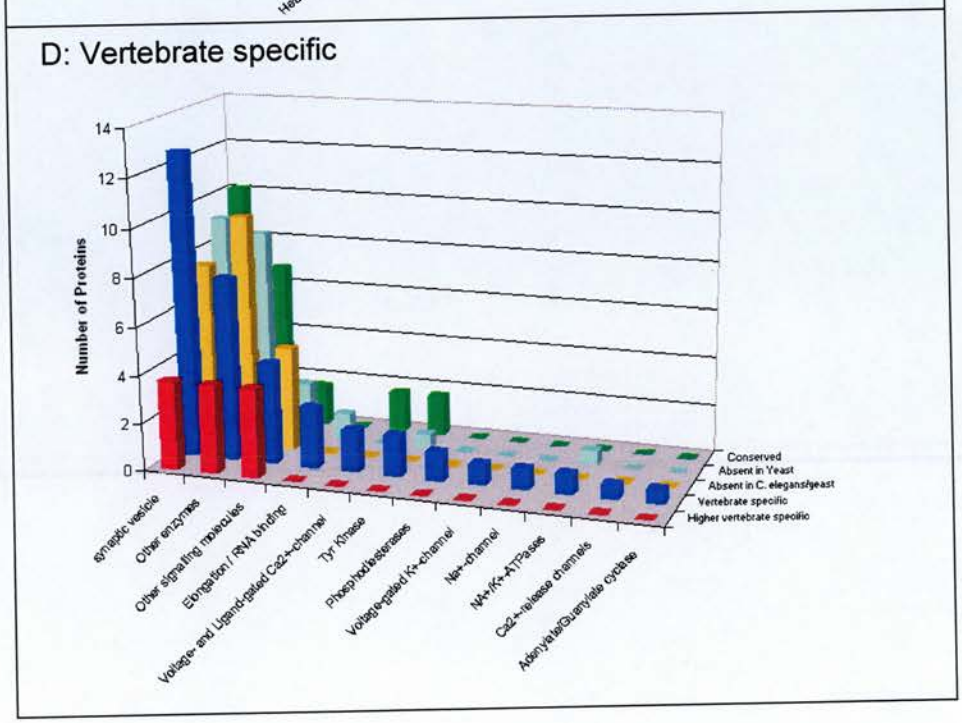
Figure 5.14 – Hierarchical clustering of PSP gene orthologs. Orthologs from C. *elegans,* A. *gambiae* (Mosquito), D. *melanogaster,* D. *rerio* (Zebrafish), S. *cerevisiae,* G. *gallus,* R. *norvegicus*, M. *musculus* and H. *sapiens* were clustered to identify taxon specific sets of genes. It can be seen that a small set of genes are conserved in all species which comprise ancient components of the PSP. Clusters with genes absent in yeast and absent in C. *elegans* and S. *cerevisiae* were found and in addition, vertebrate and higher vertebrate specific genes were identified.

Is there are relationship between these PSP evolutionary clusters and the genes which comprise them? We again used proteomic data to extract functional information from genetic data. In this case we analysed the distribution of PSP protein classes in each evolutionary cluster (Figure 5.15). The most ancient class of proteins found in this analysis was cytoskeletal proteins, being the most enriched class in the conserved cluster (Cluster 5) (Figure 5.15A). This is expected because even the simplest organisms utilise a cytoskeleton for cell support and transport. Cluster 4 corresponds to orthologs present in all taxa studied except for yeast, shows a clear enrichment (2 fold) for motor proteins (Figure 5.15B). Transport mechanisms in unicellular eukaryotes such as yeast, expanded to meet the increased demand for transport in multicellular organisms (Vale, 2003). Prokaryotes use filamentous systems for various transport processes, whereas eukaryotes emerged with an efficient system consisting of a number of molecular motors, each of which was designed to carry different types of cargo. The distinguishing feature of multicellular eukaryotes compared to unicellular ones with respect to these motors appears to be the increased diversity and regulation of these processes (Vale, 2003). Thus, the enrichment of motor proteins in cluster 4 seems not to reflect a newly acquired process in multicellular eukaryotes but rather a reflection of the increased complexity of regulation of this process to meet increased transport demands.

Cluster 3 is characterised by orthologs absent in C. *elegans* and S. *cerevisiae* and is enriched for enzymes and channels and receptors probably indicating an expansion of these types of proteins after C. *elegans* (Figure 5.15C). Cluster 2 corresponds to orthologs present in higher vertebrates only and shows enrichment for G-proteins, novel proteins, transporters and cell adhesion proteins (Figure 5.15E). Cluster 1 contains orthologs that are present in vertebrates only and there is enrichment for synaptic vesicle proteins, enzymes and signalling proteins (Figure 5.15D). Several synaptic vesicle genes have orthologs in yeast that are associated with the secretory pathway (Bennett and Scheller, 1993) however, as opposed to constitutive secretory


A: Conserved
Higher vertebrate specific
Vertebrate specific
Absent in C. elegans/yeast
Absent in Yeast
Conserved
Number of Proteins
0
5
10
15
20
25
Cytoskeletal
Ser/Thr Kinases
Ribosomal Proteins
Proteases / protease inhibitors
non-PDZ-domain containing scaffolders
Voltage-dependent anion channels
Voltage- and Ligand-gated Ca2+-channel
Tyr Kinase
Secreted / secretory / signalling
Other ligand receptors
Ca2+-ATPases
Conserved
Absent in Yeast
Absent in C. elegans/yeast
Vertebrate specific
Higher vertebrate specific
B: Absent in yeast
Number of Proteins
0
1
2
3
4
5
6
7
8
Motor Proteins
Protein Phosphatases
14-3-3
NADH/Ubiquinone Oxidoreductase
Secreted / secretory / signalling
Conserved
Absent in Yeast
Absent in C. elegans/yeast
Vertebrate specific
Higher vertebrate specific


C: Absent in C. elegans & Yeast
Higher vertebrate specific
Vertebrate specific
Absent in C. elegans/yeast
Absent in Yeast
Conserved
Number of Proteins
0
1
2
3
4
5
6
7
8
9
10
Other enzymes
Other Channels and Receptors
Development
Other signalling molecules
Transcription Elements
Heat shock / Chaperones / Chaperonins
Clathrin
Conserved
Absent in Yeast
Absent in C. elegans/yeast
Vertebrate specific
Higher vertebrate specific


D: Vertebrate specific
Number of Proteins
0
2
4
6
8
10
12
14
synaptic vesicle
Other enzymes
Other signalling molecules
Elongation / RNA binding
Voltage- and Ligand-gated Ca2+-channel
Tyr Kinase
Phosphodiesterases
Voltage-gated K+-channel
Na+-channel
NA+/K+-ATPases
Ca2+-release channels
Adenylate/Guanylate cyclase
Conserved
Absent in Yeast
Absent in C. elegans/yeast
Vertebrate specific
Higher vertebrate specific

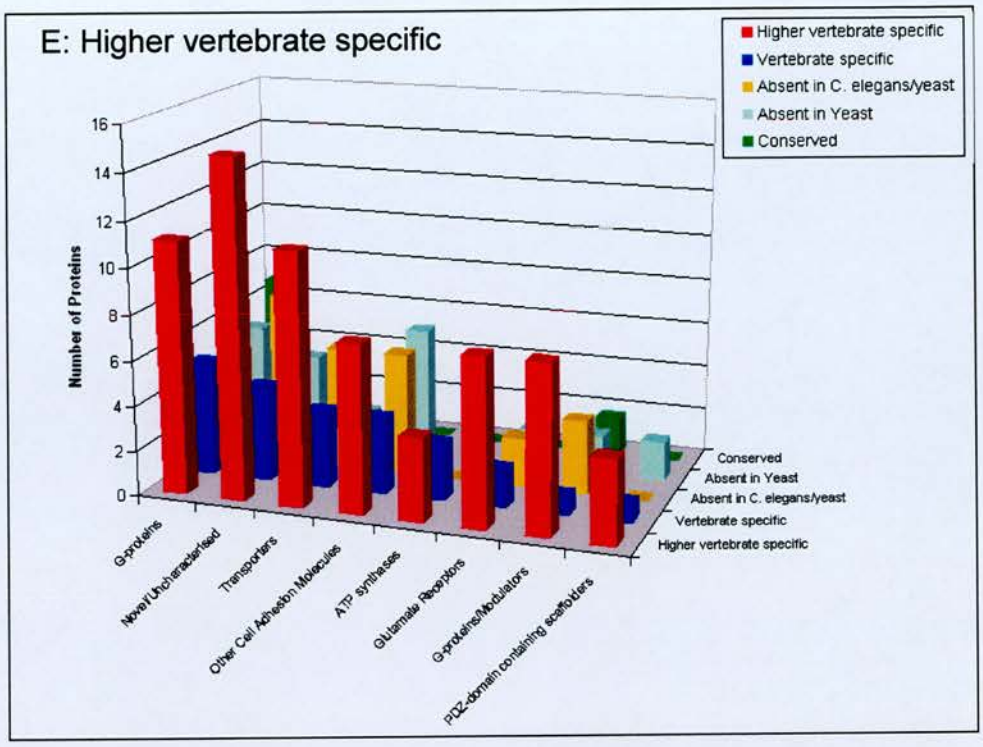


Figure 5.15 Evolution of PSP protein classes. Ortholog clusters generated as shown in Figure 5.15 were used to analyse the enrichment of specific classes of proteins in the PSP. The top enriched protein classes for each evolutionary cluster are shown from panels A to E. The most conserved protein classes are cytoskeletal and kinases and the most recently evolved classes include G-proteins, glutamate receptors and PDZ-containing scaffolders.

pathways neurosecretion is regulated by calcium. Neurons therefore contain additional regulatory components for regulated secretion and these may be contributing to the enrichment of vesicle proteins in this vertebrate specific evolutionary cluster (Volknandt, 1995).

5.4.11 Protein-protein interactions in the PSP

To facilitate analysis of the organisation of the PSP at the level of protein-interactions an in-house database called PPID (www.PPID.org) was used. This database is a mammalian (mouse, rat, human) Protein-Protein Interaction Database (PPID) describing approximately 8000 protein interactions for > 1700 proteins. 650 protein-protein interactions for 281 PSP proteins are shown in Figure 5.16. The individual proteins in this network are represented by filled circles (nodes) and the interactions are represented by connecting lines (edges). The distribution of PSP interactions show a clear bias toward proteins in MASC that can be explained by the fact the MASC set was the starting point for data mining. The majority of interaction data for non-MASC proteins in the PSP are for cPSD proteins. This is because of the large overlap between the MASC and cPSD datasets and also possibly due to the fact the highly abundant proteins may be more likely to be studied than low abundant proteins.

This interaction network was combined with phenotypic data in a manner analogous to that for integration of MASC interaction and gene expression data (section 5.4.8). An in-house database of mouse targeted mutations, in which a change in LTP or LTD was observed for a given gene knockout was used. PSP genes for which such mutant phenotypic data was available are coloured in green in Figure 5.16. 19 MASC genes, 29 cPSD and 8 Total PSD (Not in MASC or cPSD) display changes in LTP/LTD when mutated. Although, at present this PSP interaction network is incomplete, it indicates a potential way of identifying potentially interesting genes for mutagenesis and follows the observed strong

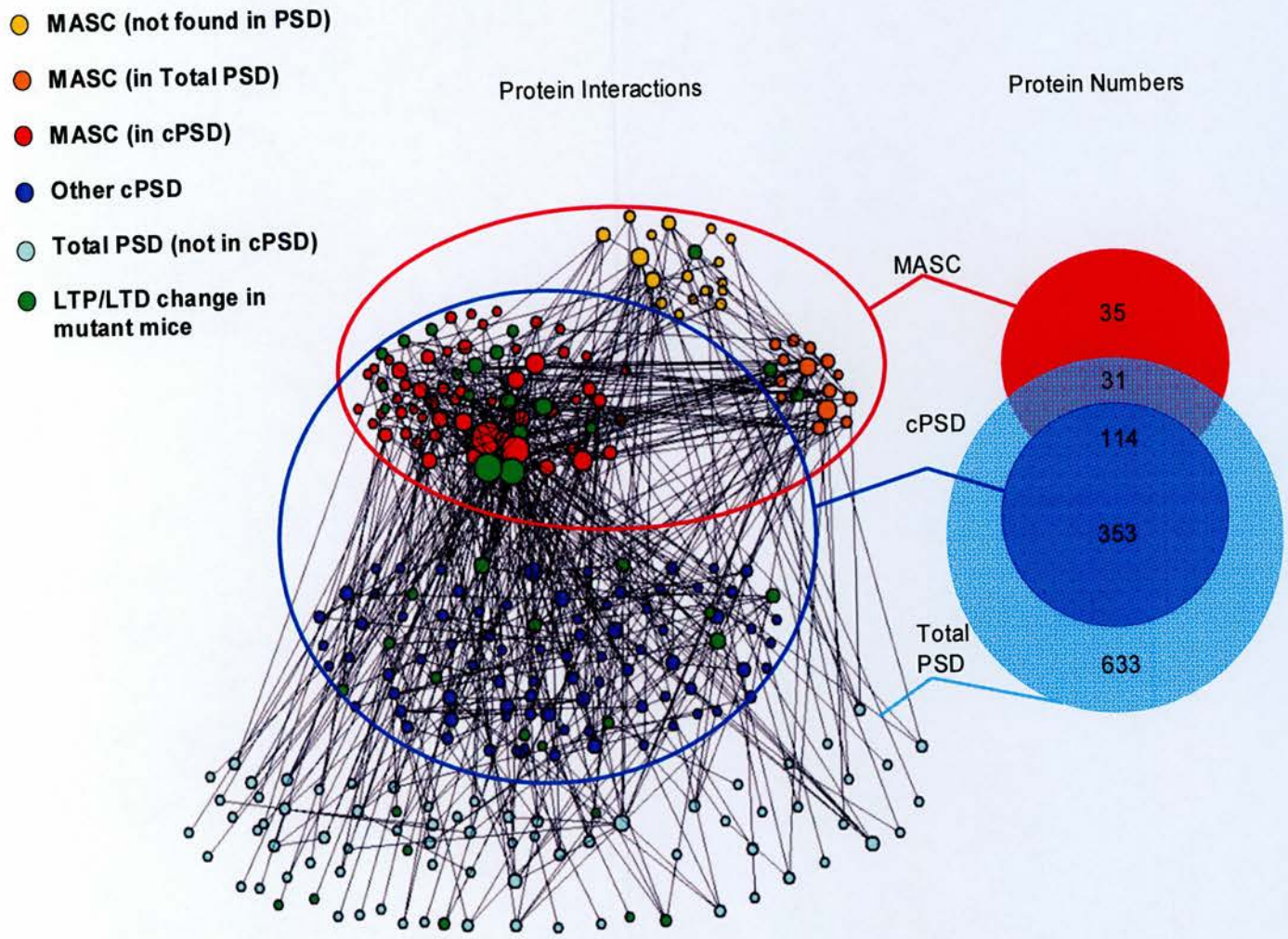


Figure 5.16 - Network illustration of protein interactions and knockout mouse phenotype in PSP proteins. 650 protein-protein interactions for 281 PSP proteins are shown. These PSP interactions are divided into MASC (Red circle), cPSD (Blue circle) and nodes outside the circles are other components of the Total PSD. Also, a venn diagram of the total numbers of proteins in the MASC, cPSD and Total PSD is shown. Proteins in this network for which an electrophysiological phenotype (change in LTP/LTD) has been reported in the literature are indicated by green nodes. The size of a node is proportional to its connectivity (number of interactors) and it can be seen that the largest nodes with an electrophysiological phenotype are in MASC

relationship between number of interactors and phenotype in S. cerevisiae (Jeong et al., 2001).

5.4.12 Construction of an Alzheimer's disease protein interaction network.

The postsynaptic proteome contains many prominent candidate proteins for neurodegenerative diseases such as Alzheimer's disease (AD). We reasoned that these proteins may have similar functions or are involved in sets of pathways that cause or affect AD. To address this we investigated the relationship between these AD-associated proteins by constructing a protein-protein interaction network. The starting point for this was a dataset of proteins, which were up- or down-regulated in mice transgenically overexpressing glycogen synthase kinase-3β (GSK-3β), a kinase with a prominent role in neurodegenerative disease (Kaytor and Orr, 2002). GSK-3β is a serine-threonine protein kinase capable of phosphorylating tau and hyperphosphorylated tau is a principle component of neurofibrillary tangles, which are characteristic of AD brains. It is regulated by serine (Inhibitory) and tyrosine (stimulatory) phosphorylation and by protein complex formation and its intracellular localisation (Figure 5.17). GSK-3β is a key regulator of neuronal plasticity and gene expression controlling the function of many metabolic, signalling and structural proteins. It also regulates cell survival, as it facilitates a variety of pro-apoptotic mechanisms [Harwood, 2001 #30]. It has been linked to all primary abnormalities associated with AD and was shown to interact with components of the plaque producing amyloid system [Grimes, 2001 #31].

A comparison between GSK-3β transgenic mice and wild type revealed 51 proteins whose expression differed significantly (Tilleman et al., 2002). Twenty-five of these proteins are present in the postsynaptic proteome, 14 of which are components of NMDAR-PSD-95 complexes that constitute MASC. There was a significant decrease in the relative abundance of cytoskeletal and energy metabolism proteins and a significant increase in proteins involved in signal transduction and oxidative stress

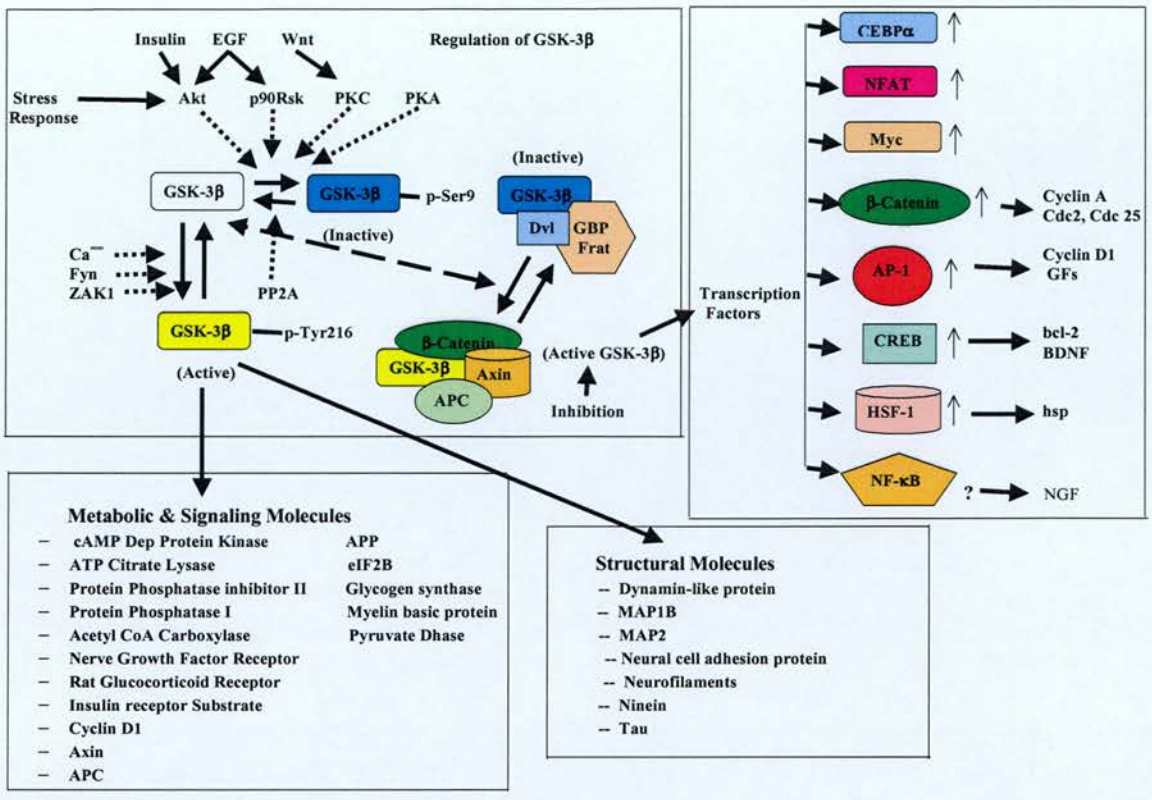


Figure 5.17 – Regulation and functions of GSK3β.
Adapted from "The multifaceted roles of glycogen synthase kinase 3beta in cellular signaling." Grimes CA, Jope RS. Prog Neurobiol. 2001 Nov;65(4):391-426.

in the GSK-3β transgenic mice. A similar study by the same group was performed on mice over expressing human Tau (Tilleman et al., 2002). Proteins were resolved by 2D-PAGE and 34 proteins whose expression levels in wild-type and tau transgenic mice differed at least 1.5 fold were identified by ESI-MS [Tilleman, 2002 #27].

A comparison of the profile of proteins with altered expression in the Tau and GSK-3β transgenic mice showed that there were a number of proteins with similar changes consistent with the mutual involvement of these two proteins in the process of neurodegeneration.

Using the data obtained from the study of GSK-3β and Tau constitutively active mutant mice by Tilleman et al, we attempted to create a global map of AD disease association. This was achieved by constructing a protein-protein interaction map (Figure 5.18) of proteins implicated in AD pathology. Proteins found to have altered expression in the GSK3–β constitutively active mutant mice were extensively searched for interactors in PPID (Protein-Protein Interaction database (Husi, 2002) and in Pubmed. In addition, these initial proteins as well as the accumulating interactors were extensively mined in PubMed to ascertain whether they had previously been implicated in AD. These proteins were deemed implicated from reports of altered expression in post-mortem brains or proteins, which have been shown to affect the basic process of amyloid deposition and NFT formation. GSK-3β phosphorylation substrates relevant to the disease were also included. This data mining was carried out until satisfactory interaction information for the initial set of proteins was obtained. This yielded 118 proteins in the interaction map, of which 97 are implicated in AD and 27 proteins were included for completeness of the interaction network. This global approach of analysing proteins implicated in AD provides a platform to evaluate mouse models of disease. It can be clearly seen that many of the proteins altered in GSK-3β constitutively active mutant mice are

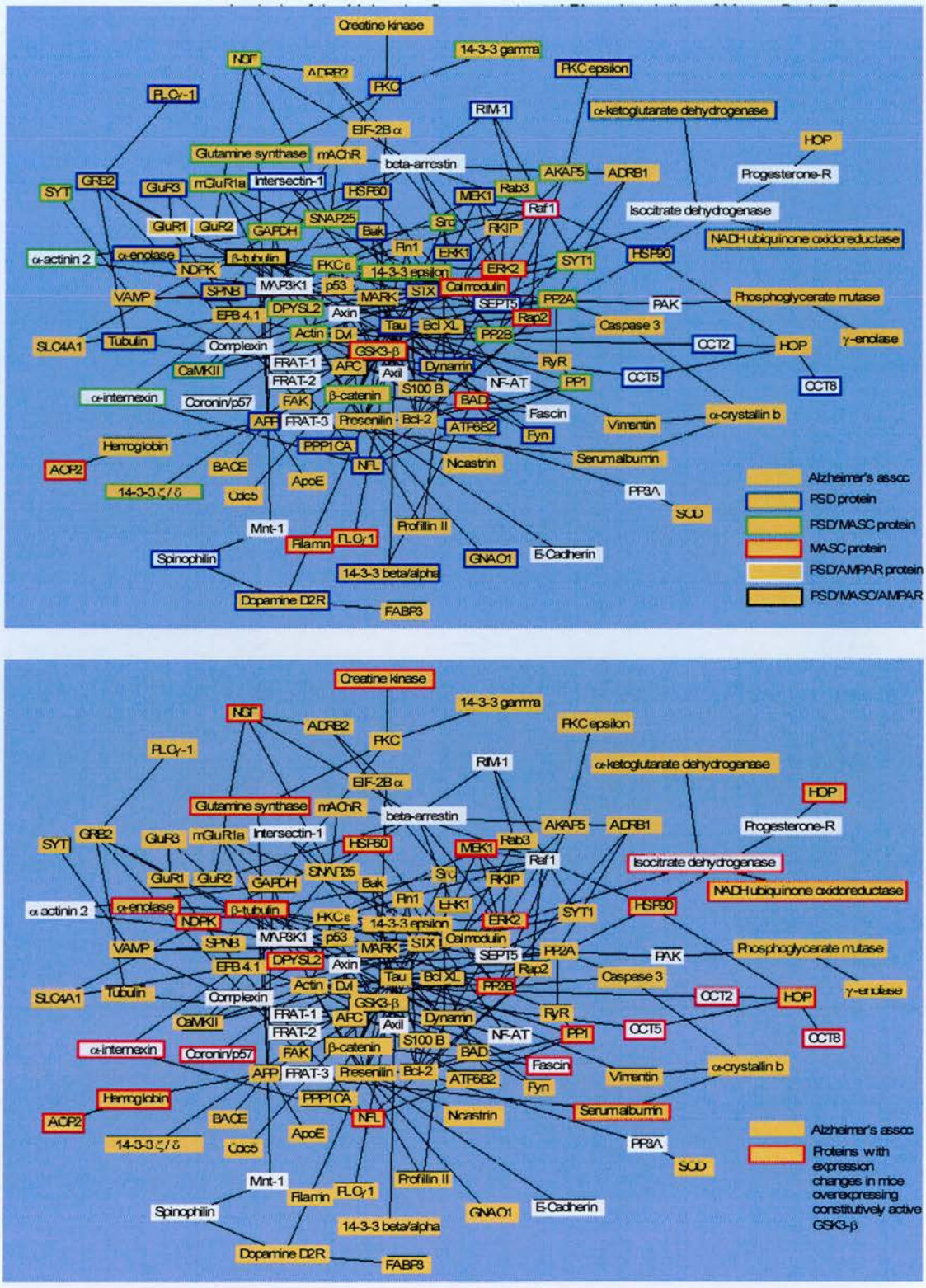


Figure 5.18 – Alzheimer's disease protein interaction network. This interaction network was constructed from literature mining for AD candidate proteins and protein interaction data from PPID. **A.** Interaction network showing the overlap between AD candidate proteins and postsynaptic proteome datasets. **B.** Evaluation of GSK3β as a mouse model for AD. Proteins with altered expression in GSK3β overexpressing mice are overlaid onto the AD protein interaction network.

implicated in AD, and interacting proteins which were not previously implicated may be novel AD associated proteins. Key proteins responsible for the pathological hallmarks of AD such as Tau, GSK-3β, β–catenin, amyloid-precursor protein, presenilin are present and are highly connected. They are closely linked to functional classes of proteins such as the cytoskeleton, membrane receptors, modulators of oxidative stress and cell death and numerous signalling molecules. Components of this network have roles in pathways, which may relate to functional impairments seen in AD. These range from synaptic transmission, stress response, lipid transport, glycolysis and cytoskeletal organisation. The number of proteins implicated in AD is quite striking and the fact that so many of these proteins can cluster together into an extensively connected network (over 280 interactions) creates confidence in this network approach. This points to the possibility of the presence of pathways and processes present in the network that are core to the pathogenesis of the disease

Further analysis of this network in relation to functional complexes at the synapse revealed that 54 out of the identified 97 proteins associated with AD are components of the PSD, 32 are found in MASC and three are present in the AMPA receptor complex (Figure 5.18). This also highlighted a large overlap of 30 proteins common to the PSD and MASC that are implicated in AD. It is probable that there are subsets of proteins within the large number of proteins associated with AD that represent distinct mechanisms leading to the formation of amyloid plaques and NFT's characteristic of AD.

5.5 DISCUSSION

5.5.1 Proteomic analysis of the postsynaptic proteome

We have presented a comprehensive analysis of isolated postsynaptic densities and embedded multi-protein complexes associated with NMDA and AMPA receptors. We have combined existing synapse proteomic data with that reported here to provide the most complete picture of the postsynaptic proteome to date. To our knowledge, the postsynaptic proteome is one of the most complicated subcellular structures analysed, but unlike other organelles, the PSP appears to hold a high degree of cellular autonomy. The composition of the PSP in terms of functional protein classes is diverse, with the necessary molecular machinery for not only classical synaptic processes and a myriad of signalling pathways, but for a host of supporting activities such as protein synthesis and degradation, transport and metabolism. Moreover, 10% of PSP proteins are novel and may participate in synaptic processes yet unknown.

The large number of proteins found at the PSD and the presence of many ribosomal, mitochondrial and nuclear proteins has prompted speculation over the purity of isolated PSD fractions used in proteomic analyses. There are numerous examples where such proteins were validated as true PSD proteins. Some proteins such as voltage-dependent anion channel 1 (VDAC1), which was originally found in the outer mitochondrial membrane, is present in the PSD (Moon et al., 1999). Nuclear proteins such as AIDA-1A, mKIAA0417 and hnRNP localise to both the nucleus and the PSD (Jordan et al., 2004). It is apparent that many proteins do not exist in discrete locations or indeed posses the same functions in different cellular contexts.

The data presented and compiled in this study is a draft of an average isolated postsynaptic density. The starting material for the analyses incorporated into our draft PSP are from different species (mouse and rat) and from different brain

regions (forebrain and whole brain) and encompass hundreds of neuronal cell types. We are now at a point where it is becoming feasible to compare different types of synapse proteomes. This could be achieved by correlation profiling (Andersen et al., 2003), a quantitative proteomic approach where synaptosomes and isolated PSD's are isotopically tagged ((ICAT or iTRAQ) and the resultant quantitative MS data would show enrichment or depletion of proteins in the PSD compared to synaptosomes. This type of analysis would provide confident assignment of true PSD proteins, highlight contaminating proteins and would validate the PSD localisation of proteins currently considered restricted to other cellular compartments.

In addition, isolated PSD's or protein complexes from different brain regions or cultured cell types could be quantitatively compared using isotopic labelling. Alternatively, peptides derived from SILAC (Stable Isotope Labelling with Amino acids in Cell culture) labelled cells could be used to quantitatively compare the expression of proteins in many tissue samples (Ishihama et al., 2005). Using these approaches, proteins enriched in a particular cell type or brain region would be identified as being potentially more relevant to the physiology of that particular cellular or tissue type. It is at this stage where systematic immuno-cytochemistry of these enriched proteins would be feasible and would provide complementary information regarding their specific protein localisation at the PSD and throughout the cell type in question. Current projects such as Gene Expression Nervous System Atlas (GENSAT) in which large-scale in-situ hybridisation and BAC transgenic reporter genes are being used to map gene expression in neuronal cell types will provide complimentary information to the proteomic experiments suggested here.

5.5.2 PSP organisation

Mass spectrometry-based analysis of proteomes is routinely producing long lists of proteins that describe the composition of a particular proteome. The usual type of

bioinformatic analysis carried out on this data involves categorising the protein components according to functional class (using varied criteria) and by Gene Ontology (GO) terms. The usefulness of this basic approach is limited, and the value of such proteome lists is debatable without the appropriate analysis to extract biologically meaningful information, which describes the proteome in question. We believe that integration of diverse data types with proteome data and analysis of this integrated data provides an appropriate way to address the complex biology of a proteome. In this study, we used data from multiple levels of cellular complexity and integrated it with our proteomic data to ask how the postsynaptic proteome is organised?

5.5.2.1 Genes

The genes that code for proteins in the postsynaptic proteome are diverse and as the synapse itself is a cell-cell contact that specialised through evolution, we investigated the possible origins of these genes. We performed ortholog analysis for PSP genes using data available for 696 PSP genes, which included the G2C PSD and MASC datasets. Hierarchical clustering of PSP gene orthologs from C. *elegans*, A. *gambiae*, D. *melanogaster*, D. *rerio*, S. *cerevisiae*, G. *gallus*, R. *norvegicus*, M. *musculus* and H. *sapiens* was performed to identify taxon specific genes (Figure 5.14). Five major evolutionary clusters were found, which included higher vertebrate specific, vertebrate specific, absent in C. *elegans* and S. *cerevisiae*, absent in S. *cerevisiae* only and finally genes conserved in all phyla analysed. This clustering result clearly showed that components of the PSP evolved in stages. To gain more insight into these evolutionary clusters we analysed the distribution of PSP protein classes in each cluster. The most ancient class of proteins, the first building blocks of the synapse is cytoskeletal proteins. Genes associated with protein transport are the

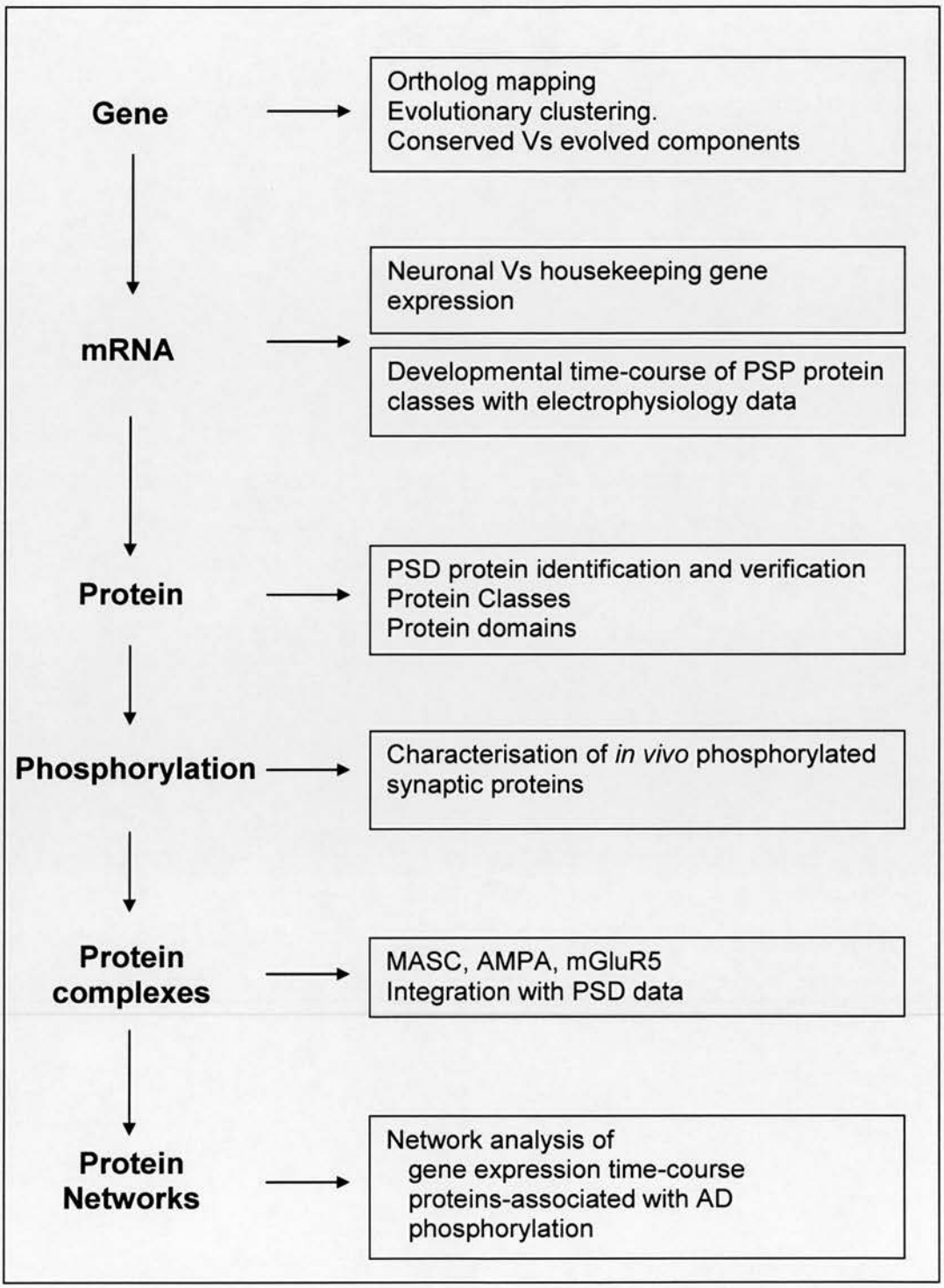


Figure 5.19 – Scheme of PSP complexity and the analysis applied to each level of organisation

next acquisition in evolutionary time, allowing trafficking of newly synthesised proteins out to the dendrite.

Next, in the cluster representing genes absent in C. *elegans* and S. *cerevisiae,* we see enrichment for enzymes, channels and receptors indicating an expansion of these types of proteins after C. *elegans*. In the vertebrate only cluster, there is an enrichment for synaptic vesicle proteins, enzymes and signalling proteins, presumably necessary to deal with the increasing complexity of synapse function. In the higher vertebrate cluster, there is enrichment for G-proteins, novel proteins, transporters and cell adhesion proteins; these novel proteins would be good candidates for further characterisation. This analysis has provided insight into the timing of the appearance of sets of synaptic genes and also the appearance of the functions associated with these genes.

5.5.2.2 mRNA expression

The next level of complexity that was investigated is the transcription of PSP genes into mRNA. Two data sources were used for this analysis; tissue gene expression in adult mice from the Gene Expression Atlas, and a time-course of genes expression in developing primary hippocampal neuronal culture generated in-house. First, we asked are there PSP-specific genes. To address this, we mined the gene expression atlas and retrieved data for 464 unique PSP genes across 45 mouse tissues (13 brain/neuronal tissues). This resulted in three broad gene expression clustered and to assign general functions to the genes in these clusters we looked a significantly enriched GO terms associated with genes in each cluster. The first cluster contained genes whose expression was lower in neuronal tissues than the other tissues analysed. This cluster was enriched for GO terms associated with biosynthesis processes especially translation with a number of GO terms referring to ribosome or protein biosynthesis. Similarly, cluster three contained genes whose expression is lower in the 13 neuronal tissues than the other tissues included in this analysis and

the majority of GO terms enriched in this cluster correspond to catabolic and metabolic processes. However, cluster two showed strong expression in neuronal compared to non-neuronal tissues and the majority of GO terms enriched in this cluster correspond to classical synaptic functions, from ion transport to a variety of signalling processes.

This tissue expression analysis of the PSP reveals yet another level of organisation and hints at the evolutionary origins of the synapse. Of the PSP genes included in this analysis it appears that approximately a quarter of these are candidate neuronal specific genes or at least prominent neuronal genes with strong enrichment of synapse-associated GO biological processes. The remainder of PSP genes in this analysis are not specific to neuronal tissues and have generalised housekeeping functions. This supports the view that the synapse is home to a wide variety of cellular processes, from protein biosynthesis to metabolic pathways which, provide a supporting role for the more specialised synaptic processes of ion transport and receptor signalling.

The differentiation of neural precursor cells into synaptically connected neuronal networks with electrical activity can be monitored using primary culture of embryonic mouse hippocampus neurons. Therefore, we used mRNA levels to indicate the timing of expression of PSP genes in developing neuronal cultures. Phase 1 is characterised by high expression of neural specific transcription factor genes, which decreases to its average expression at the end of this phase (2.5 days). Phase 1 also displays dramatic expression of ribosomal genes culminating in 97 % of these genes being expressed by day two and the highest expression of genes associated with the process of neurogenesis (Figure 5.11A). The coordinated up-regulation of components of transcriptional and translational machinery within the first 2 days is expected and necessary for the massive growth of projections/spines that starts at day 4 . After production of ribosomes, expression of genes involved in transport of proteins from the cell body to the growing projections is essential and

the peak expression of such genes characterises phase two of this developmental time course (Figure 5.11B). This is closely followed by a plateau of the steadily increasing expression of G-proteins and modulators in agreement with the notion that G-protein regulated processes are required during neurogenesis as well as in mature synapses (Luo, 2000).

Phase 3 represents the main phase of up-regulation of the core synaptic machinery including receptors and channels, scaffolding proteins and cell adhesion molecules (Figure 5.11C). These three classes of proteins along with the core PSD display a similar profile of up-regulation over time. After the expression, trafficking and assembly of these synaptic molecules by day 10, it can be seen that sufficient synapses have formed and that the necessary molecular machinery are in place to produce electrical activity. It is this synaptic functionality, as measured by electrical activity and the increase thereof, which defines the last phase characterised by gene expression data. The steady increase in activity is accompanied by a plateau in expression of the classes of proteins thus far described. However, the expression of cytoskeletal proteins, kinases and phosphatases that may be regulating the ongoing activity are still increasing in phase four (Figure 5.11D). Activity-dependent genes, i.e. genes that are regulated by NMDA receptor activity in the hippocampus or cortex (supp table x) show a sustained up regulation in this phase and the increase in expression shows a profile similar to that of the total mean frequency of electrical activity (Figure 5.11E). This would indicate that the electrical activity, including the NMDA receptor-mediated component, is controlling the expression of these genes, which are known to be involved in synaptic plasticity.

5.5.2.3 Proteins

Next, the proteins that were characterised in our proteomic data and from other studies were analysed in terms of functional classes and proteins domains. In the PSP there are representatives of many classes of proteins representing a broad range

of cell biological functions were observed: membrane bound receptors, adhesion proteins and channels; a plethora of signalling proteins and adaptors; and proteins involved with regulation of transport, RNA metabolism, translation and transcription.

Protein domains are structured units within proteins and many of these domains have associated functions, so analysis of the occurrence of different domains in a proteome can provide functional insight. The enrichment of the top 10 domains found in the cPSD was compared to the Total and our PSD dataset (Figure 5.6). The greatest enrichment in these datasets is for protein interaction domains, in agreement with the notion that the cPSD contains abundant proteins such as synaptic scaffolders. Protein domains associated with kinase function were the most common in the PSP correlating with enrichment in other domains commonly found in protein kinases, which allow them to interact in networks (Manning et al., 2002). In addition, domains involved with Ca^{2+}-dependent signalling and members of the Ras GTPase superfamily were also highly abundant. The same signalling domains enriched at the synapse are known to have expanded in the genome with the evolutionary step from single to multi-cellular organisms (Manning et al., 2002). 62% of domains found in proteins contained in the Total PSD are present in S. *cerevisiae*, indicating that a large proportion of the basic building blocks of mammalian synaptic proteins are conserved from before the metazoan expansion. The complement of these domains increases after this expansion with C. *elegans*, D. *rerio* and D. *melanogaster* having 85, 86 and 87% of Total PSD protein domains respectively. This would indicate that after this expansion there was an increase in the order of 20-30% in the appearance of new domains as well as domain shuffling to produce new proteins.

5.5.2.4 Protein Phosphorylation

The next level of complexity in the postsynaptic proteome was addressed in chapter 3. Protein phosphorylation is an essential regulator of protein function and is

particularly important in the postsynaptic proteome where very complex signalling pathways are believed to occur. Identification of phosphorylation sites, their cognate kinases and functional significance is critically important for the understanding of the dynamic organisation of the PSP. We investigated phosphorylation in the PSP by identification of phosphorylation sites, their cognate kinases (for a selected set) and construction of a simple three-dimensional network (Figure 3.13). This network consisted of protein interactions within MASC with kinase-substrate interactions superimposed. This basic representation of how phosphorylation can be considered as an extra dimension in protein organisation hints at possible complexity of a complete kinase-substrate interaction map of the PSP.

5.5.2.5 Protein complexes

NMDAR, AMPAR complex and mGluR5 receptor complexes are contained in the postsynaptic proteome and although they were not purified directly from synaptic preparations the majority of their components are in the PSP. In total, there are 80 channels and receptors in the PSP and although these are very unlikely to exist together in a single synapse, it is clear that many of these are likely to be components of other signalling complexes. The idea that a proteome is composed of such functional units of multiprotein complexes is supported by systematic studies in yeast (Ho et al., 2002), (Gavin et al., 2002). The co-operative action of groups of proteins in multiprotein complexes has led to the idea that proteomes are composed of molecular machines (Gavin and Superti-Furga, 2003). These molecular machines have distinct functions but may be composed of many of the same proteins that are present in other machines. We have observed this in the PSP where 21 proteins are common to MASC and mGlur5 complexes. These two complexes may physically coupled as there is data that shows that they are functionally linked (Kotecha et al., 2003). We have addressed organisation within MASC in terms of evolution, gene expression, phosphorylation and networks, as detailed in those sections.

5.5.2.6 Protein networks

Protein interactions are essential to the functionality of proteomes because proteins rarely act alone, but rather in complexes. Large-scale mapping of protein interactions in the yeast proteome reveals highly complex networks of protein interactions (Fromont-Racine et al., 1997), (Gavin et al., 2002). These networks have not only been useful for clustering functionally related proteins, but were recently subject to analysis with mathematical tools developed for non-molecular models such as social networks and the World Wide Web (Han et al., 2004). Interacting proteins are more likely to be involved in similar cellular functions and processes and are more likely to be co-expressed and co-regulated by sets of transcription factors. It has been shown that interacting proteins are more likely to be present in the same expression clusters than random sets of proteins in yeast (Ge et al., 2001). We investigated this by analysing the primary neuronal culture time-course gene expression data in relation to the MASC protein-protein interaction network.

This analysis revealed that there were two classes of proteins in this expression/interaction network. One group (~ 25 %) is highly expressed early in the primary culture experiment and their expression decreases with neuronal development. However, the majority of MASC genes in this network however, start with a low expression level but increase in a similar profile to high expression levels at the end of the time course. Although, this does not prove that co-expressed synaptic proteins are more likely to interact with one another but when compared to the time course analysis of PSP protein classes (Figure 5.11) it can be seen that the majority of the components of the MASC complex seem to be co-regulated. This apparent co-regulation of a set of highly interconnected proteins does seem to support the functional link between the transcriptome and the interactome. Models which exploit gene expression data to predict protein interactions are currently emerging and the inclusion of expression data for orthologs seems to improve the predictive capabilities of such models (Bhardwaj and Lu, 2005). A more detailed

analysis of MASC in this setting, in conjunction with such ortholog data would be interesting and potentially useful for completing interaction maps.

5.5.2.6.1 Network analysis of proteins implicated in Alzheimer's disease

The main approach to analyse proteome changes in caused by brain pathology is to use comparative 2D-PAGE separation of constituent proteins in diseased versus normal samples. This is usually followed by gel image analysis to identify alterations in the intensity of specific spots compared to the normal state and subsequent excision of the abnormal spots and identification by mass spectrometry or peptide sequencing. The number of comparative proteomic studies of disease and disease models has increased over the last number of years. This has led to the accumulation of lists of proteins whose expression levels have been shown to be altered in the particular system, but with little integration of these lists to provide a global view. This is emerging problem in the study of complex diseases, not only at the level of proteomics but also at the level of study of individual protein function in relation to disease.

The pathogenesis of neurodegenerative diseases such as AD is likely to consist of a number of initial pathways converging to form the distinctive pathology of amyloid plaques and NFT's (neurofibrillary tangles) characteristic of AD. Proteins implicated in this pathogenesis may also be characteristic of the stage of progression of the disease and may be transiently altered. The components and formation of amyloid plaques and NFT's are reasonably well understood but the many pathways leading to this end point are still elusive. It is clear that there is a need to integrate the large amount of data concerning proteins involved in the pathogenesis of AD into a coherent global view of the disease. It may be quite likely that an understanding of the fundamental components and mechanisms of the pathways

leading to AD pathology will be gained from a holistic view that can be obtained by using a proteomic strategy combined with bioinformatics and network analysis.

The large number of implicated proteins present in the MASC and AMPAR complex indicates that altered signalling from these membrane receptors may be responsible for the loss of synaptic function and cognitive decline associated with AD. Recently, its has been shown that application of amyloid-beta promotes endocytosis of NMDA receptors in cortical neurons, possibly explaining the reduced glutamatergic transmission and inhibition of synaptic plasticity observed in transgenic mice overexpressing amyloid precursor protein (APP). Elucidating the involvement of particular pathways in a systematic way should begin to uncover these various molecular routes to the diseased state. This approach of network integration of components of a complex disease should provide a coherent view of the disease process at the molecular level. Network analysis of AD as well as many other complex diseases might indicate new leads into the mechanisms of pathology and areas of therapeutic intervention.

5.5.3 Conclusions

We have performed comprehensive bioinformatic analysis of the PSP and found elements of organisation at every level of complexity. This analysis and annotation of the PSP has added substantial value to the dataset and has provided insights into how the PSP is organised. This PSP dataset is a candidate gene list for the Genes to Cognition research programme in which the function of a large number of these genes is studied by systematic targeted mutations in mice (Grant, 2003). Comprehensive phenotypic analysis of these mutants in terms of biochemistry, electrophysiology, cell biology and behaviour is allowing functional annotation of PSP components. This approach together with SNP (Single-Nucleotide Polymorphism) screening in humans for all genes in the PSP will produce systematic data relevant to synaptic physiology and disease. Establishment this PSP

dataset (available at www.genes2cognition.org/ science/proteomics/synapse.html) is essential for the progression of synapse proteomics and with new quantitative proteomic strategies which may be applied to activated or diseased states of the synapse, will also provide a benchmark for analysis of dynamic aspects of the synapse proteome.

CHAPTER 6

DISCUSSION AND CONCLUSIONS

6. Discussion and conclusions

The postsynaptic proteome, which encompasses the postsynaptic density and embedded multi-protein complexes, is the fundamental functional unit of neuronal cells. Many of its component proteins are involved in functions specific to the brain and dysregulation of these components characterise numerous neurological disorders, from metal retardation to neurodegeneration. Protein phosphorylation is an essential regulator of synaptic function and dysregulation of phosphorylation in signalling pathways has also been implicated in nervous system disorders. Therefore, global characterisation of the components and phosphorylation of the postsynaptic proteome will undoubtedly shed light on basic processes of synaptic function as well as synaptic dysfunction in disease.

The objective of this project was to develop and apply current proteomic and bioinformatic approaches to map and understand the molecular environment of central nervous system synapses.

6.1 – Method development: new tools to tackle the phosphoproteome

In 2002, at the same time as the work described in this thesis began, the first phosphoproteomic dataset was published. Ficarro and co-workers had modified the standard IMAC protocol, which suffered from non-specific binding of non-phosphorylated peptides, and introduced chemical methyl-esterification of peptide to neutralise acidic side chains (Ficarro et al., 2002). They applied this methodology to the yeast phosphoproteome and identified 383 phosphorylation sites. We reasoned that a 2-step enrichment of proteins and phosphoproteins would increase the likelihood of obtaining greater phosphoproteome coverage. We adapted the principle of peptide metal affinity chromatography and used it to target

phosphoproteins. A key element of our protein IMAC approach was to use high urea concentrations to denature synaptosomal proteins. This allowed inactivation of any residual phosphatase activity, protein denaturation that would allow greater protein solubility and more efficient binding to the IMAC resin.

Many phosphorylation sites were identified directly from this protein IMAC fraction using precursor ion scanning functions. However, there was significant gain in phosphoproteome coverage when this was combined with a peptide IMAC step. This tandem affinity isolation of phosphorylated species was superior to single protein or single peptide IMAC isolations. Comparable single peptide IMAC purifications from isolated PSD's only resulted in identification of approximately one-third of that observed from tandem-IMAC purifications (Trinidad et al., 2005), supporting the validity of our approach. Our optimised protocol has identified up to 401 phosphorylation events on 173 phosphopeptides in a single 4-hour LC-MS/MS analysis. This advance in technology compared to traditional approaches such as ^{32}P ATP labelling and Edman sequencing, has significantly decreased the time and the increased the sensitivity of phosphorylation site identification. Indeed, the advent of such approaches has made global phospho-profiling experiments feasible.

6.2 – Application of phosphoproteomic methods

6.2.2 – The synapse phosphoproteome

The complexity of the synapse in terms of its components and the processes it carries out necessitates a complex and robust regulatory system. Phosphorylation at the synapse has been studied for over 35 years and the identification of phosphorylation sites has been an arduous process. The study of synaptic phosphorylation on a protein-by-protein level was inherently biased with attention focused on channels and receptors, which were the earliest class of proteins

characterised at the synapse by pharmacological methods. Undoubtedly, phosphorylation of this class of proteins is important, but identification of phosphorylation sites in a global unbiased manner is necessary to expand our understanding of synaptic signalling beyond local regulation of channel function.

The aim of this study was to establish an *in vivo* map of the synapse phosphoproteome. We detected over 650 phosphorylation events corresponding to 331 sites (289 have been unambiguously assigned), 92% of which are novel. These represent 79 proteins, half of which are novel phosphoproteins. An additional 149 candidate phospho-proteins were identified by profiling the composition of the protein IMAC enrichment. This degree of novelty in terms of newly identified phosphoproteins as well as phosphorylation sites highlights a previously unrecognised complexity in synaptic phosphorylation.

To examine the significance of this complexity we examined the representation of phosphorylation data across synaptic proteins and important synaptic sub-components and organelles. Our dataset contained postsynaptic components including the postsynaptic density (134 proteins) and NMDA-receptor associated proteins (44 proteins) and presynaptic components, of which the most striking examples were Piccolo-Bassoon transport vesicles. A prominent feature of the synapse phosphoproteome was the appearance of many phosphorylated postsynaptic scaffold proteins. Major components of the NMDA receptor complex such as PSD-95, Chapsyn-110 and SynGAP were all phosphorylated indicating that the scaffolding functions of these proteins may be regulated by phosphorylation. This would suggest that the PSD scaffold, which provides a molecular matrix to support and connect synaptic membrane receptors with intercellular signalling pathways, is itself subject to regulation.

Would such a dynamic synaptic scaffold be able to regulate signalling proteins in a spatial manner? We investigated this by looking at the relationship between kinases

and scaffolding proteins in a network model of the NRC. Using *in vitro* kinase data from peptide arrays and protein interaction data, we observed that some kinases such as PKA, act on sets of interacting proteins. PKA was linked to four interacting substrates (PSD-95, PSD-93, SAPAP1, SAPAP2) via AKAP79, a kinase anchor protein that recruits PKA to the postsynaptic membrane in proximity to its physiological substrates. This novel approach of layering kinase-substrate data onto protein-protein interaction networks in the third-dimension provided an insight into how phosphorylation may regulate synaptic signalling by acting on the synaptic scaffold.

Another theme that emerged from the study of the synapse phosphoproteome was that of kinase multiplicity. Data derived from *in silico* predictions of cognate kinases for identified phosphorylation sites and *in vitro* phosphorylation of peptide arrays suggested that a small number of kinases can phosphorylate many proteins and each substrate is phosphorylated by many kinases. This phenomenon may represent signal integration, i.e. when signalling pathways converge and result in multi-site phosphorylation of substrate. Another possibility that is not mutually exclusive is that kinase multiplicity represents robustness in signalling networks. The existence of a number of variant signalling pathways involving different kinases/ phosphatases or substrates, which have the same downstream target (e.g. transcription factor activation), would ensure signal transduction robustness. Finally, the complexity of synaptic signalling may require subtle fine tuning, which may be mediated by combinations of different kinases working on the same substrates, resulting in slightly altered activity or localisation of the substrate.

Interestingly, we were able to characterise signalling components of the mGluR5-IP3 receptor-signalling pathway from a global analysis of synaptic phosphorylation. We found novel phosphorylation sites on well-characterised proteins such as mGluR5 and PLC-β, which are functionally and physically linked. In addition, we found potentially very interesting regulatory phosphorylation sites on IRBIT, a poorly

characterised signalling protein that binds the IP3 receptor in an activity-dependent manner. This complex seems to facilitate a signalling pathway from mGluR5 to modulation of intracellular calcium, which is known to regulate a myriad of intercellular signalling activities. Identification of these phosphorylation sites in proteins along the signalling cascade should provide an excellent starting point for the investigation of signal transduction in this pathway.

6.2.1 – The cytosolic phosphoproteome

We applied our optimised tandem metal-affinity purification approach to purify phosphorylated species from cytosolic extract from mouse forebrains. This cytosolic forebrain preparation was chosen to get an overview of phosphorylation of soluble brain phosphoproteins. Compared to the synapse phosphoproteome dataset that was derived from multiple analyses of single protein, single peptide and double IMAC experiments, the cytosolic dataset was the result of just two experiments. These two analyses focused on a tandem IMAC (double) purification from cytosolic extract with different on-line RP separations. This comparison identified that a PepMap column is best suited for LC-MS/MS analysis of phosphopeptides with a 47% increase in phosphorylation event identifications using the PepMap column compared to the BetaMax neutral column.

As water-soluble cytosolic proteins are hydrophilic we reasoned that they might contain intrinsically disordered sequence, which also tends to be hydrophilic. We found that the cytosolic phosphoproteome dataset was rich in sequence disorder and contains many putative natively disordered proteins. The majority of these had RNA or tubulin binding functions and the combination of disorder and phosphorylation is likely to be essential for their function. We identified 37 phosphorylation sites on nine protein kinases, the majority of which are novel and are not located in kinase domains indicating secondary regulatory functions. Finally, in our global study of cytosolic phosphorylation we have characterised a

number of phosphoproteins with disease-associations, which may have related roles in the pathogenesis of neurodegenerative disease.

6.3 The postsynaptic proteome

6.3.1 Profiling postsynaptic proteomes

The last five years has seen tremendous growth in knowledge of the protein components of mammalian synapses. The first investigation of the synapse proteome appeared in 2000, with the identification of the components of the NMDA receptor complex (77 proteins) (Husi et al., 2000). The large number of proteins in this complex was at first controversial, as this exceeded the total number of known PSD proteins. This report was soon followed by the first MS-based characterisation of the PSD, which revealed just 31 proteins in the PSD, most of which were already known. Our analysis of the PSD and protein complexes and that of other PSD proteomic studies has subsequently validated the components of the NRC. The data integration performed on these PSD and protein complex datasets has resulted in a confidence-ranked database of postsynaptic proteins which totals 1124. This 10-fold increase in known postsynaptic components reflects the technological advances of MS-based protein identification compared to traditional methods.

The most important and striking point to come out of theses analyses is that the PSP is incredibly complex. The vast amount numbers of proteins in the PSP presents a daunting challenge to systematically attribute synaptic function to each protein. However, for the first time we are in a position where we actually know the identity of the majority of proteins in the PSP and are able to ask questions concerning the entire postsynaptic proteome, as opposed to discrete subsets of proteins in the past. We are therefore well equipped to start taking a systems biologic approach to study synapse function. In fact, this PSP database is a candidate gene list for the Genes to Cognition research programme in which the function of a large number of these

genes is studied by systematic targeted mutations in mice (Grant, 2003). Comprehensive phenotypic analysis of these mutants in terms of biochemistry, electrophysiology, cell biology and behaviour is allowing functional annotation of PSP components. This approach together with SNP (Single-Nucleotide Polymorphism) screening in humans for all genes in the PSP will produce systematic data relevant to synaptic physiology and disease.

6.3.2 Molecular organisation of the postsynaptic proteome

The emerging field of systems biology has yet to be given a clear definition but at its most basic, it aims to integrate biological data to facilitate global representation and modelling of biological phenomena. In order to address some aspects of synaptic biology we used our database of postsynaptic proteins as a common platform onto which we layered and integrated multiple data types. This encompassed many levels of the hierarchy of cellular complexity from gene evolution and expression, protein classification and domain architecture, post-translation modifications and multi-protein complexes. Protein-protein interaction networks were used to integrate the proteomic data with gene expression, phosphorylation, phenotypic data in knockout mice and AD-protein association data. Networks are central to representation of systems biology data and although in this study we did not use graph-theory modelling, some interesting trends were observed.

We observed a theme of organisation at each level of complexity. At the genetic level this organisation was the acquisition of synaptic genes through evolution, a significant proportion of which exhibit neuronal specific expression as seen in the tissue transcriptome analysis. The integration of our PSP proteomic data with developmental gene expression and electrophysiological data from cultured neurons provided insights into how a synapse might be constructed as a neuron develops. The timing of expression of key components of the synapse correlates with the onset of electrical activity at around 10 days in culture. The concomitant

up-regulation of activity-dependent genes with electrical activity and the overall timing of expression of different protein classes are intuitive. This would indicate that despite complete correlation of mRNA with protein expression, that the temporal model of assembly is plausible. Such a model could be verified using quantitative proteomic analysis of the same cultured neurons. The proteome coverage would not be as complete as that obtainable for the transcriptome but enough protein expression data for correlation analysis would be achievable.

Integration of phenotypic data with PSP proteomic data and representation in protein interaction networks provides an interesting way to try to connect a very specific electrophysiological phenotype in observed in a number of knockout mice. There seems to be a relationship between phenotype and connectivity of the gene knocked out. In systematic deletion studies in yeast, the most severe phenotypes were observed for deleted proteins that had many protein interaction partners (Jeong et al., 2001). Currently, neither the protein interaction data nor systematic mutations in mouse are available for the MASC to comprehensively test this model. However, both of these issues are being addressed and it is expected that within the next five years that sufficient data will be available to explore these issues comprehensively.

Complex disorders such as AD have implicated many genes/proteins in their causation or pathology. It is likely that a number of cellular pathways are involved in this disease process and integration of data on AD-implicated proteins and their connections in a protein interaction network may allow identification of common pathways. In addition, such an approach is useful for evaluation of animal models of disease. We used protein expression change data observed in GSK3β overexpressing mice and overlaid it onto an AD-associated protein interaction network. Many proteins with altered expression in the mouse model were present in the AD interaction network, indicating that the pathway disrupted in the mouse model is involved in at least some of the pathology. Such analysis of mouse models

in a systematic manner should highlight perturbations in pathways responsible for the disease.

6.4 Neuroproteomics perspective

Now that the field of proteomics is coming to the end of its developmental phase, the application of proteomic methodologies to biological problems should provide even more leaps of knowledge in the next five years. Analysis of the phosphoproteome is not currently routine, however phosphoproteomic methodologies are maturing. It is anticipated that full phosphoproteome coverage is not achievable from analyses of "basal states" in tissue because some phosphorylation sites are of such low abundance in the absence of a specific stimulation. Therefore, the real challenge is to move from a discovery phase of phosphorylation site identification to a profiling phase of phosphorylation in response to stimulation of cells or tissues with drugs and inhibitors etc. Similarly, the recent advent of quantitative mass spectrometry will facilitate profiling of dynamic proteome states. The use of bioinformatics in the analysis of this data will be of the utmost importance and is an area that will need much attention if the value of such data is to be fully appreciated. One can envisage that the application of these new approaches using well defined chemical (NMDA) or electrical stimulation paradigms (LTP) in cultured neurons or tissue slices, will undoubted reveal novel and interesting aspects of the complexity of synaptic plasticity

6.5 Conclusions

The work detailed in this thesis has for the first time, described the constituents and phosphorylation of multiple brain proteomes. 652 *in vivo* phosphorylation sites were identified directly from brain purifications, more than that identified by traditional approaches in the last 35 years of study of in the field of neuronal signalling. For the first time a comprehensive database of the components of the postsynaptic

proteome was assembled and systematic molecular organisation of the PSP was revealed. Establishment of such maps of proteins and phosphorylation at the synapse will provide a basis for future research of synaptic physiology and disease.

INDEX OF FIGURES

INDEX OF TABLES

Details of Appendices 1 – 12

Appendix 1 Synaptic proteins identified in IMAC purifications identified by mass spectrometry.

Appendix 2 Unique phosphopeptide sequences detected in multiple IMAC purifications from synaptosomes by LC-MS/MS analysis.

Appendix 3 Non-redundant phosphopeptide list from the synapse phospho-proteome with unique phosphorylation site assignation.

Appendix 4 Unique phosphopeptide sequences detected in double IMAC purifications from forebrain cytosolic fraction by LC-MS/MS analysis.

Appendix 5 Non-redundant phosphopeptide list with unique phosphorylation site assignation for the cytosolic phosphoproteome.

Appendix 6 Overview of G2C PSP datasets.

Appendix 7 Integration of PSP datasets.

Appendix 8 Summary of integrated PSD and multiprotein complex datasets.

Appendix 9 **Collins, MO**, Husi, H, and Grant, SGN. (2004) Proteomics in the Neurosciences. In Proteomics: Biomedical and Pharmaceutical Applications. Edited by Hondermarck, H. (Kluwer Academic Press).

Appendix 10 **Collins MO**, Yu L, Husi H, Blackstock WP, Choudhary JS, Grant SG. (2005) Robust enrichment of phosphorylated species in complex mixtures by sequential protein and peptide metal-affinity chromatography and analysis by tandem mass spectrometry. Science STKE. 2005(298):pl6.

Appendix 11 **Collins MO**, Yu L, Coba MP, Husi H, Campuzano I, Blackstock WP, Choudhary JS, Grant SG. (2005) Proteomic analysis of in vivo phosphorylated synaptic proteins. Journal of Biological Chemistry. 280(7):5972-82.

Appendix 12 **Collins, MO**, Husi , H, Yu, L, Brandon, JM, Anderson, CNG, Blackstock, WP, Choudhary, JS, Grant, SGN. (2005) Molecular characterisation and comparison of the components and multi-protein complexes in the postsynaptic proteome. Journal of Neurochemistry. (In press)

Appendix 1

No	Accession	Mr (kD)	Rel Abun Rank	Protein Name	PSD Proteins	NRC Proteins	Protein Identification Peptides Detected	Whole Synaptosomes Phosphopeptides Detected	Urea soluble Synaptosomes Phosphopeptides Detected	Urea insoluble Synaptosomes Phosphopeptides Detected	Known Phosphoproteins Reference #	Known cognate kinases (literature)	Scansite phosphorylation hits	TTM # of helices
	Channels/receptors													
1	Q9R196	132		Calcium-activated potassium channel alpha subunit 1 (BK channel)						2	1	PKA, PKC	c, b, f, g	6
2	Q9EPT9	58	67	L-type calcium channel beta 1A subunit	+		4				2	PKA	t, i	
3	Q8R3Z5	65	73	L-type calcium channel beta 1B subunit			4				2	PKA	c, t, i, w	
4	Q8BJY3	132		Metabotrophic glutamate receptor, mGluR5	+	+		1		2	3	PKC	e	6
5	Q64284	41	123	Voltage gated potassium channel, beta 2 subunit	+		1		1				g, n	
6	Q8R0S4	55		Voltage-dependent L-type calcium channel beta-4 subunit					1				t, d	
7	P70207	210		Plexin 2	+					1			a, s, u, v, j	1
	Transporters/transporting ATPases													
8	P56564	60	148	Excitatory amino acid transporter 1			1				4	PKC	d, c, s	10
9	P43006	62	149	Excitatory amino acid transporter 2	+	+	1				4	PKC	u	9
10	Q9QW01	18	35	Glutamate/aspartate transporter, GLU-T			2							2
11	P53986	53		Monocarboxylate transporter 1 (SLC16A1)						1			t	11
12	Q91YY9	106	40	Na+/K+ transporting ATPase alpha-2 chain			11				5	PKA, PKC	a, q, j	8
13	Q8CGD9	112	61	Na+/K+ transporting ATPase alpha-3 chain	+		8				5	PKA, PKC	q	5
14	P14094	35	33	Na+/K+ transporting ATPase beta-1 chain	+		4						m	1
15	Q9R0K7	133	153	PM calcium-transporting ATPase 2	+	+	2				6	PKA	i, f, g, o	8
16	Q91YQ9	74		Solute carrier family 20-member 1						1			s, u	8
17	Q9EPR4	70		Solute carrier family 23, member 2				1					t	10
18	Q8BUN9	74		Solute carrier family 24						1			s, v, b, i	11
19	Q9Z1G4	97	168	Vacuolar proton pump subunit 1	+		1						s	6
	Scaffolding and adaptors													
20	Q92319	49	74	Abl interactor 1			3				7		a, j	
21	O88881	67	154	Begain-1	+		1						q	
22	Q8K3M6	110	6	CAST	+		53						v	
23	P70587	167	176	Densin-180	+		1				8	CaMKII	a, c, t, u	
24	Q8IZQ1	110	102	DISKS large-associated protein 1 (DAP1)	+	+	4				9	CaMKII	q, i	
25	Q8IJ42	122	122	DISKS large-associated protein 2 (DAP2)		+	3		2				b, i, f, g	
26	AAO89219	106	11	DISKS large-associated protein 3 (DAP3)		+	39		6				d, c, q	
27	P97839	108	110	DISKS large-associated protein 4 (DAP4)	+	+	3						d, l, a, q, s, b, i	
28	O08599	68	3	Munc-18-1 (Syntaxin binding protein 1)	+	+	48		7		10, 11	Cdk5, PKC	q, u	
29	O35413	134	32	nArgBP2	+		16		1				c, q, s, v	
30	Q923C6	83	44	N-ethylmaleimide sensitive factor	+	+	8				12	PKC		
31	Q9QYX7	548	111	Piccolo protein	+		15	1	18	4			d, e, l, a, q, s, t, u, b, f, g, j, m, w	
32	Q91XM9	110	19	PSD-93/Chapsyn-110	+	+	23	1			13	Fyn	c, s, b, f, g	
33	Q62108	80		PSD-95	+	+			1	1	14	Cdk5	h, f, g	
34	O35168	180	37	RIM-1	+		19		2	1	15	CaMKII	d, a, c, h, q, s, t, i, f, g	
35	Q9JIR4	180	177	RIM-1B			1				15	CaMKII	d, a, c, h, q, s, t, i, f, g	
36	Q9EQZ7	173	165	RIM-2	+		2				15	CaMKII	d, a, c, h, q, s, t, f, g	
37	Q9WV48	226	15	Shank1	+	+	53		1				d, l, a, c, q, v, b, f, g, j, o	
38	Q9UPX8	159	66	Shank2	+	+	11						l, a, c, q, f, j	
39	O35526	33	30	Syntaxin 1A	+		4				16	CK1, CK2	v	1
40	Q9ERQ9	62	150	WAVE-1	+		1				17	MAPK	c, j	
	Kinases													
41	Q63092	54	142	Calmodulin-binding protein	+		1							
42	P11798	54	7	CaM kinase type II alpha chain	+	+	26	1			18	Autophos	d, e	
43	P28652	60	10	CaM kinase type II beta chain	+	+	24		1		18	Autophos	c, u	
44	Q9CZE2	54	101	CaM kinase type II delta chain	+		2				18	Autophos	c, v	
45	Q923T9	60	69	CaM kinase type II gamma chain	+		4				18	Autophos	c, v	
46	P49025	183	128	Citron	+	+	4						c, q, t, b, f, g	
47	Q04447	43	46	Creatine kinase, B chain	+		4				19			
48	Q9Z188	65		Dual-specificity tyrosine-phosphorylation regulated kinase 1B					1		20		q, c, d, j	
49	Q9WV60	47	93	Glycogen synthase kinase-3 beta	+	+	2				21, 22, 23	p70 D6 kinase, Rsk-1, PKC	a	
50	P70336	161	175	p160 ROCK-2	+		1				24	Autophos	c, t, b	
51	O88377	47	94	Phosphatidylinositol 5-phosphate 4-kinase beta	+		2				25	EGFRK	m	
52	P10644	43		PKA type I-alpha regulatory chain				1			26		r, v	
53	Q8K1M3	46		PKA type II-alpha regulatory chain	+					1	27		q, d	
54	P31324	46	131	PKA type II-beta regulatory chain	+	+	1				28	Autophos	q, d	
55	O88664	116	172	Ser/Thr protein kinase TAO1			1				29	MEK3	g	
56	Q9JLM8	84	52	Serine/threonine-protein kinase DCAMKL1	+		7				30	Autophos	l, c, q, u, o	
57	Q9NQU5	75	113	Serine/threonine-protein kinase PAK 6			2				31	Autophos	a, f, g, o	
58	Q9UKE5	154	57	TRAF2 and NCK interacting kinase (TNIK)			12				32		e, q	
59	O75962	342	146	Triple functional domain protein, Trio	+		6				33, 34	FAK	l, s, v, b, i, j, m	

Appendix 1

No	Accession	Mr (kD)	Ref Abun Rank	Protein Name	PSD Proteins	NRC Proteins	Protein Identification Peptides Detected	Whole Synaptosomes Phosphopeptides Detected	Urea soluble Synaptosomes Phosphopeptides Detected	Urea insoluble Synaptosomes Phosphopeptides Detected	Known Phosphoproteins Reference #	Known cognate kinases (literature)	Scansite phosphorylation hits	TTM # of helices
Phosphatases														
60	P20652	59	68	PP2B (Calcineurin) catalytic subunit, alpha isoform	+	+	4				35	PKC, CaMKII	g	
61	Q9EQ17	212	179	Protein-tyrosine phosphatase LAR	+		1				36	TrkA	a, h, i, f, g, j	1
62	Q13332	217	180	Protein-tyrosine phosphatase sigma	+		1						a, c, t, u, h, j	2
63	O60228	190	167	Sbf-1 dual-specificty phosphatase	+		2						d, e, l, a, x, f	
Small G-proteins and modulators														
64	Q15673	70	157	EFA6	+		1						l, h, f, g	
65	BAC65687	60	27	mKIAA0942 protein (EFA6R)			8		23	2			t, f	
66	Q61018	40	85	Guanine nucleotide-binding protein G(I), alpha-1 subunit			2						d	
67	P04901	37	112	Guanine nucleotide-binding protein G(I)/G(S)/G(T) beta subunit 1	+		1							
68	P18872	40	26	Guanine nucleotide-binding protein G(O), alpha subunit 1	+		6							
69	P18873	40	24	Guanine nucleotide-binding protein G(O), alpha subunit 2	+		7						d	
70	P21279	41	60	Guanine nucleotide-binding protein G(q), alpha subunit	+		3							
71	Q8IH66	63		Guanine nucleotide-binding protein 3 (Alastin)						2			c, v, i	2
72	O94863	95	133	KIAA0763 PROTEIN	+		2						d, n, f, g	
73	Q9Y2I9	71	51	NRip11	+		6						a, c, f, g	
74	Q8R361	70	50	Rab11 interacting protein Rip11			6		2		37		d, l, q	
75	P47708	75	160	Rabphilin-3A	+		1				38	PKA, CaMKII	v, b, i	
76	NP_796353	190	178	RICS/Rho GTPase-activating protein	+		1				39	CaMKII	a, c, s, n	
77	Q9D677	79	118	Regulator of G-protein signaling 12			2						l, s	
78	O54829	55	145	Regulator of G-protein signaling 7	+		1				40			
79	O54828	77	161	Regulator of G-protein signaling 9			1				41	PKA	a, q, u, w	
80	Q9ES28	80		Rho guanine nucleotide exchange factor 7 (beta-PIX)	+				1		42	ERK/PAK2	q, t, l, g, f, j	
81	O55131	51	56	Septin 7	+		4						c, s	
Other signaling molecules														
82	P11576	28	104	14-3-3 protein eta	+	+	1				43	SDK1/PKC delta		
83	P51830	151		Adenylate cyclase, type IX						1	44	PKC	s, t, g	11
84	P02593	17	79	Calmodulin	+	+	1				45, 45	CK2, INSRK, EGFRK, Src		
85	P57716	78	162	Nicastrin			1						c	1
86	Q9QUH6	143	12	SynGAP	+	+	52		1		47, 48	CaMKII	a, c, q, r, x, h, o	
87	Q80SW1	59		IP3R binding protein released with inositol 1,4,5-triphosphate						1	49		d, e	
88	Q8VDM7	73	99	Phosphoinositol 3-phosphate-binding protein-2			3						e, q, t, v	
89	Q9Z1B3	138	83	Phospholipase C-beta-1	+	+	7		1		50	ERK1/2	d, e, c, v	
90	P13795	23	48	Synaptosomal-associated protein 25	+	+	2				51	PKA		
91	P80316	51	138	T-complex protein 1, epsilon subunit	+		1						b, i	
Other enzymes														
92	Q9DCX2	19	38	ATP synthase D chain, mitochondrial			2						c	
93	Q9ERA8	33	14	ATP synthase gamma-subunit	+	+	9						d, t	
94	P56818	70	62	Beta-secretase			5				52	CK2	q	1
95	Q9DAM4	52	100	Citrate Synthase	+		2						a, u	
96	P17742	18	36	Cyclophilin A			2						t	
97	P00009	11	13	Cytochrome c, somatic			4							
98	P05064	39	115	Fructose-biphosphate aldolase A	+	+	1				53	Autophos	d, a, c	
99	P06745	63	151	Glucose-6-phosphate isomerase			1				54	CK2	i	
100	P16858	36	80	Glyceraldehyde 3-phosphate dehydrogenase	+	+	2				55	PKC iota/lambda		
101	P70404	43		Isocitrate dehydrogenase [NAD] subunit gamma	+				1					
102	P17710	102	78	Hexokinase, type I	+		6						r	
103	Q8C1D1	40	119	Hypothetical acyl-CoA N-acyltransferase			1						a, h	
104	Q99JP9	39	116	NADH-ubiquinone oxidoreductase	+	+	1						s	
105	P17183	47	132	Neural enolase	+		1				56			
106	P35486	43		Pyruvate dehydrogenase E1 component alpha subunit	+				1		57	PDK		
107	Q8C5C0	64	152	SAM domain and HD domain protein			1						l	
108	Q91WP8	51	139	Similar to mitochondrial phosphoprotein, MIPP65			1						c, l, u, i	
109	Q9WUM5	35	34	Succinyl-CoA ligase, alpha-chain	+		4				58		s	
110	P51863	40	120	Vacuolar ATP synthase subunit D	+	+	1							
111	P50518	27	59	Vacuolar ATP synthase subunit E	+		2							
Cell adhesion and cytoskeletal														
112	NP_878259	24	98	Brain-specific protein p25 alpha	+		1		3		59	PKC, TPK II	q	
113	Q8BFZ3	42	125	Actin	+		1							
114	P02571	42	63	Actin, cytoplasmic 2	+	+	3							
115	Q8K4G5	78	58	Actin-binding LIM protein 1 long isoform	+		6		2		60		q, f, g, j	
116	P26231	100		Alpha-1 catenin				1			61			
117	P97582	90	91	Ankyrin	+		4							
118	Q02357	204	107	Ankyrin 1	+		6				62	CK2	q, r, t, u, h, i	
119	Q01484	430	28	Ankyrin 2	+		56		6		63		l, a, c, q, h, i, f, g, o	
120	Q61307	212	39	Ankyrin 3 epithelial	+		22				62	CK2	a, q, s, v, o	
121	O70511	270	182	Ankyrin 3 G isoform	+		1						l, a, h, q, s, t, v, i, f, g, j	
122	Q9Z2H5	98	169	Band 4.1-like protein 1	+		1	1	2	1			b, f, g,	
123	O70318	110	144	Band 4.1-like protein 2			2				59		q, b, i, f, g	

Appendix 1

No	Accession	Mr (kD)	Rel Abun Rank	Protein Name	PSD Proteins	NRC Proteins	Protein Identification Peptides Detected	Whole Synaptosomes Phosphopeptides Detected	Urea soluble Synaptosomes Phosphopeptides Detected	Urea insoluble Synaptosomes Phosphopeptides Detected	Known Phosphoproteins Reference #	Known cognate kinases (literature)	Scansite phosphorylation hits	TTM # of helices
124	Q9WV92	103	140	Band 4.1-like protein 3	+		2	4	3				v, b, i, g	
125	O88737	419	20	Bassoon	+	+	87	8	27	4			d, e, l, a, c, h, q, b, i, f, g, j, o, m	
126	Q02248	85	164	Beta-catenin	+	+	1				64, 65, 66	Src, GSK3, CK1	f, c, q, i, o	
127	Q91ZE6	288	183	Beta-spectrin 4	+		1						a, c, q, b, o	
128	P35564	65		Calnexin precursor	+			1			67		f	1
129	Q8IWR2	50	53	Class II beta tubulin isotype	+		4						v	
130	P17426	108		Adapter-related protein complex 2 alpha 1 subunit	+			2	1	7	68		f, c	
131	Q6QR53	119		Adapter-related protein complex 3 beta 3					1		69	CK1	u, f, s, r, o,	
132	P17427	104	170	Clathrin coat assembly protein 2 alpha 2 subunit	+		1						f, c	
133	Q61548	92	129	Clathrin coat assembly protein AP180	+		2				70	CK2	f	
134	P20172	17	9	Clathrin coat assembly protein AP50	+		7				71	AP50 kinase		
135	BAC65475	192	134	Clathrin heavy chain 1	+	+	4				72	Src		
136	Q9Z0H8	115	114	CLIP-115	+		3						a, q, v, b, f, g	
137	Q99JD4	141	173	CLIP-associating protein CLASP2	+		1						a, b, q, i, b, i, f, g	
138	P12960	113	92	Contactin-1	+		5						f, s	
139	O35927	135		Delta-catenin	+				1	1	73	Src family kinases	u, w	
140	Q9WV69	45		Dematin					2		74	PKA	a, s, b, e, f, j	
141	Q9QXS6	72		Drebrin	+	+			1		75		s, t, e, i, d, f, j	
142	P60904	22		DnaJ homolog subfamily C member 5 (Cysteine string protein)					1				c	
143	Q9JI.04	166		Formin 2					1		76		n, v, i, b, d, f, g, f	
144	P08109	71	2	Heat shock cognate 71 kDa protein	+	+	51						c, p	
145	Q61735	35	108	Integrin-associated protein			1							5
146	P28738	109	143	Kinasin heavy chain	+		2				77	PKC, PKA	b, i	
147	O55129	96	45	Microtubule-associated protein (Stop)	+		9				78	CaMKII	d, f	
148	XP_194040	113	72	Microtubule-associated protein 1A	+		7		12	1	79	MAP1A kinase	t, b, i	
149	P14873	270	166	Microtubule-associated protein 1B	+		5	7	24	2	80, 81, 82	CK2, GSK3 beta, JNK1	e, l, c, s, t, i, f, g, j	
150	P20357	199	8	Microtubule-associated protein 2	+	+	89		7		83, 84, 82	EGFRK, INSRK, JNK1	l, a, c, q, t, u, f, g, j, o, w	
151	P27546	118	22	Microtubule-associated protein 4	+		22		1		85	Cdc2	l, r, s, b, f, g	
152	P10637	76	4	Microtubule-associated protein Tau	+	+	49	3	28	3	86	GSK3a, b, PKA, PKC, CaMKII	l, a, b, i, f, g, o	
153	P04370	27	1	Myelin basic protein	+		115		5		87, 88, 89	PKA, PKC, CaMKII, ERK2	c	
154	P02688	21	87	Myelin basic protein S			1		1					
155	Q9JMH9	231	171	Myosin containing PDZ domain	+		2						e, a, q, s, u, v, b	
156	Q62812	226	181	Myosin heavy chain, non muscle type A	+		1				90, 91	PKC, CK2	t, u, v, b, i, j	
157	Q61879	229	147	Myosin heavy chain, non muscle type B	+	+	4				90, 91	PKC, CK2		
158	Q99104	216	82	Myosin Va	+		11				92	CaMKII	a, q, t	
159	P13595	119		Neural cell adhesion molecule 1	+				1				p, h, d, l, g, f	2
160	Q810U3	138		Neurofascin precursor	+					2	93		d, a, b	2
161	P08553	96		Neurofilament triplet M protein	+			1			94	GSK-3 alpha	b, c, e	
162	Q99569	134		Plakophilin 4	+			1		3			f, q, t, b, i, g, p	
163	Q920P4	42		Paralemmin	+					2	95		i, f, g, o	
164	Q8K4S4	84		Reticulon 1A	+	+			2	1	96		d, l, b, i, j	2
165	UPI0000193BD1	425		Similar to ankyrin 2, neuronal				2	4	5			d, l, a, c, q, s, b, i, f, g, o	
166	XP_234989	45	127	Similar to Ccn3 protein			1						c, f, g	
167	XP_212807	71		Similar to Heat shock cognate 71 kDa protein					1					
168	Q8CFP5	91	31	Similar to Microtubule-associated protein 4			11						l, r, s, f, g	
169	Q81139	67	155	Similar to myosin heavy chain IX			1						t	
170	Q9QWI6	131	89	SNAP-25-interacting protein	+		6		1				l, q, f, g, j	
171	P16546	285	174	Spectrin alpha chain	+	+	2				97	Src	m	
172	Q62261	274	29	Spectrin beta chain	+	+	34		3		98	PKA	l	
173	XP_129018	71	109	Spir-1 protein			2			1	99	JNK	d, a, q, s, t, u	
174	Q64332	63	55	Synapsin II	+		5				100	CaMKII	q	
175	Q91WB8	33	75	Synaptophysin	+		2				101	pp60src		4
176	P46096	47	95	Synaptotagmin I	+	+	2				102	PKC, CaMKII, CK2	r, s, t, v	1
177	P05209	50	76	Tubulin alpha-1 chain	+	+	3						s, t	
178	P05213	50	77	Tubulin alpha-2 chain	+	+	3				103	CaMKII	s, t	
179	Q922F4	50	42	Tubulin beta	+		5				103, 104	CaMKII, CK2	v	
180	P05217	50	54	Tubulin beta-2 chain	+		4						v	
	Nuclear/transcriptional													
181	P22752	14	5	Histone H2A.1			7				105			
182	P17317	13	25	Histone H2A.z			2				106			
183	Q64477	15	21	Histone H2B			3				107	PKC	a, v	
184	P32067	48	71	Lupus La protein homolog			3						s, t	
185	Q8CE21	74	159	Nuclear and coiled-body phosphoprotein			1				108	CK2	d, e, s, v, f, g	
186	Q8VDJ6	44	90	PTB-associated splicing factor			2				109, 110	PKC alpha	l, t, v, j	
187	Q8R4E6	40	121	Pur-gamma A-form			1						l	
188	Q8C1H9	28	17	Splicing factor SFRS1			6						l, a, h, f, g	
189	Q9D0B0	28	23	Splicing factor, arginine/serine-rich 9 (Putative)			5				111		a, h, p	
190	Q9EQN4	108	141	TSC22-related inducible leucine zipper 1b			2						d, e, b, f, o	
	Translational													
191	P10126	50	135	Elongation factor 1-alpha 1	+		1				112, 113	PKC, Rho Kinase	a, s	
192	P27706	50	136	Elongation factor 1-alpha 2	+		1				112, 113	PKC, Rho Kinase	a, s	

Appendix 1

No	Accession	Mr (kD)	Rel Abun Rank	Protein Name	PSD Proteins	NRC Proteins	Protein Identification Peptides Detected	Whole Synaptosomes Phosphopeptides Detected	Urea soluble Synaptosomes Phosphopeptides Detected	Urea insoluble Synaptosomes Phosphopeptides Detected	Known Phosphoproteins Reference #	Known cognate kinases (literature)	Scansite phosphorylation hits	TTM # of helices
193	P16254	14	65	Signal recognition particle 14 kDa protein			1						d	
	Novel/uncharacterised proteins													
194	Q8BKY9	61		1110030H18Rik				1					m, q, u, v	12
195	Q8CIB7	37		5730405D09Rik protein				1					a, g, q	
196	Q8C026	73	158	4931426N11RIK protein			1						l, c, q, o	
197	Q9D3Z0	20	86	4933426L22RIK protein			1						h, q	
198	Q8R502	92	130	AD158			2			2			l, c, u	4
199	Q8K070	45		CDNA sequence BC034054					1				h, h, r, i, l, g, f	
200	Q8C145	84	124	Estrogen regulated LIV-1 protein			2						u, i	6
201	Q8BXF0	56		G2 PROTEIN	+					2			a, u, v	1
202	Q8BKM7	29	41	Hypothetical histidine-rich region containing protein			3							
203	Q8CCL6	68	156	Hypothetical protein			1						d, c, a, f, g, j, k, m, w	
204	Q8C3Q5	39	16	Hypothetical protein (D430041B17)			9		3				l, a, i, g	
205	Q9H3X1	23	49	Hypothetical protein			2							
206	XP_303766	84	103	Hypothetical protein			3						d, a, h, q	
207	Q8C3E1	22	47	Hypothetical protein 1700054N08RIK			2						g	
208	Q9CPT5	21	88	Hypothetical protein D13WSU177E			1							
209	Q7L0D6	32		Hypothetical protein KIAA0773				1					h, l	
210	XP_283703	53		Similar to SNIP-a protein						1			l, j, o	
211	Q9H6K5	60	106	Hypothetical protein FLJ22184			2						l, j, o	
212	O60275	170	81	Hypothetical protein KIAA0522	+		9						d, l, a, q, h, i, o, k	
213	Q8R0S2	62	18	KIAA0522 PROTEIN			13						a	
214	O75133	152	137	KIAA0638 PROTEIN			3						l, a, s, i, f, j	
215	O94930	47	96	KIAA0844 PROTEIN			2							
216	Q9P2M9	50	43	KIAA1317 PROTEIN			5						q	
217	Q9C0J0	78	163	KIAA1673 PROTEIN			1						d, l, a, o	
218	Q80WM4	43	126	link protein 4 precursor			1				114		s	1
219	Q8CHA8	115	105	KIAA1217 PROTEIN			4						l, a, f, j, o	
220	Q8BX42	42	64	RIKEN CDNA 2900042E17 GENE			3							5
221	Q8VD37	86		RIKEN cDNA 3110007P09	+				11				e, l, u, f, g, j, k	
222	Q8CID5	89		RIKEN cDNA 5730419I09				1					c	
223	Q8C0E9	44		RIKEN full-length enriched library, clone:6030438I23					1				t, i, h, g, f	
224	Q8CIL3	93	70	SIMILAR TO CDNA SEQUENCE BC019977			6						l, a, q, f, g, j	
225	Q8CCP1	60	84	SIMILAR TO ELKS			3							
226	XP_220918	128		similar to KIAA1501 protein					1				l, a, q, h, i, f, g	
227	Q8BYP7	39	117	Weakly similar to CDNA: FLJ21159 FIS (Fragment)			1						q, t, i, g	
228	Q8BWS5	80		Weakly similar to interacting protein 6				1		2			c, s, f, g, o	

Known phosphoprotein references

1 Tian, L., Duncan, R. R., Hammond, M. S., Coghill, L. S., Wen, H., Rusinova, R., Clark, A. G., Levitan, I. B. & Shipston, M. J. (2001) *J Biol Chem* **276,** 7717-20.
2 Chien, A. J. & Hosey, M. M. (1998) *J Bioenerg Biomembr* **30,** 377-86.
3 Alagarsamy, S., Rouse, S. T., Junge, C., Hubert, G. W., Gutman, D., Smith, Y. & Conn, P. J. (2002) *Pharmacol Biochem Behav* **73,** 299-306.
4 Casado, M., Bendahan, A., Zafra, F., Danbolt, N. C., Aragon, C., Gimenez, C. & Kanner, B. I. (1993) *J Biol Chem* **258,** 27313-7.
5 Mahmmoud, Y. A. & Cornelius, F. (2002) *Biophys J* **82,** 1907-19.
6 James, P. H., Pruschy, M., Vorherr, T. E., Penniston, J. T. & Carafoli, E. (1989) *Biochemistry* **28,** 4253-8.
7 Courtney, K. D., Grove, M., Vandongen, H., Vandongen, A., LaMantia, A. S. & Pendergast, A. M. (2000) *Mol Cell Neurosci* **16,** 244-57.
8 Walikonis, R. S., Oguni, A., Khorosheva, E. M., Jeng, C. J., Asuncion, F. J. & Kennedy, M. B. (2001) *J Neurosci* **21,** 423-33.
9 Yoshimura Y, Shinkawa T, Taoka M, Kobayashi K, Isobe T, Yamauchi T. (2002) *Biochem Biophys Res Commun* **250,** 948-54
10 Fletcher, A. I., Shuang, R., Giovannucci, D. R., Zhang, L., Bittner, M. A. & Stuenkel, E. L. (1999) *J Biol Chem* **274,** 4027-35.
11 Fujita, Y., Sasaki, T., Fukui, K., Kotani, H., Kimura, T., Hata, Y., Sudhof, T. C., Scheller, R. H. & Takai, Y. (1996) *J Biol Chem* **271,** 7265-8.
12 Matveeva, E. A., Whiteheart, S. W., Vanaman, T. C. & Slevin, J. T. (2001) *J Biol Chem* **276,** 12174-81.
13 Nada S, Shima T, Yanai H, Husi H, Grant SG, Okada M, Akiyama T. (2003) *J Biol Chem.* **278,** 47610-21.
14 Morabito MA, Sheng M, Tsai LH. (2004) *J Neurosci.* **24,** 865-76.
15 Sun, L., Bittner, M. A. & Holz, R. W. (2003) *J Biol Chem* **278,** 38301-9.
16 Dubois, T., Kerai, P., Learmonth, M., Cronshaw, A. & Aitken, A. (2002) *Eur J Biochem* **269,** 909-14.
17 Miki, H., Fukuda, M., Nishida, E. & Takenawa, T. (1999) *J Biol Chem* **274,** 27605-9.
18 Colbran, R. J. (1993) *J Biol Chem* **268,** 7163-70.
19 Mahadevan, L. C., Whatley, S. A., Leung, T. K. & Lim, L. (1984) *Biochem J* **222,** 139-44.
20 Becker W, Joost HG. (1999) *Prog Nucleic Acid Res Mol Biol* **62,** 1-17.
21 Goode, N., Hughes, K., Woodgett, J. R. & Parker, P. J. (1992) *J Biol Chem* **267,** 16878-82.
22 Hughes, K., Nikolakaki, E., Plyte, S. E., Totty, N. F. & Woodgett, J. R. (1993) *Embo J* **12,** 803-8.
23 Stambolic, V. & Woodgett, J. R. (1994) *Biochem J* **303,** 701-4.
24 Singh, U. S., Kunar, M. T., Kao, Y. L. & Baker, K. M. (2001) *Embo J* **20,** 2413-23.
25 Itoh, T., Ijuin, T. & Takenawa, T. (1998) *J Biol Chem* **273,** 20292-9.
26 Boeshans KM, Resing KA, Hunt JB, Ahn NG, Shabb JB. (1999) *Protein Sci.* **7.** 1515-22.
27 Tournier S, Gerbaud P, Anderson WB, Lohmann SM, Evain-Brion D, Raynaud F. (1995) *J Cell Biochem.* **57.** 647-54.
28 Budillon, A., Cereseto, A., Kondrashin, A., Nesterova, M., Merlo, G., Clair, T. & Cho-Chung, Y. S. (1995) *Proc Natl Acad Sci U S A* **92,** 10634-8.
29 Hutchison, M., Berman, K. S. & Cobb, M. H. (1998) *J Biol Chem* **273,** 28625-32.
30 Lin, P. T., Gleeson, J. G., Corbo, J. C., Flanagan, L. & Walsh, C. A. (2000) *J Neurosci* **20,** 9152-61.
31 Dan, C., Nath, N., Liberto, M. & Minden, A. (2002) *Mol Cell Biol* **22,** 567-77.

Scansite Kinase abbreviations

a Akt Kinase
b ATM Kinase
c Calmodulin dependent Kinase 2
d Casein Kinase 1
e Casein Kinase 2
f Cdc2 Kinase
g Cdk5 Kinase
h Clk2 Kinase
l DNA PK
k EGFR Kinase
j Erk1 Kirase
l GSK3 Knase
m Insulin Receptor Kinase
n Itk Kinase
o p38 MA?K
p PDGFR Kinase
q Protein Kinase A
r PKC alpha/beta/gamma
s PKC delta
t PKC epsilon
u PKC mu
v PKC zeta
w Src kinase

32 Brill LM, Salomon AR, Ficarro SB, Mukherji M, Stettler-Gill M, Peters EC. (2004) *Anal Chem* **76**, 2763-72
33 Debant, A., Serra-Pages, C., Seipel, K., O'Brien, S., Tang, M., Park, S. H. & Streuli, M. (1996) *Proc Natl Acad Sci U S A* **93,** 5466-71.
34 Medley, Q. G., Buchbinder, E. G., Tachibana, K., Ngo, H., Serra-Pages, C. & Streuli, M. (2003) *J Biol Chem* **278,** 13265-70.
35 Hashimoto, Y. & Soderling, T. R. (1989) *J Biol Chem* **264,** 16524-9.
36 Kypta, R. M., Su, H. & Reichardt, L. F. (1996) *J Cell Biol* **134,** 1519-29.
37 Prekeris, R., Klumperman, J. & Scheller, R. H. (2000) *Mol Cell* **6,** 1437-48.
38 Okabe T, Nakamura T, Nishimura YN, Kohu K, Ohwada S, Morishita Y, Akiyama T. (2003) *J Biol Chem.* **278**, 9920-7.
39 Foletti, D. L., Blitzer, J. T. & Scheller, R. H. (2001) *J Neurosci* **21,** 5473-83.
40 Benzing, T., Yaffe, M. B., Arnould, T., Sellin, L., Schermer, B., Schilling, B., Schreiber, R., Kunzelmann, K., Leparc, G. G., Kim, E. & Walz, G. (2000) *J Biol Chem* **275,** 28167-72.
41 Balasubramanian, N., Levay, K., Keren-Raifman, T., Faurobert, E. & Slepak, V. Z. (2001) *Biochemistry* **40,** 12619-27.
42 Shin EY, Shin KS, Lee CS, Woo KN, Quan SH, Soung NK, Kim YG, Cha CI, Kim SR, Park D, Bokoch GM, Kim EG. (2002) *J Biol Chem* **277,** 44417-30.
43 Megidish, T., Cooper, J., Zhang, L., Fu, H. & Hakomori, S. (1998) *J Biol Chem* **273,** 21834-45.
44 Cumbay MG, Watts VJ. (2004) J Pharmacol Exp Ther. Mar 2
45 Benaim, G. & Villalobo, A. (2002) *Eur J Biochem* **269,** 3619-31.
46 Joyal, J. L., Crimmins, D. L., Thoma, R. S. & Sacks, D. B. (1996) *Biochemistry* **35,** 6267-75.
47 Chen, H. J., Rojas-Soto, M., Oguni, A. & Kennedy, M. B. (1998) *Neuron* **20,** 895-904.
48 Pei, L., Teves, R. L., Wallace, M. C. & Gurd, J. W. (2001) *J Cereb Blood Flow Metab* **21,** 955-63.
49 Dekker JW, Budhia S, Angel NZ, Cooper BJ, Clark GJ, Hart DN, Kato M. (2002) *Immunogenetics* **53**, 993-1001.
50 Xu, A., Suh, P. G., Marmy-Conus, N., Pearson, R. B., Seok, O. Y., Cocco, L. & Gilmour, R. S. (2001) *Mol Cell Biol* **21,** 2981-90.
51 Risinger, C. & Bennett, M. K. (1999) *J Neurochem* **72,** 614-24.
52 Walter, J., Fluhrer, R., Hartung, B., Willem, M., Kaether, C., Capell, A., Lammich, S., Multhaup, G. & Haass, C. (2001) *J Biol Chem* **276,** 14634-41.
53 Sygusch, J., Beaudry, D. & Allaire, M. (1990) *Arch Biochem Biophys* **283,** 227-33.
54 Haga, A., Niinaka, Y. & Raz, A. (2000) *Biochim Biophys Acta* **1480,** 235-44.
55 Tisdale, E. J. (2002) *J Biol Chem* **277,** 3334-41.
56 Cooper, J. A., Esch, F. S., Taylor, S. S. & Hunter, T. (1984) *J Biol Chem* **259,** 7835-41.
57 Korotchkina LG, Patel MS. (2001) *J Biol Chem* **276**, 37223-9.
58 Kuo, T. H., Giacomelli, F. & Wiener, J. (1986) *Biochem Biophys Res Commun* **139,** 56-63.
59 Yokozeki, T., Homma, K., Kuroda, S., Kikkawa, U., Ohno, S., Takahashi, M., Imahori, K. & Kanaho, Y. (1998) *J Neurochem* **71,** 410-7.
60 Roof DJ, Hayes A, Adamian M, Chishti AH, Li T. (1997) *J Cell Biol* **138**, 575-88.
61 Beausoleil SA, Jedrychowski M, Schwartz D, Elias JE, Villen J, Li J, Cohn MA, Cantley LC, Gygi SP. (2004) *Proc Natl Acad Sci U S A.* Aug 9
62 Ghosh, S., Dorsey, F. C. & Cox, J. V. (2002) *J Cell Sci* **115,** 4107-15.
63 Cianci CD, Giorgi M, Morrow JS. (1988) *J Cell Biochem* **37**, 301-15.
64 Amit, S., Hatzubai, A., Birman, Y., Andersen, J. S., Ben-Shushan, E., Mann, M., Ben-Neriah, Y. & Alkalay, I. (2002) *Genes Dev* **16,** 1066-76.
65 Hagen, T. & Vidal-Puig, A. (2002) *Biochem Biophys Res Commun* **294,** 324-8.
66 Roura, S., Miravet, S., Piedra, J., Garcia de Herreros, A. & Dunach, M. (1999) *J Biol Chem* **274,** 36734-40.
67 Wong HN, Ward MA, Bell AW, Chevet E, Bains S, Blackstock WP, Solari R, Thomas DY, Bergeron JJ. (1998) *J Biol Chem.* **273.** 17227-35
68 Wilde A, Brodsky FM. (1996) *J Cell Biol* **135**, 635-45.
69 Faundez VV, Kelly RB. (2000) *Mol Biol Cell* **11**. 2591-604.
70 Wilde, A. & Brodsky, F. M. (1996) *J Cell Biol* **135,** 635-45.
71 Hao, W., Luo, Z., Zheng, L., Prasad, K. & Lafer, E. M. (1999) *J Biol Chem* **274,** 22785-94.
72 Wilde, A., Beattie, E. C., Lem, L., Riethof, D. A., Liu, S. H., Mobley, W. C., Soriano, P. & Brodsky, F. M. (1999) *Cell* **96,** 677-87.
73 Martinez MC, Ochiishi T, Majewski M, Kosik KS. (2003) *J Cell Biol* **162,** 99-111.
74 Khanna R, Chang SH, Andrabi S, Azam M, Kim A, Rivera A, Brugnara C, Low PS, Liu SC, Chishti AH. (2002) *Proc Natl Acad Sci U S A* **99.** 6637-42.
75 Gronborg M, Kristiansen TZ, Stensballe A, Andersen JS, Ohara O, Mann M, Jensen ON, Pandey A. (2002) *Mol Cell Proteomics* **1.** 517-27.
76 Vogt TF, Jackson-Grusby L, Rush J, Leder P. (1993) *Proc Natl Acad Sci U S A* **90.** 5554-8.
77 Matthies, H. J., Miller, R. J. & Palfrey, H. C. (1993) *J Biol Chem* **268,** 11176-87.
78 Job, D., Rauch, C. T., Fischer, E. H. & Margolis, R. L. (1983) *Proc Natl Acad Sci U S A* **80,** 3894-8.
79 Fujii, T., Watanabe, M. & Nakamura, A. (1996) *Neurochem Int* **28,** 535-44.
80 Avila, J., Ulloa, L., Gonzalez, J., Moreno, F. & Diaz-Nido, J. (1994) *Cell Mol Biol Res* **40,** 573-9.
81 Goold RG, Gordon-Weeks PR. (2003) *J Neurochem* **87**, 935-46.
82 Chang L, Jones Y, Ellisman MH, Goldstein LS, Karin M. (2003) *Dev Cell* **4**, 521-33.
83 Akiyama, T., Kadowaki, T., Nishida, E., Kadooka, T., Ogawara, H., Fukami, Y., Sakai, H., Takaku, F. & Kasuga, M. (1986) *J Biol Chem* **261,** 14797-803.
84 Kadowaki, T., Fujita-Yamaguchi, Y., Nishida, E., Takaku, F., Akiyama, T., Kathuria, S., Akanuma, Y. & Kasuga, M. (1985) *J Biol Chem* **260,** 4016-20.
85 Kitazawa, H., Iida, J., Uchida, A., Haino-Fukushima, K., Itoh, T. J., Hotani, H., Ookata, K., Murofushi, H., Bulinski, J. C., Kishimoto, T. & Hisanaga, S. (2000) *Cell Struct Funct* **25,** 33-9.
86 Morishima-Kawashima, M., Hasegawa, M., Takio, K., Suzuki, M., Yoshida, H., Titani, K. & Ihara, Y. (1995) *J Biol Chem* **270,** 823-9.
87 Kishimoto, A., Nishiyama, K., Nakanishi, H., Uratsuji, Y., Nomura, H., Takeyama, Y. & Nishizuka, Y. (1985) *J Biol Chem* **260,** 12492-9.
88 Shoji, S., Ohnishi, J., Funakoshi, T., Fukunaga, K., Miyamoto, E., Ueki, H. & Kubota, Y. (1987) *J Biochem (Tokyo)* **102,** 1113-20.
89 Hirschberg D, Radmark O, Jornvall H, Bergman T. (2003) *J Protein Chem* **22**, 177-81.
90 Conti, M. A., Sellers, J. R., Adelstein, R. S. & Elzinga, M. (1991) *Biochemistry* **30,** 966-70.
91 Moussavi, R. S., Kelley, C. A. & Adelstein, R. S. (1993) *Mol Cell Biochem* **127-128,** 219-27.
92 Karcher, R. L., Roland, J. T., Zappacosta, F., Huddleston, M. J., Annan, R. S., Carr, S. A. & Gelfand, V. I. (2001) *Science* **293,** 1317-20.
93 Garver TD, Ren Q, Tuvia S, Bennett V. (1997) *J Cell Biol* **137**, 703-14.
94 Yang SD, Huang JJ, Huang TJ. (1995) *J Neurochem* **64.** 1848-54.
95 Kutzleb C, Sanders G, Yamamoto R, Wang X, Lichte B, Petrasch-Parwez E, Kilimann MW. (1998) *J Cell Biol* **143**, 795-813.
96 van de Velde HJ, Roebroek AJ, Senden NH, Ramaekers FC, Van de Ven WJ. (1994) *J Cell Sci* **107**, 2403-16.
97 Nicolas, G., Fournier, C. M., Galand, C., Malbert-Colas, L., Bournier, O., Kroviarski, Y., Bourgeois, M., Camonis, J. H., Dhermy, D., Grandchamp, B. & Lecomte, M. C. (2002) *Mol Cell Biol* **22,** 3527-36.
98 Sihag, R. K. (1998) *J Neurochem* **71,** 2220-8.
99 Otto, I. M., Raabe, T., Rennefahrt, U. E., Bork, P., Rapp, U. R. & Kerkhoff, E. (2000) *Curr Biol* **10,** 345-8.
100 Hosaka, M., Hammer, R. E. & Sudhof, T. C. (1999) *Neuron* **24,** 377-87.
101 Hilfiker, S., Pieribone, V. A., Nordstedt, C., Greengard, P. & Czernik, A. J. (1999) *J Neurochem* **73,** 921-32.
102 Jena, B. P., Webster, P., Geibel, J. P., Van den Pol, A. N. & Sritharan, K. C. (1997) *Cell Biol Int* **21,** 469-76.
103 Wandosell, F., Serrano, L., Hernandez, M. A. & Avila, J. (1986) *J Biol Chem* **261,** 10332-9.
104 Serrano, L., Diaz-Nido, J., Wandosell, F. & Avila, J. (1987) *J Cell Biol* **105,** 1731-9.
105 Ajiro, K. (2000) *J Biol Chem* **275,** 439-43.
106 Madigan, J. P., Chotkowski, H. L. & Glaser, R. L. (2002) *Nucleic Acids Res* **30,** 3698-705.
107 Ajiro, K., Yoda, K., Utsumi, K. & Nishikawa, Y. (1996) *J Biol Chem* **271,** 13197-201
108 Pai, C. Y., Chen, H. K., Sheu, H. L. & Yeh, N. H. (1995) *J Cell Sci* **108,** 1911-20.
109 Shav-Tal, Y., Cohen, M., Lapter, S., Dye, B., Patton, J. G., Vandekerckhove, J. & Zipori, D. (2001) *Mol Biol Cell* **12,** 2328-40.
110 Rosenberger, U., Lehmann, I., Weise, C., Franke, P., Hucho, F. & Buchner, K. (2002) *J Cell Biochem* **86,** 394-402.
111 Yeakley, J. M., Tronchere, H., Olesen, J., Dyck, J. A., Wang, H. Y. & Fu, X. D. (1999) *J Cell Biol* **145,** 447-55.
112 Peters, H. I., Chang, Y. W. & Traugh, J. A. (1995) *Eur J Biochem* **234,** 550-6.
113 Izawa, T., Fukata, Y., Kimura, T., Iwamatsu, A., Dohi, K. & Kaibuchi, K. (2000) *Biochem Biophys Res Commun* **278,** 72-8.
114 Oester, D. A., Caterson, B. & Schwartz, E. R. (1986) *Biochem Biophys Res Commun* **137,** 599-605.

Appendix 1 - Synaptic proteins identified in IMAC purifications identified by mass spectrometry. Mr, relative molecular mass; Rel Abun Rank, rank of relative abundance (number of peptides detected for each protein normalized to molecular weight: PSD, postsynaptic density proteins; NRC, NMDA receptor-signaling complex. Peptides detected, total number of matching peptides sequenced by MS/MS for a specific protein. Proteins previously shown to be phosphorylated and known kinases responsible for this phosphorylation were identified from the literature. Proteins identified in this dataset, which have been shown to be present in the post-synaptic density and in NRC complexes, are shown. Scansite (high stringency and a percentile score of greater than 0.2%) and NetphosK kinase hits are included and are represented by a letter corresponding to the following kinases: a, Akt Kinase; b, ATM Kinase; c, Calmodulin dependent Kinase 2; d, Casein Kinase 1; e, Casein Kinase 2; f, Cdc2 Kinase; g, Cdk5 Kinase; h, Clk2 Kinase; I, DNA PK; k, EGFR Kinase; j, Erk1 Kinase; l, GSK3 Kinase; m, Insulin Receptor Kinase; n, Itk Kinase; o, p38 MAPK; p, PDGFR Kinase; q, Protein Kinase A; r, PKC alpha/beta/gamma; s, PKC delta; t, PKC epsilon; u, PKC mu; v, PKC zeta; w, Src kinase.

Appendix 2

Accession	Protein Name	Phosphopeptide sequences	Residues	Mod. Id	Mr	m/z	Biochemistry			
							PR	DO	PEw	PEi
Q8BKY9	1110030H18Rik	SEDDAASGEHDIQIEGVR	25-42	2 pS	2184.85	1093.43/729.29			+	
Q8CI87	5730405I09Rik protein	SASADNLILPR	299-309	2 pS	1263.58	632.80			+	
Q8K4G5	Actin-binding LIM protein 1 isoform a	TLSPTPSAEGVQDVR	473-487	1 pS	1699.75	850.95	+	+		
		STSQGSINSPVYSR	494-507	1 pS/T	781.84	1561.66		+		
Q8R502	AD158	SQSLKSIPEK	210-219	2 pS	1303.68	652.8476				+
		SQSLKSIPEKFVVDK	210-224	2 pS	1905.93	636.34/954.00				+
P17426	Adapter-related protein complex 2 alpha 1 subunit	GGVEPTPSTVSTPSPSADLLGLR	642-664	3 pS/T	2519.21	1260.62				+
		GVEPTPSTVSTPSPSADLLGLR	643-664	3 pS/T	2462.23	1232.12/821.74				+
		VEPTPSTVSTPSPSADLLGLR	644-664	3 pS/T	2405.21	1203.61				+
		TPSTVSTPSPSADLLGLR	647-664	3 pS/T	2065.99	1034				+
		STVSTPSPSADLLGLR	649-664	3 pS/T	1867.90	934.97		+		+
		TVSTPSPSADLLGLR	649-664	2 pS/T	1700.73	851.37			+	
		TVSTPSPSADLLGLR	650-664	3 pS/T	1780.83	891.42				+
		TPSPSADLLGLR	653-664	2 pS/T	1413.71	707.86			+	+
Q6QR53	Adapter-related protein complex 3 beta 3	AFYGSEEDEAKGPGSEEAATAALPAR	268-293	1 pS	2703.11	902.04		+		
P51830	Adenylate cyclase, type IX	TASPGSVSDLAQTVK	608-622	2 pS	1647.82	824.91				+
P26231	Alpha-1 catenin	SRTSVQTEDDQLIAGQSAR	652-670	2 pS	2276.93	1139.47			+	
Q01484	Ankyrin 2	SDSNASFLR	29-37	1 pS	1075.43	538.72		+		
		SDSNASFLR	29-37	1 pS	1075.43	538.72		+		
		SHLVNEVPVLASPDLLSEVSEMKQDLIK	1456-1483	1 pS	3169.28	1057.4331	+			
		SLESSPVEPKMK	2435-2446	2 pS	1532.76	767.39		+		
		KRFTPEEEMFK	2547-2557	1 pT	1520.67	761.34/507.90		+		
		RFTPEEEMFK	2548-2557	1 pT	1392.57	697.29		+		
Q9Z2H5	Band 4.1-like protein 1	SEAEEGEVRTPTK	466-478	2 pS/T	1661.71	831.86			+	
		RLPSSPASPSPKGTPEK	537-553	2 pS/T	1894.86	632.63		+		
		SLPELDRDKSDSETEGLVFAR	639-659	1 pS	2443.08	815.37		+		
		DKSDSETEGLVFAR	646-659	2 pS	1782.82	892.4178				+
Q9WV92	Band 4.1-like protein 3	QLEVQQFEDDKLSQR	84-98	1 pS	2005.86	1003.94/669.63		+		
		GISQTNLITTVTPEKK	484-499	1 pS	1808.90	905.46		+		
		GISQTNLITTVTPEKK	484-499	2 pS/T	1916.89	959.45			+	
		TDTAADGETSATESDQEEDAEIK	534-556	2 pS	2712.01	1357.01/905.01			+	
		TSATESDQEEDAEIK	542-556	2 pS/T	1909.70	955.86			+	
		TESDQEEDAEIK	545-556	1 pS	1570.64	786.33			+	
		APMIEPLVPEETKQSSGEKLMDGSEILSLLESAR	612-645	3 pS/T	3955.75	990.23	+			
O88737	Bassoon	MGNEASLEGGAGEGPLPPGGSGLGPGPGAGKPPSALAGGGQLPVAGAAR	1-49	1 pS	4341.33	1448.1163	+			
		SKTPPSNLSPIEDASPTEELR	1100-1120	4 pS/T	2657.22	886.75				+
		SKTPPSNLSPIEDASPTEELR	1100-1120	3 pS/T	2507.00	1254.51		+	+	
		TPPSNLSPIEDASPTEELR	1102-1120	2 pS	2211.94	1106.98			+	
		TPPSNLSPIEDASPTEELRQAAEMEELHR	1102-1130	2 pS	3518.94	880.74		+		
		PPSNLSPIEDASPTEELR	1103-1120	2 pS	2180.90	1091.46			+	
		SNLSPIEDASPTEELR	1105-1120	2 pS	1986.78	994.40			+	
		NLSPIEDASPTEELR	1106-1120	2 pS	1899.77	950.89			+	
		LSPIEDASPTEELR	1107-1120	2 pS	1685.84	893.93		+	+	+
		SPIEDASPTEELR	1108-1120	2 pS	1672.86	837.48		+	+	+
		SPSTSSTIHSYGQPPTTANYGSQTEELPHAPSGPPGSGR	1414-1452	2 pS/T/Y	4139.84	1035.97	+			
		SSPPLSPSTPSESPTFSPGKLGPR	1481-1504	2 pS	2597.34	866.79		+		
		SSPPLSPSTPSESPTFSPGKLGPR	1481-1504	4 pS	2757.3	920.11		+		
		SESPTFSPGKLGPR	1491-1504	2 pS/T	1646.84	824.43		+		
		ATAEFSTQTPSLTLSSDIPR	1505-1524	2 pS/T	2323.92	1162.9686	+			
		LVWQQSSQEAPIMVITLASDASSQTR	1548-1573	2 pS	3048.22	1017.0797	+			
		HSYSLGFADGR	2026-2036	1 pS	1288.52	645.27		+		
		HGSGSGGPDLVQYQPQHGPGLSAPQGLAPLR	2120-2150	2 pS	3280.62	1094.55	+			
		SPVSPLSPHR	2805-2813	2 pS	1249.59	625.8		+		

Appendix 2

Accession	Protein Name	Phosphopeptide sequences	Residues	Mod. Id	Mr	m/z	Biochemistry			
							PR	DO	PEw	PEi
O88737	Bassoon	LLDTSFASSER	2815-2825	1 pS	1304.56	653.29		+		
		SLSDPKPLSP	2858-2867	2 pS	1227.58	614.8		+		
		SLSDPKPLSPTA	2858-2869	2 pS	1399.7	700.86		+		
		SLSDPKPLSPTAE	2858-2870	2 pS	1542.74	772.38		+		
		SLSDPKPLSPTAEE	2858-2871	2 pS	1685.86	843.93		+		
		SLSDPKPLSPTAEESAK	2858-2874	1 pS	1835.04	918.53	+	+		
		SLSDPKPLSPTAEESAK	2858-2874	2 pS	1971.85	986.93			+	
		SLSDPKPLSPTAEESAKER	2858-2876	2 pS	2000.98	708.0235	+	+		+
		SLSDPKPLSPTAEESAKER	2858-2876	3 pS	2351.18	784.73		+		
		SLSDPKPLSPTAEESAKERF	2858-2877	2 pS	2418.27	807.10		+		
		SDPKPLSPTAEESAKER	2860-2876	2 pS	2071.05	691.36		+		
		TLPSPPPEEAHLPLAGQVPSQLYAASLLQR	2905-2934	1 pS	3259.33	1087.4509	+			
		SSVSQSPAPTYPSDSHYTSLEQNVPR	3213-3238	2 pS	3036.36	1013.1268	+			
		ISKFSPIEEAKDVESDLASYPPPTVSSSLTSR	3378-3409	1 pS	3516.33	1173.1168	+			
NP_878259	Brain-specific protein p25 alpha	AANKTPPKSPGDPAR	11-25	1 pT	1585.74	529.59		+		
		AANKTPPKSPGDPAR	11-25	2 pS/T	1665.70	833.86/556.24		+		
		RLSLESEGANEGATAAPELSALEEAFRR	29-56	1 pS	3053.39	1018.80		+		
P11798	Calcium/calmodulin-dependent protein kinase type II alpha	ESSESTNTTIEDEDTKVR	329-346	2 pS	2297.87	1149.94/766.97			+	
P28652	Calcium/calmodulin-dependent protein kinase type II beta	NFSVGRQTTAPATMSTAASGTTMGLVEQAK	313-342	1 pS	3136.74	785.19	+			
Q9R196	Calcium-activated potassium channel alpha subunit 1	RLEDEQPPTLSPK	701-713	2 pS/T	1724.82	863.42				+
		SSPDNSPVHGMLR	923-935	2 pS	1583.72	792.87				+
P35564	Calnexin precursor	QKSDAEEDGVTGSQDEEDSKPK	551-572	2 pS	2664.06	889.03			+	
P10644	cAMP-dependent protein kinase type I-alpha regulatory chain	TDSREDEISPPPPNPVVK	75-92	2 pS	2205.97	1103.99			+	
Q8K1M3	cAMP-dependent protein kinase type II-alpha regulatory chain	SDSEDADLEVPVPSK	75-91	2 pS	1830.89	916.45				+
Q8K070	CDNA sequence BC034054	SLDEDEPPPSPLAR	108-121	1 pS	1601.68	801.85		+		
O35927	Delta catenin-2	ALQSPEHHIDPIYEDR	409-424	1 pS	1998.86	1000.43/667.29		+		
		SPSIDSIQKDPR	529-540	2 pS	1543.77	772.8928				+
Q9WV69	Dematin	STSPPPSPEVWAESR	90-104	1 pS	1705.70	853.86		+		
		KGAEEEEEEEDDDSEEEIKAIR	213-234	1 pS	2658.04	887.02		+		
Q8BJ42	DISKS large-associated protein 2/SAPAP2	SGSYIKAMGDEESG	438-451	2 pS	1645.72	823.87		+		
		SGSYIKAMGDEESGESD	438-454	3 pS	2084.83	1043.42		+		
AAO89219	DISKS large-associated protein 3/SAPAP3	AGLGHLSPEGPLSLSEGPSSVGPEGGPGGVGAGGGSSTFPR	52-92	1 pS	3751.76	1251.59	+			
		SGSYIKAMGDEESG	392-406	2 pS	1645.6	823.87		+		
		SGSYIKAMGDEESGDSD	392-409	3 pS	2070.86	1036.44		+		
		SGSYIKAMGDEESGDSDG	392-410	3 pS	2127.9	1064.96		+		
		GSPKTSPKALAR	409-420	2 pS	1385.72	693.87		+		
		SPKTSPKALAR	410-420	2 pS	1328.68	665.35		+		
P60904	DnaJ homolog subfamily C member 5 (Cysteine string protein)	SLSTSGESLYHVLGLDK	8-24	1 pS	1884.86	943.44		+		
Q9QXS6	Drebrin	LSSPVLHR	140-147	1 pS	987.49	494.75		+		
Q9Z188	Dual-specificity tyrosine-phosphorylation regulated kinase 1B	IYQYIQSR	270-277	1 pY	1149.52	575.77		+		
Q9JL04	Formin 2	SIQTSPTEEGRIL	720-732	1 pS	1551.75	776.88		+		
Q8BXF0	G2 protein	SRQSLNSPSPGETEMDLLVTR	466-486	3 pS	2612.28	871.77				+
		SRQSLNSPSPGETEMDLLVTR	466-486	4 pS	2692.25	898.4213				+
Q8BJY3	Glutamate receptor, metabotropic 5	SPSPISTLSHLAG	135-151	2 pS	1439.69	720.85				+
		SPSPISTLSHLAGSAGR	135-155	2 pS	1810.9	906.46/604.63			+	+
Q8BH66	Guanine nucleotide-binding protein 3	SSDWSSEEEEPVRK	18-31	2 pS	1907.88	954.9475				+
		SSDWSSEEEEPVRK	18-31	3 pS	1987.83	994.92				+
Q8C3Q5	Hypothetical protein (D430041B17)	SSSLTPGLGGPDSMAPR	47-63	1 pS	1708.74	855.3783	+			
		LVSQEHLLLSSPEALR	227-242	1 pS	1870.95	936.4837	+			
		QGTLEQLQFIPGHHLPQHLR	329-348	1 pT	2428.2	810.4094	+			
Q7L0D6	Hypothetical protein KIAA0773	DTDAYSDLSDGEKEAR	77-92	2 pS	2028.76	1015.39/677.26			+	
Q80SW1	IP3R binding protein released with inositol 1,4,5-trisphosphate	SYTDSSDDEVSPR	80-92	3 pS/T	1766.7	884.36				+
P70404	Isocitrate dehydrogenase [NAD] subunit gamma	HTPDIGGQGTTSQAIQDIIR	362-381	1 pT	2187.03	1094.48	+			
XP_194040	Microtubule-associated protein 1A	ELAISSPEDLTQDFEELKR	760-778	2 pS	2379.03	1190.57/794.07	+	+		
		AELEEMEEVHPSDEEEEETKAESFYQK	894-918	1 pS	3321.31	1108.11		+		
		TISPTSSLEEDKGFK	1131-1145	2 pS	1853.93	927.97/618.97		+		
		TISPTSSLEEDKGFK	1131-1145	3 pS	1933.92	967.97		+		+

Appendix 2

Accession	Protein Name	Phosphopeptide sequences	Residues	Mod. Id	Mr	m/z	Biochemistry			
							PR	DO	PEw	PEi
XP_194040	Microtubule-associated protein 1A	EMTLDQKSPEK	1811-1821	1 pS	1384.58	693.30		+		
		ALGLEESPEEEGKAR	1838-1852	1 pS	1693.73	847.94	+	+		
		ETSPTRGEPVPAWEGKSPEQEVR	1870-1892	1 pS	1819.81	910.91/607.61		+		
		ETSPTRGEPVPAWEGKSPEQEVR	1870-1892	2 pS/T	2725.12	1363.58/909.38		+		
		GEPVPAWEGKSPEQEVR	1876-1892	1 pS	1973.87	987.94		+		
		WLAESPVGLPPEEEDKLTR	1981-1999	1 pS	2245.06	749.36		+		
		SPFEIISPPASPPEMTG	2000-2016	2 pS	1957.98	980		+		
		SPFEIISPPASPPEMTGQR	2000-2018	2 pS	2199.91	1100.96/734.32		+		
P14873	Microtubule-associated protein 1B	SLMSSPEDLTK	825-835	2 pS	1408.65	705.3315				+
		SLMSSPEDLTKDFEELKAEEIDVAK	825-849	2 pS	2983.23	995.42		+		
		DVMSDETNNEETESPSQEFVNITK	1138-1161	1 pS/T	2822.14	1412.08/941.71		+		
		DVSDERLSPAKSPSLSPSPPSPIEK	1240-1264	4pS	3009.18	1004.07			+	
		SPSLSPSPPSPIEK	1251-1264	1 pS	1501.69	751.85		+		
		SPSLSPSPPSPIEK	1251-1264	2 pS	1581.65	791.83		+	+	
		SVSPGVTQAVVEEHCASPEEK	1291-1311	1 pS	2261.99	1132.05	+	+		
		TLEVVSPSQSVTGSAGHTPYYQSPTDEK	1312-1339	1 pS + 1 pY	3124.31	1042.44		+		
		TLEVVSPSQSVTGSAGHTPYYQSPTDEK	1312-1339	3 pS	3204.25	1069.09		+		
		ASLSPMDEPVPDSESPVEK	1370-1388	1 pS	2093.09	1047.55	+			
		ASLSPMDEPVPDSESPVEK	1370-1388	2 pS	2172.83	1087.42		+		
		ASLSPMDEPVPDSESPVEK	1370-1388	2 pS	2256.88	1129.45			+	
		SESPVEKVLSPLR	1382-1394	2 pS	1641.88	821.95/548.30		+		
		VLSPLRSPPLLGSESPYEDFLSADSK	1389-1414	2 pS	2963.33	988.79		+		
		QGFPDRESPVSDLTSTGLYQDKQEEK	1431-1456	1 pS/T	3033.31	1012.11		+		
		SISPPDFSPK	1614-1623	2 pS	1261.52	631.77			+	
		SISPPDFSPKTAK	1614-1626	2 pS	1561.77	521.6		+		
		VQSLEGEKLSPK	1766-1777	2 pS	1515.69	758.85			+	
		VQSLEGEKLSPKSDISPLTPR	1766-1786	4 pS/T	2599.81	867.61		+		
		VQSLEGEKLSPKSDISPLTPR	1766-1786	4 pS/T	2656.31	886.44				+
		SDISPLTPR	1778-1786	1 pS	1064.49	533.25		+		
		SDISPLTPR	1778-1786	2 pS/T	1172.50	587.26			+	
		ESSPLYSPGFSDSTSAAK	1787-1804	1 pS	1909.78	955.90		+		
		ESSPLYSPGFSDSTSAAK	1787-1804	2 pS	1989.75	995.88		+	+	
		ETAAAHQASSSPPIDAATAEPYGFR	1805-1829	1 pS/T	2624.4	875.81	+			
		ETAAAHQASSSPPIDAATAEPYGFR	1805-1829	1 pT	2624.11	1313.13/875.71	+	+		
		TPGDFNYAYQKPENAAGSPDEEDYDYESQEK	1860-1890	1 pS	3636.37	1213.13		+		
		TIKSPCDSGYSYETIEK	1908-1924	1 pS	1999.82	1000.92		+		
		TTRTPEEGGYSYEISEK	1942-1958	1 pT	2025.81	1013.91/676.27	+	+		
		TTRSPDTSAYCVETMEK	2027-2043	1 pS	2061.79	1031.9/688.23		+		
		TELSPSFINPNPLEWFAGEEPTEESEKPLTQSGGAPPPSGGK	2091-2132	1 pS	4487.99	1497.00		+		
P20357	Microtubule associated protein 2	GHDLSPLASDILTNTSGSMDEGDDYLPPTTPAVEK	722-756	1 pY	3722.68	1241.9012	+			
		ETSPETSLIQDEVALK	1159-1174	1 pS	1838.84	920.4294		+		
		DGSPDAPATPEKEEVAFSEYK	1351-1371	2 pS	2425.92	1213.97/809.65		+		
		DKVTDGISKSPEKR	1530-1543	1 pS	1638.79	547.27		+		
		ARVDHGAEIITQSPSR	1771-1786	1 pS	1815.86	606.29	+	+		
		VDHGAEIITQSPSR	1773-1786	1 pS	1588.7	795.39	+	+		
		RLSNVSSSGSINLLESPQLATLAEDVTAALAK	1794-1825	3 pS	3494.58	1165.8684	+			
P27546	Microtubule-associated protein 4	AAVGVTGNDITTPPNKEPPPSPEKK	647-671	2 pS/T	2703.22	902.08		+		
P10637	Microtubule-associated protein Tau	SGERSGYSSPGSPG	482-495	3 pS	1591.5	796.8		+		
		SGYSSPGSPGTPGSR	486-500	2 pS	1566.58	784.3375		+		
		SGYSSPGSPGTPGSR	486-500	3 pS	1646.66	824.3354		+		
		SGYSSPGSPGTPGSR	486-500	1 pS	1472.59	737.36	+			
		SGYSSPGSPGTPGSR	486-500	1 pS	1472.71	737.36	+	+		
		SGYSSPGSPGTPGSRSR	486-502	2 pS	1809.82	905.92		+		
		AKTDHGAEIVYKSPVVS	675-691	2 pS	2002.03	668.35		+		
		AKTDHGAEIVYKSPVVSG	675-692	2 pS	2058.94	1030.55/687.35		+		
		AKTDHGAEIVYKSPVVSGD	675-693	2 pS	2188.13	1095.07/730.37		+		
		AKTDHGAEIVYKSPVVSGDTSPR	675-697	3 pS	2652.91	885.31	+	+		

Appendix 2

Accession	Protein Name	Phosphopeptide sequences	Residues	Mod. Id	Mr	m/z	Biochemistry			
							PR	DO	PEw	PEi
P10637	Microtubule-associated protein Tau	AKTDHGAEIVYKSPVVSGDTSPR	675-697	3 pS	2709.36	904.1293				+
		TDHGAEIVYKSPVVS	677-691	2 pS	1802.91	902.46		+		
		TDHGAEIVYKSPVVSG	677-692	2 pS	1859.93	930.97/620.99		+	+	
		TDHGAEIVYKSPVVSGD	677-693	2 pS	1988.98	995.50/664.00		+	+	
		TDHGAEIVYKSPVVSGDTSPR	677-697	3 pS/T	2453.97	819.00		+	+	
		TDHGAEIVYKSPVVSGDTSPR	677-697	1 pT	2294.11	765.71	+			
		TDHGAEIVYKSPVVSGDTSPR	677-697	2 pS/T	2374.00	1188.01/792.34		+		
		TDHGAEIVYKSPVVSGDTSPRHL	677-699	3 pS	2760.38	921.13				+
		SPVVSGDTSPR	687-697	1 pS	1180.49	591.25		+		
		SPVVSGDTSPR	687-697	1 pS/T	2652.91	885.31	+			
		SPVVSGDTSPR	687-697	2 pS/T	1288.59	645.3				+
		SPVVSGDTSPRH	687-698	2 pS/T	1425.65	713.83		+		
		SPVVSGDTSPRHL	687-699	2 pS/T	1538.76	770.39/513.92		+		
		SPVVSGDTSPRHL	687-699	3 pS/T	1618.73	810.37		+		
		SPVVSGDTSPRHLSNVSSTG	687-706	2 pS/T	2171.07	1086.54		+		
		SPVVSGDTSPRHLSNVSSTG	687-706	3 pS/T	2251.04	1126.53		+		
		HLSNVSSTGSIDMVDSPQLATLADEVSASLAK	698-729	1 pS/T	3322.21	1108.4116	+			
		HLSNVSSTGSIDMVDSPQLATLADEVSASLAK	698-729	1 pT	3322.81	1108.61	+			
		HLSNVSSTGSIDMVDSPQLATLADEVSASLAK	698-729	1 pS/T	3323.73	1108.92	+			
		HLSNVSSTGSIDMVDSPQLATLADEVSASLAK	698-729	2 pS/T	3403.85	1135.62	+			
		HLSNVSSTGSIDMVDSPQLATLADEVSASLAK	698-729	3 pS/T	3483.61	1162.21	+			
BAC65687	mKIAA0942 protein	LEWAVDDEEKKKSPSEGTDEK	12-32	1 pS	2499.07	834.03		+		
		IGSTTNPFLDIPHDPNAAVYK	44-64	1 pS/T	2348.91	1175.4617	+			
		IGSTTNPFLDIPHDPNAAVYK	44-64	1 pS/T	2349.1	1175.64	+			
		IGSTTNPFLDIPHDPNAAVYK	44-64	1 pS	2349.07	1175.55/784.03		+		
		DEYKPEKSLSDEDLK	105-119	1 pS	1874.82	625.93		+		
		KSHSSPSLNPDASPVTAKVKR	236-256	1 pS	2323.18	772.04		+		
		KSHSSPSLNPDASPVTAK	236-253	2 pS	1981.84	991.92/661.62		+		
		SHSSPSLNPDASPVTAK	237-253	2 pS	1853.75	927.8839	+	+		
		SHSSPSLNPDASPVTAK	237-253	3 pS	1933.58	967.7955	+			
		SHSSPSLNPDASPVTAK	237-253	2 pS	1881.9	941.96/628.32		+		+
		SHSSPSLNPDASPVTAK	237-253	3 pS	1961.89	981.95/654.96		+		+
		SHSSPSLNPDASPVTAK	237-253	1 pS	1773.79	887.91	+			
		SHSSPSLNPDASPVTAK	237-253	1 pS	1773.75	887.88/592.26	+	+		
		SHSSPSLNPDASPVTAK	237-253	2 pS	1853.76	927.94	+			
		SHSSPSLNPDASPVTAK	237-253	2 pS	1853.81	927.91	+	+		
		SHSSPSLNPDASPVTAK	237-253	3 pS	1933.73	967.93	+			
		SHSSPSLNPDASPVTAK	237-253	2 pS	1853.74	618.92		+		
		SHSSPSLNPDASPVTAKVKR	237-256	2 pS	2265.18	756.07		+		
		SHSSPSLNPDASPVTAKVKR	237-256	3 pS	2345.15	782.7233		+		
		SHSSPSLNPDASPVTAKVKRNVSER	237-261	3 pS	2944.32	737.13		+		
		SSPSLNPDASPVTAK	239-253	2 pS	1657.7	829.91		+		
		SSPSLNPDASPVTAKVKR	239-254	2 pS	2040.96	681.36		+		
		SSPSLNPDASPVTAKR	239-254	3 pS	2121.06	708.03		+		
P53986	Monocarboxylate transporter 1 (SLC16A1)	LKSKESLQEAGK	208-219	2 pS	1518.66	760.35/507.23				+
O08599	Munc-18-1	TAKVRAAHVFFTDSCPDALFNELVK	96-120	1 pS/T	2965.78	742.45	+			
		AAHVFFTDSCPDALFNELVK	101-120	1 pT	2303.04	1152.5298	+			
		SQVYSLDSADSFQSFYSPHK	142-161	2 pS + 1 pY	2531.29	1266.65	+			
O08599	Munc-18-1	DPSSPVLHELTFQAMSYDLLPIENDVYK	238-265	2 pS/T	3423.5	1141.85	+			
		STHLHLAEDCMK	345-356	2 pS/T	1543.56	773.36	+			
		EDSEIITNMAHLGVPIVTDSTLR	435-457	1 pS/T	2633.82	878.4	+			
		SSASFSTTAVSAR	506-518	1 pS	1350.79	676.29		+		
P04370	Myelin basic protein (generic)	SKYLATASTMDHAR	144-157	1 pS	1630.92	816.4684	+			
		TTHYGSLPQK	196-205	1 pS	1210.54	606.27		+		
		NIVTPRTPPPSQGK	223-236	1 pT	1570.78	786.4019	+			
P04370-6, 8, 11	Myelin basic protein (6,8,11)	FSWGGRDSR	138-146	1 pS	1146.25	574.23		+		
P04370-4, 5, 9, 10	Myelin basic protein isoform (isoforms 4, 5, 7, 9, 10)	FSWGAEGQKPGFGYGGR	138-154	1pS	1879.79	940.90/627.60	+	+		

Appendix 2

Accession	Protein Name	Phosphopeptide sequences	Residues	Mod. Id	Mr	m/z	Biochemistry			
							PR	DO	PEw	PEi
P02688	Myelin basic protein S	PQKSQRTQDENPVVHFFK	97-114	1 pS/T	2264.07	1132.7	+			
O35413	nArgBP2	RGSLPDNSILHR	842-853	1 pS/T	1443.68	482.24	+			
P13595	Neural cell adhesion molecule 1	AAFSKDESKEPIVEVR	767-782	1 pS	1883.87	628.96		+		
Q810U3	Neurofascin precursor	DKEETEGNESSEA	1217-1229	2 pS	1681.69	841.85				+
		DKEETEGNESSEATSPVNAIY	1217-1237	3 pS/T	2607.19	1304.60/870.26				+
P08553	Neurofilament triplet M protein	EEAEEKEEEPEAEKSPVKSPEAK	488-510	2 pS	2912.26	971.76			+	
Q9Z0P4	Paralemmin	SETLVNAQQTPLGTPK	132-147	2 pT	1870.97	936.4928				+
		TSTPVRSPGGSTMMK	151-165	2 pS/T	1709.66	855.84				+
Q9Z1B3	Phospholipase C-beta-1	TYLTAGQLAGNSSVEMYR	333-350	2 pT+1 pY	2199.83	1100.57	+			
Q9QYX7	Piccolo	SSPSSPSDLAK	1307-1317	2 pS	1262.56	632.29				+
		EHKENPELVDDLSPR	1397-1411	1 pS	1856.80	619.94		+		
		HSWHDEDDETFDESPELK	1532-1549	1 pS	2294.85	765.96		+		
		ERPKTPPSNLSPIEDASPTEELR	1726-1748	3 pS/T	2802.17	935.06		+	+	
		TPPSNLSPIEDASPTEELR	1730-1748	2 pS	2211.94	1106.98		+		
		TPPSNLSPIEDASPTEELRQAAEMEELHR	1730-1758	2 pS	3518.94	880.74		+		
		LSPIEDASPTEELR	1735-1748	2 pS	1785.84	893.93		+		+
		SPIEDASPTEELR	1736-1748	2 pS	1672.7	837.48		+		+
		VYKLPTAVSLYSPTDEQSVMQK	1788-1809	1 pS	2563.22	855.4135	+			
		LPTAVSLYSPTDEQSVMQK	1791-1809	1 pS	2173.13	1087.57	+			
		SMSDPKPLSPTADE	3578-3590	2 pS	1689.77	845.89		+		
		SMSDPKPLSPTADESSR	3578-3594	2 pS	1963.75	982.88		+		
		QTTLYLEPKITSTYEVIR	3863-3880	2 pT	2357.24	786.756	+			
		GRIPIVAQNSEEESPLSPVGQPMGMAR	4247-4273	2 pS	3065.66	1022.8953		+		
		GRIPIVAQNSEEESPLSPVGQPMGMAR	4247-4273	3 pS	3145.5	1049.51		+		+
		SEEESPLSPVGQPMGMAR	4256-4273	2 pS	2116.95	1059.48		+		
		EESPLSPVGQPMGMAR	4258-4273	2 pS	1886.9	944.46		+		
		SPLSPVGQPMGMAR	4260-4273	2 pS	1600.65	801.39		+		
		DQFGSSHSLPEVQQHMR	4288-4304	1 pS	2061.86	688.29		+		
Q99569	Plakophilin 4	STSSSPALLGIR	1045-1056	2 pS	1361.67	681.84			+	+
		STSSSPALLGIR	1045-1056	3 pS	1441.62	721.82				+
		TRSPSIDSIQKDPR	527-540	3 pS	1880.89	941.45				+
P70207	Plexin 2	YTGSPDSLR	1617-1625	2 pS	1182.5	592.26				+
Q91XM9	PSD-93/Chapsyn-110	AISLEGEPR	412-420	1 pS	1092.53	547.27			+	
Q62108	PSD-95/SAP90	EQLMNSSLGSGTASLR	409-424	2 pS	1837.86	919.9391				+
		EQLMNSSLGSGTASLRSNPKRGFYIR	409-434	2 pS + 1 pY	3210.67	1071.23		+		
P35486	Pyruvate dehydrogenase E1 component alpha subunit	YHGHSMSDPGVSYR	289-302	1 pS	1671.64	558.22		+		
O35168	Rab3 effector, RIM-1	LQTHDESSLPLPQPSPFMPR	881-900	1 pS	2356.12	786.3815	+			
		SAKSSDSDVSDVSAISR	1259-1275	3 pS	2005.89	1003.95/669.64				+
		SSDSDVSDVSAISR	1262-1275	2 pS	1639.64	820.88		+		
Q8R361	Rab11 interacting protein Rip11	SPSHSSWLSTEGGR	278-291	1 pS	1566.68	784.35		+		
		TYSDEASQLR	305-314	1 pS	1248.50	625.26		+		
Q9QUH6	Ras GTPase-activating protein SynGAP	GYMMRDLNSSIDLQSFMAR	742-760	1 pY	2388.99	797.41	+			
Q8K4S4	Reticulon 1A	SVSEDELIAAIK	350-361	2 pS	1489.81	745.88		+		+
		SQGKGSVSEDELIAAIK	345-361	2 pS	1947.06	974.54		+		
Q9ES28	Rho guanine nucleotide exchange factor 7 (beta-PIX)	KPSDEEFAVR	514-523	1 pS	1256.53	629.27		+		
Q8VD37	RIKEN cDNA 3110007P09	NLSSEEVARPR	167-177	1 pS	1336.60	669.31		+		
		RSTPTPELTSK	178-188	2 pT	1375.56	688.79		+		
		TVPATPPRTGSPLTVATGNDQAATEAK	255-281	2 pS/T	2810.24	937.76		+		
		TGSPLTVATGNDQAATEAK	263-281	1 pS	1910.82	956.42		+		
		TGSPLTVATGNDQAATEAK	263-281	1 pS/T	1911.03	956.52	+			
		IEKLPSISDLDSIFGPVLSPK	282-302	2 pS	2414.16	1208.09	+	+		
		WVHFSDASPEHVTPELTPR	312-330	1 pS	2284.01	1143.01	+	+		
		WVHFSDASPEHVTPELTPR	312-330	2 pS/T	2363.90	1183.00	+	+		
		WVHFSDASPEHVTPELTPR	312-330	3 pS/T	2443.94	815.65		+		
		HVPSPLNLEEVQK	368-380	1 pS	1568.76	785.39		+		
Q8CID5	RIKEN cDNA 5730419I09	SQSESSDEVTELDLSHGK	459-476	2 pS	2190.82	1096.42/731.29			+	
Q8C0E9	RIKEN full-length enriched library, clone:6030438H23	ISNSSEFSAK	41-50	1 pS	1148.48	575.25		+		

Appendix 2

Accession	Protein Name	Phosphopeptide sequences	Residues	Mod. Id	Mr	m/z	Biochemistry			
							PR	DO	PEw	PEi
Q9WV48	Shank1	HYTVGSYDSFDAPSLIDGIDSGSDYIIKEK	633-662	1 pS	3371.51	1124.8461	+			
UPI0000193BD1	Similar to ankyrin 2, neuronal	LGYISVVDTLKVVTEEVTTT	796-815	3 pS/T	2462.3	821.77		+		
		KGSSEESVDEDRGLVPEPLPTAK	1693-1715	3 pS	2777.38	926.8013				+
		NEKLSPVSPSAK	1807-1819	2 pS	1443.61	722.8143		+	+	+
		SPQGLELPLPNR	2355-2366	1 pS	1427.71	714.86			+	
		KVSSSSSESEPELTQLSK	2609-2626	6 pS	2458.01	1230.0142				+
		VQPPSPLPSSIDSNSSPEEATQFQPIVPK	2639-2668	4 pS	3451.57	1151.5311				+
		SIDSNSSPEEATQFQPIVPK	2648-2668	3 pS	2469.02	824.04				+
		SVYSGQDDESPESSPEEQK	3207-3226	3 pS	2435.06	1218.54		+		
		ARSYIETETESR	3351-3362	1 pS	1520.64	761.32/507.89		+		
XP_212807	Similar to Heat shock cognate 71 kDa protein	FSSSTQASIEIDSLYEGIDFYTSITR	273-298	1 pS	3009.16	1004.06	+			
XP_220918	Similar to KIAA1501 protein	TSPSSPTFTFALSR	561-573	2 pS	1671.84	836.93		+		
XP_283703	Similar to SNIP-a protein	IGGKSPPPPPPPPRR	163-177	1 pS	1628.9	543.9752				+
Q9QWI6	SNAP-25-interacting protein	RGSDELTVPR	1052-1061	1 pS	1208.55	605.28		+		
Q91YQ9	Solute carrier family 20, member 1	SSPSESPLMEKK	268-279	2 pS	1520.6	761.31/507.87				+
Q9EPR4	Solute carrier family 23, member 2	SSLAETLDSTGSLDPQR	69-85	2 pS	1991.80	996.91			+	
Q8BUN9	Solute carrier family 24	GGSSASLHNSLMR	332-344	2 pS	1489.68	745.8486				+
Q62261	Spectrin beta chain, brain 1	FSMVRDLMLWMEDVIR	1917-1932	1 pS	2120.01	707.68		+		
		VSEEAESQQWDTSKGDQVSQNGLPAEQGSPR	2109-2139	2 pS	3503.39	1168.80		+		
		TSSKESSPVPSPTLDRK	2158-2174	3 pS	2054.82	1028.42		+		
XP_129018	Spir-1 protein	APTLAELDSSDSEEEK	388-403	3 pS	2057.88	1029.95				+
Q64284	Voltage-gated potassium channel beta-2 subunit	QTGSPGMIYSTR	17-28	1 pS	1376.58	689.29	+	+		
Q8R0S4	Voltage-dependent L-type calcium channel beta-4 subunit	ATHTSSTPMTPLLGR	367-382	2 pS/T	1222.54	612.28		+		
Q8BWS5	Weakly similar to interacting protein 6	SVSTSPSILPAYLK	323-336	2 pS	1635.85	818.9348			+	+

Appendix 2 - Unique phosphopeptide sequences detected in multiple IMAC purifications from synaptosomes by LC-MS/MS analysis. PR (red): Protein IMAC on urea soluble fraction, DO (blue): Double IMAC (protein then peptide) on the urea soluble fraction, PEw (orange): peptide IMAC of whole synaptosomal digest, PEi (green): peptide IMAC of urea insoluble membrane. Unambiguously assigned sites of phosphorylation are denoted in the sequence as S/T/Y (red) and ambiguous sites are just coloured in red. Mod Id: number of phosphates detected.

Appendix 3

Accession	Protein Name	Phosphopeptide sequences	Residues	Mod. Id	MS defined Phospho-sites	Predicted Phospho-sites	Scansite/NetPhosK Predicted cognate Kinase	Scansite Phospho Binding site	Pfam Domain	Domain Residues
Q8BKY9	1110030H18Rik	pSEDDAApSGEHDIQIEGVR	25-42	2 pS	S25		CK2 (0.4471)		-	
					S31		CK1 (0.4804), CK2 (0.5837)		-	
Q8C187	5730405I09Rik protein	pSApSADNLILPR	299-309	2 pS	S299		PKA (0.3877)	14-3-3 mode 1 (0.3712)	-	
					S301		Akt (0.4284)	14-3-3 mode 1 (0.4642)	-	
Q8K4G5	Actin-binding LIM protein 1 isoform a	TLpSPTPSAEGYQDVR	473-487	1 pS	S475 (known)		Cdk5 (0.2710) (PMID: 15302935)		-	
		STSQGSINSPVYSR	494-507	1 pS/T		T495	Cdc2 (0.51) np		-	
Q8R502	AD158	SQpSLKpSIPEK	210-219	2 pS	S212		PKA (0.1632), CaMKII (0.2807)	14-3-3 Mode 1 (0.1372)	-	
					S215		PKC epsilon (0.2710)		-	
P17426	Adapter-related protein complex 2 alpha 1- subunit	pTVpSpTPpSPSADLLGLR	650-664	4 pS/T	T650		Cdc2 (0.51) np		-	
					S652		PKC (0.71), Cdc2 (0.53) np		-	
					T653		p38 MAPK (0.1131), Cdk5 (0.2788)		-	
					S655		ERK1 (0.4552)		-	
Q6QR53	Adapter-related protein complex 3 beta 2	AFYGpSEEDEAKGPGSEEAATAALPAR	268-293	1 pS	S272		CK2 (0.3915)		Adaptin_N	35-591
P51830	Adenylate cyclase, type IX	TApSPGpSVSDLAQTVK	608-622	2 pS	S610		Cdk5 (0.60)		-	
					S613		CK1 (0.2218)		-	
P26231	Alpha-1 catenin	pSRTpSVQTEDDQLIAGQSAR	652-670	2 pS	S652		CK1 (0.4936)	PDK1 Binding (0.9408)	Vinculin	20-868
					S655		PKC epsilon (0.4830), PKA (0.5626)		-	
Q01484	Ankyrin 2	SDpSNApSFLR	29-37	2 pS	S31		Cdc2 (0.49), CaMKII (0.44) np		-	
					S34		PKC delta (0.4742)		-	
		SHLVNEVPVLApSPDLLSEVSEMKQDLIK	1456-1483	1 pS	S1467		p38 MAPK (0.0823)		-	
		pSLESpSPVEPKMK	2435-2446	2 pS	S2435		PKA (0.64) np		-	
					S2439		Cdk5 (0.61), GSK3 (0.51)		-	
		KRFpTPEEEMFK	2547-2557	1 pT	T2550		PKA (0.2982)		-	
Q9Z2H5	Band 4.1-like protein 1	pSEAEEGEVRpTPTK	466-478	2 pS/T	S466		CK2 (0.4992)		DUF1168	456-483
					T475		Cdk5 (0.4634), Cdc2 (0.4928)		DUF1168	456-483
		RLPSSPASPpSPKGpTPEK	537-553	2 pS/T	S546		ERK1 (0.5046)		Prog_receptor	389-823
					T550		Cdk5 (0.4463), Cdc2 (0.4799)		Prog_receptor	389-823
		DKpSDpSETEGLVFAR	646-659	2 pS	S648		PKA (0.1672), CK2 (0.1687)		Prog_receptor	389-823
					S650		CK2 (0.64) np	14-3-3 Mode 1 (0.1483)	Prog_receptor	389-823
Q9WV92	Band 4.1-like protein 3	QLEVQQFEDDKLpSQR	84-98	1 pS	S96		PKC (0.79), ATM (0.52)		-	
		GIpSQTNLITTVTPEKK	484-499	1 pS	S486		DNA PK (0.4976)		ERM	317-615
		GISQTNLITTVpTPEKK	484-499	2 pS/T	T495		Cdk5 (0.4257), Cdc2 (0.4917)		ERM	317-615
		TDTAADGETpSATEpSDQEEDAEIK	534-556	2 pS	S543		CK2 (0.5024)		SAB	538-605
					S547		CK2 (0.3824)		SAB	538-605
		APMIEPLVPEEpTKQpSpSGEKLMDGSEILSLLESAR	612-645	3 pS/T	T623		CK2 (0.52) np		-	
					S626		CK1 (0.2153)		-	
					S627		CaMKII (0.2624)		-	
O88737	Bassoon	MGNEApSLEGGAGEGPLPPGGSGLGPGPGAGKPPSALAGGGQLPVAGAAR	1-49	1 pS	S6		CK2 (0.59) np		-	
		SKpTPPpSNLpSPIEDApSPTEELR	1100-1120	4 pS/T	T1102		Akt (0.4673)	14-3-3 Mode 1 (0.2440)	-	
					S1105		Cdc2 (0.53) np		-	
					S1108		p38 MAPK (0.1282)		-	
					S1114		p38 MAPK (0.1326)		-	
		SPSTSSTIHSYGQPPTTANYGSQTEELPHAPSGPPGSGR	1414-1452	2 pS/T/Y		S1419	CK1 (0.4653)		-	
						Y1424	Abl kinase (0.6024)	Abl SH2 (0.4169)	Alpha_TIF	1423-1448
		pSpSPPIpSPSTPSESPTFpSPGKLGPR	1481-1504	4 pS	S1481		Cdc2 (0.56) np		-	
					S1482		p38 MAPK (0.54), GSK3 (0.52) np		-	
					S1486		p38 MAPK (0.0915)		-	
					S1497		Cdk5 (0.75), CK1 (0.52) np		RNA_pol_Rpbl_R	1487-1501
		SFpSPpTFSPGKLGPR	1491-1504	2 pS/T	S1493		p38 MAPK (0.1203)		RNA_pol_Rpbl_R	1487-1501
					S1495		GSK3 (0.46) np		RNA_pol_Rpbl_R	1487-1501
		ATAEFpSTQTPSLTLSSDIPR	1505-1524	2 pS/T	S1510		Cdc2 (0.49), CaMKII (0.43) np		-	
						T1511	DNAPK (0.57) np		-	
		LVWQQpSpSQEAPIMVITLASDASSQTR	1548-1573	2 pS	S1553		CK2 (0.50) np		-	
					S1554		DNAPK (0.64), ATM (0.60) np		-	
		HSYpSLGFADGR	2026-2036	1 pS	S2029		PKC epsilon (0.4149)		-	
		HGpSGpSGGPDLVQYQPQHGPGLSAPQGLAPLR	2120-2150	2 pS	S2122		PKA (0.67) np		-	
					S2124		PKC (0.56)		-	
		SPVpSPLpSPHIR	2805-2813	2 pS	S2808		p38 MAPK (0.0829)		RNA_pol_Rpbl_R	2797-2812
					S2811		p38 MAPK (0.0837)		RNA_pol_Rpbl_R	2797-2812
		LLDTSFApSSER	2815-2825	1 pS	S2822		CaMKII (0.45)		-	
		pSLpSDPKPLpSPTAEESAKER	2858-2876	3 pS	S2858		Cdc2 (0.61) np		-	
					S2860		Rsk (0.53) np	14-3-3 Mode 1 (0.0946)	-	
					S2866		p38 MAPK (0.0895)		-	
		TLPpSPPPEEAHLPLAGQVPSQLYAASLLQR	2905-2934	1 pS	S2908		p38 MAPK (0.1151)		-	
		SSVSQSPAPTYPSDSHYTSLEQNVPR	3212-3237	2 pS		S3213	PKC (0.77) np		-	
						S3215	DNAPK (0.60) np		-	
		ISKFSPIEEAKDVESDLASYPPPTVSSSLTSR	3378-3409	1 pS		S3382	CK2 (0.52)		Phage_Coat_B	3373-3385

Appendix 3

Accession	Protein Name	Phosphopeptide sequences	Residues	Mod. Id	MS defined Phospho-sites	Predicted Phospho-sites	Scansite/NetPhosK Predicted cognate Kinase	Scansite Binding site	Pfam Domain	Domain Residues
NP_878259	Brain-specific protein p25 alpha	AANKpTPPKpSPGDPAR	11-25	2 pS/T	T15		Cdk5 (0.4299)		-	
					S19		ERK1 (0.4868), Cdc2 (0.4870)		-	
		RLpSLESEGANEGATAAPELSALEEAFRR	29-56	1 pS	S31		PKA (0.3242)		-	
P11798	Calcium/calmodulin-dependent protein kinase type II alpha	ESSEpSTNTTIEDEDTKVR	329-346	2 pS		S331	CK1 (0.58), CK2 (0.54) np	14-3-3 mode 1 (0.5875)	Myco_19_kDa	297-339
					S333		CK2 (0.6077)		Myco_19_kDa	297-339
P28652	Calcium/calmodulin-dependent protein kinase type II beta	NFpSVGRQTTAPATMSTAASGTTMGLVEQAK	313-342	1 pS	S315 (known)		Autophosphorylation (PMID: 2162839)	14-3-3 mode 1 (0.1482)	-	
Q9R196	Calcium-activated potassium channel alpha subunit 1	RLEDEQPPpTLpSPK	701-713	2 pS/T	T709		PKC (0.62) np		BK_channel_a	375-1141
					S711		Cdk5 (0.1029), Cdc2 (0.1689)		BK_channel_a	375-1141
		pSpSPDNSPVHGMLR	923-935	2 pS	S923		GSK3 (0.46) np		BK_channel_a	375-1141
					S924		GSK3 (0.51) np		BK_channel_a	375-1141
P35564	Calnexin precursor	QKpSDAEEDGVTGpSQDEEDSKPK	551-572	2 pS	S553 (known)		CK2 (PMID: 9642293)		Nucleoplasmin	431-591
					S563 (known)		CK2 (PMID: 9642293)		Nucleoplasmin	431-591
P10644	cAMP-dependent protein kinase type I-alpha reg chain	TDSREDEISPPPPNPVVK	75-92	2 pS/T		S77	CK2 (0.5089), DNAPK (0.5219)		-	
						S83 (known)	ERK1 (0.4948) (PMID: 10422841)		-	
Q8K1M3	cAMP-dependent protein kinase type II-alpha reg chain	pSDpSEDADLEVPVPSK	75-91	2 pS	S75		CK2 (0.1203), CK1 (0.2079)		-	
					S77		CK2 (0.1390), DNA PK (0.3624)		-	
Q8K070	cDNA sequence BC034054	pSLDEDEPPPSPLAR	108-121	1 pS	S108		PKA (0.4472)		-	
O35927	Delta catenin-2	ALQpSPEHHIDPIYEDR	409-424	1 pS	S412		Cdk5 (0.4327), Cdc2 (0.4932)		-	
		SPpSIDpSIQKDPR	529-540	2 pS	S531		PKA (0.1416), CaMKII (0.2662)		-	
					S534		Cdc2 (0.47), CamKII (0.42) np		-	
Q9WV69	Dematin	STpSPPSPEVWAESR	90-104	1 pS	S92		ERK1 (0.4387)		-	
		KGAEEEEEEEDDDpSEEEIKAIR	213-234	1 pS	S226		CK2 (0.3238)		-	
Q8BU42	DISKS large-associated protein 2/SAPAP2	SGpSYIKAMGDEEpSGEpSD	438-454	3 pS	S440		PKA (0.1411), PKC delta (0.2879)		-	
					S450		CK2 (0.1342)		-	
					S453		CK2 (0.1721), CK1 (0.1846)		-	
AAO89219	DISKS large-associated protein 3/SAPAP3	AGLGHLpSPEGPLSLSEGPSSVGPEGGPGGVGAGGGSSTFPR	52-92	1 pS	S58		Cdk5 (0.54), p38 MAPK (0.56) np		-	
		SGpSYIKAMGDEEpSGDpSD	392-408	3 pS	S394		PKA (0.1411), PKC delta (0.2879)		-	
					S404		CK2 (0.1240)		-	
					S407		CK1 (0.1711), CK2 (0.1834)		-	
		pSPKTpSPKALAR	410-420	2 pS	S410		p38 MAPK (0.1253), Cdk5 (0.1989)		-	
					S414		Cdk5 (0.2209), Cdc2 (0.2854)		-	
P60904	DnaJ homolog subfamily C member 5 (Cysteine string protein)	SLpSTSGESLYHVLGLDK	8-24	1 pS	S10 (known)		Akt (0.4928) (PMID: 14997482)	14-3-3 mode 1 (0.3954)	-	
Q9QXS6	Drebrin	LSpSPVLHR	140-147	1 pS	S142 (known)		Cdk5 (0.5332) (PMID: 15302935)		TolA	131-447
Q9Z188	Dual-specificity tyrosine-phosphorylation regulated kinase 1B	IVQpYIQSR	270-277	1 pY	Y273 (known)		Autophosphorylation		Pkinase	112-392
Q9JL04	Formin 2	SIQTpSPTEEGRIL	720-732	1 pS	S724		Cdk5 (0.53), CK2 (0.52) np		-	
Q8BXF0	G2 protein	pSRQpSLNpSPpSPGETEMDLLVTR	466-486	4 pS	S466		GSK3 (0.48) np		-	
					S469		PKA (0.1510), PKC epsilon (0.2937)		-	
					S472		p38 MAPK (0.0979), ERK1 (0.4316)		-	
					S478		DNAPK (0.3722)		-	
Q8BJY3	Glutamate receptor, metabotropic 5	pSPpSPISTLSHLAGSAGR	135-151	2 pS	S135		p38 MAPK (0.1146), PKA (0.1646)		Drf-FH1	1012-1175 (rat)
					S137		p38 MAPK (0.1198), PKA (0.1737)		Drf-FH1	1012-1175 (rat)
Q8BH66	Guanine nucleotide-binding protein 3	pSSDWpSpSEEEEPVRK	18-31	3 pS	S18		CK2 (0.51) np	PDK1 binding (0.3563)	-	
					S22		CK2 (0.1348), CK1 (0.2205)		-	
					S23		CK2 (0.1182), DNAPK (0.3334)		-	
Q8C3Q5	Hypothetical protein (D430041B17)	SSSLTPGLGGPDSMAPR	47-63	1 pS		S47	Clk2 (0.3294), Akt (0.4122)	14-3-3 Mode 1 (0.1169)	-	
						S49	PKA (0.1535), Akt (0.1775)	14-3-3 Mode 1 (0.0746)	-	
		LVSQEHLLLSSPEALR	227-242	1 pS		S237	p38 MAPK (0.0971)		-	
		QGpTLEQLQFIPGHHLPQHLR	329-348	1 pT	T331		PKA (0.1573), CaMKII (0.2509)		-	
Q7L0D6	Hypothetical protein KIAA0773	DTDAVpSDLpSDGEKEAR	77-92	2 pS	S82		CK1 (0.4974), CK2 (0.5519)		-	
					S85		CK1 (0.4652), CK2 (0.5519)		-	
Q80SW1	IP3R binding protein released with inositol 1,4,5-trisphosphate	SYpTDpSpSDDEVSPR	80-92	3 pS/T	T82		CK2 (0.1713)	PDK1 binding (0.3518)	-	
					S84		CK2 (0.1398)		-	
					S85		CK2 (0.1460), CK1 (0.2307)		-	
P70404	Isocitrate dehydrogenase [NAD] subunit gamma	HpTPDIGGQGTTSQAIQDIIR	362-381	1 pT	T363		Cdc2 (0.3395)		Iso_dh	57-380
XP_194040	Microtubule-associated protein 1A	ELALpSpSPEDLTQDFEELKR	760-778	2 pS	S764		CK2 (0.61) np		Myosyn_tail_1	306-936
					S765		p38 MAPK (0.1090)		Myosyn_tail_1	306-936
		AELEEMEEVHPpSDEEEEETKAESFYQK	894-918	1 pS	S905		CK1 (0.3896)		Myosyn_tail_1	306-936
		TIpSPTpSpSLEEDKGFK	1131-1145	3 pS	S1133		p38 MAPK (0.56), Cdk5 (0.52) np		-	
					S1136		CK2 (0.1826)		-	
					S1137		DNAPK (0.57), CK1 (0.56) np		-	
		EMTLDQKpSPEK	1811-1821	1 pS	S1818		Cdk5 (0.4689)		-	
		ALGLEEpSPEEEGKAR	1838-1852	1 pS	S1844		CK2 (0.1467)		-	
		ETpSPTRGEPVPAWEGKpSPEQEVR	1870-1892	2 pS	S1872		Cdk5 (0.55) np		-	
					S1886		GSK3 (0.49), p38 MAPK (0.47)		-	
		WLAEpSPVGLPPEEEDKLTR	1981-1999	1 pS	S1985		Cdk5 (0.54), CK1 (0.53) np		-	
		SPFEIIpSPPApSPPEMTG	2000-2016	2 pS	S2006		p38 MAPK (0.1017)		-	
					S2010		p38 MAPK (0.0974)		-	
P14873	Microtubule-associated protein 1B	SLMpSpSPEDLTK	825-835	2 pS	S828		CK2 (0.55), CK1 (0.54) np		-	
					S829		Cdk5 (0.57), GSK3 (0.51) np		-	
		DVMSDETNNEETESPSQEFVNITK	1138-1161	1 pS/T		T1151	CK2 (0.4918)		-	

Appendix 3

Accession	Protein Name	Phosphopeptide sequences	Residues	Mod. Id	MS defined Phospho-sites	Predicted Phospho-sites	Scansite/NetPhosK Predicted cognate Kinase	Scansite Binding site	Pfam Domain	Domain Residues
P14873	Microtubule-associated protein 1B	DVSDERLpSPAKSPSLSPSPPpSPIEK	1240-1264	4 pS	S1247		GSK3 (0.4848)		-	
						S1251	GSK3 (0.2969)		-	
					S1260		Cdk5 (0.4900)		-	
		SPSLpSPSPPpSPIEK	1251-1264	2 pS	S1255		ERK1 (0.3455)		-	
		SVSPGVTQAVVEEHCApSPEEK	1291-1311	1 pS	S1307		p38 MAPK (0.59), CKII (0.54) np		-	
		TLEVVpSPSQSVTGSAGHITpYYQpSPTDEK	1312-1339	3 pS + 1 pY	S1317		Cdk5 (0.53), GSK3 (0.50) np		-	
						S1319	DNA PK (0.63)		-	
					Y1331		INSR (0.44), Src (0.40) np		-	
					S1334		Cdk5 (0.62), p38 MAPK		-	
		ApSLpSPMDEPVPDSESPVEK	1370-1388	2 pS	S1371		PKA (0.71)		-	
					S1373		p38 MAPK (0.1227)		-	
		ASLSPMDEPVPDpSEpSPVEK	1370-1388	2 pS	S1382		CK2 (0.55) np		-	
					S1384		p38 MAPK (0.1184)		-	
		VLpSPLRpSPPLLGSESPVEDFLSADSK	1389-1414	2 pS	S1391		p38 MAPK (0.0950)		-	
					S1395		GSK3 (0.4814)		-	
		QGFPDRESPVSDLTSTGLYQDKQEEK	1431-1456	1 pS/T		S1438	Cdk5 (0.53), p38 MAPK (0.52) np		-	
		SIpSPPDFpSPKTAK	1614-1626	2 pS	S1616		Cdk5 (0.57) np		-	
					S1621		Cdk5 (0.63) np		-	
		VQpSLEGEKLpSPK	1766-1777	2 pS	S1768		Cdc2 (0.47), CK2 (0.45)		-	
					S1775		Cdk5 (0.2372), Cdc2 (0.2949)		-	
		VQSLEGEKLpSPKpSDIpSPLpTPR	1766-1786	4 pS/T	S1778		PKA (0.51) np		-	
					S1781		p38 MAPK (0.1178)		-	
					T1784		p38 MAPK (0.0874)		-	
		ESSPLYpSPGFSDSTSAAK	1787-1804	2 pS	S1793		Cdk5 (0.67), GSK3 (0.54) np		-	
		FpTAAAHQASSSPPIDAATAEPYGFR	1805-1829	1 pT	T1806		PKG (0.54) np		-	
		TPGDFNYAYQKPENAAGpSPDEEDYDYESQEK	1860-1890	1 pS	S1877		CK2 (0.4574)		MAP1B neuraxin	1875-1891
		TIKpSPCDSGYSYETIEK	1908-1924	1 pS	S1911		GSK (0.53), p38 MAPK (0.53) np		MAP1B neuraxin	1909-1925
		TTRTPEEGGYSYEISEK	1942-1958	1 pT		T1945	p38 MAPK (0.1331)		MAP1B neuraxin	1943-1959
		TTRpSPDTSAYCVETMEK	2027-2043	1 pS	S2030		p38 MAPK (0.51) np		MAP1B neuraxin	2028-2044
		TELpSPSFINPNPLEWFAGEEPTEESEKPLTQSGGAPPPSGGK	2091-2132	1 pS	S2094		GSK3 (0.49), p38 MAPK (0.47) np		-	
P20357	Microtubule-associated protein 2	GHDLSPLASDILTNTSGSMDEGDDpYLPPTTPAVEK	722-756	1 pY	Y746		EGFR (0.58), INSR (0.53) np	Itk SH2 (0.1727)	-	
		ETpSPETSLIQDEVALK	1159-1174	1 pS	S1161		CK1 (0.53) np		Myosyn tail 1	895-1585
		DGpSPDAPApTPEKEEVAFSEYK	1350-1370	2 pS/T	S1352		Cdk5 (0.54) np		Myosyn tail 1	895-1585
					T1358		p38 MAPK (0.61), Cdk5 (0.54) np		Myosyn tail 1	895-1585
		DKVTDGISKpSPEKR	1530-1543	1 pS	S1539		Cdk5 (0.3127), Cdc2 (0.3441)		Myosyn tail 1	895-1585
		VDHGAEITQpSPSR	1773-1786	1 pS	S1783		p38 MAPK (0.1244)		-	
		RLSNVSSSGSINLLESPQLATLAEDVTAALAK	1794-1825	3 pS		S1799	CK1 (0.58) np		-	
						S1801	Cdc2 (0.61) np		-	
						S1803	CaMKII (0.46), GSK3 (0.45) np		-	
P27546	Microtubule-associated protein 4	AAVGVTGNDITTPPNKEPPPpSPEKK	647-671	2 pS/T		T658	Cdk5 (0.5304)		-	
					S667 (known)		Cdc2 (PMID 9398320)		-	
P10637	Microtubule-associated protein Tau	SGERSGYpSpSPGpSPG	482-495	3 pS	S489 (known)		Cdk5 (PMID 14642273)		-	
					S490 (known)		Cdk5 (PMID 14642273)		-	
					S493 (known)		Cdk5, PKA (PMID 14642273, 10460255)		-	
		AKTDHGAEIVYKpSPVVpSGDTpSPR	675-697	3 pS	S687 (known)		GSK3/Cdk5 (PMID 8245032)		-	
					S691 (known)		GSK3/Cdk5 (PMID 8245032)		-	
					S695 (known)		PKC/Cdk5/Cdc2 (PMID 8245032)		-	
		SPVVpSGDpTpSPRHL	687-699	3 pS/T	T694 (known)		PKC/Cdk5/Cdc2 (PMID 14642273)		-	
		SPVVSGDpTpSPRHLSNVpSSTG	687-706	3 pS/T	S703		CK1 (0.57) np		-	
		HLSNVSSpTGSIDMVDSPQLATLADEVSASLAK	698-729	3 pS/T		S704 (known)			-	
						S700 (known)	PKA (PMID 10460255)		-	
					T705		Cdc2 (0.56) np		-	
BAC65687	mKIAA0942 protein	LEWAVDDEEKKKpSPSEGTDEK	12-32	1 pS	S24		PKA (0.4201)		-	
		IGpSTTNPFLDIPHDPNAAVYK	44-64	1 pS	S46		PKC delta (0.3156)		-	
		DEYKPEKpSLSDEDLK	105-119	1 pS	S112		PKC epsilon (0.3460)		PH	63-165
		pSHpSpSPSLNPDASPVTAKVKRNVSER	237-261	3 pS	S237		Cdc2 (0.48) np		-	
					S239		PKA (0.53) np	14-3-3 Mode 1 (0.0952)	-	
					S240		Cdk5 (0.2560), Cdc2 (0.3207)		-	
		SHSpSPpSLNPDApSPVTAK	237-253	3 pS	S242		DNAPK (0.45) np		-	
					S248		Cdc2 (0.3394)		Gemini coat	245-262
P53986	Monocarboxylate transporter 1 (SLC16A1)	LKpSKFpSLQEAGK	208-219	2 pS	S210		ATM (0.4358)		7tm 5	78-348
					S213		PKC epsilon (0.1783), PKC delta (0.2689)		7tm 5	78-348
O08599	Munc-18-1	RAHVFFpTDSCPDALFNELVK	101-120	1 pT	T107		Cdc2 (0.57) np		Sec1	29-583
		pSQVpYpSLDSADSFQSFYSPHK	142-161	2 pS + 1 pY	S142		DNAPK (0.3330), GSK3 (0.7626)		Sec1	29-583
					Y145		SRC (0.43), INSR (0.43)		Sec1	29-583
					S146		Akt (0.4730)		Sec1	29-583
		DPSSPVLHELTFQAMSYDLLPIENDVYK	238-265	2 pS/T		S241	p38 MAPK (0.1040), Cdk5 (0.2318)		Sec1	29-583
						T248	PKC mu (0.2889)		Sec1	29-583
		pSpTHLHLAEDCMK	345-356	2 pS/T	S345		PKA (0.46) np		Sec1	29-583
					T346		CaMKII (0.47) np		Sec1	29-583
		EDSEIITNMAHLGVPIVTDSTLR	435-457	1 pS/T		S437	CK2 (0.47)		Sec1	29-583
		SSASFSTTAVSAR	506-518	1 pS		S507	PKC epsilon (0.4112)		Sec1	29-583

Appendix 3

Accession	Protein Name	Phosphopeptide sequences	Residues	Mod. Id	MS defined Phospho-sites	Predicted Phospho-sites	Scansite/NetPhosK Predicted cognate Kinase	Scansite Binding site	Pfam Domain	Domain Residues
P04370	Myelin basic protein (generic)	pSKYLATASTMDHAR	144-157	1 pS	S144 (known)		PKA/PKC (PMID:2413024)		-	
		TTHYGpSLPQK	196-205	1 pS	S201		PKC a, b, g (0.5004), PKC delta (0.5180)		Myelin_MBP	149-250
		NIVTPRpTPPPSQGK	223-236	1 pT	T229 (known)		ERK2		Myelin_MBP	149-250
P04370-6, 8, 11	Myelin basic protein (isoforms 6,8,11)	FSWGGRDpSR	138-146	1 pS	S145		GSK3 (0.50), CK1 (0.49) np		Myelin_MBP	16-154
P04370-4, 5, 9, 10	Myelin basic protein isoforms (isoforms 4, 5, 7, 9, 10)	FpSWGAEGQKPGFGYGGR	138-154	1pS	S139 (known)		PKA (0.1514)	PDK1 binding (0.2781)	DUF1199	132-148
P02688	Myelin basic protein S	PQKSQRTQDENPVVHFFK	97-114	1 pS/T		T103	DNAPK (0.2828), CaMKII (0.2926)		Myelin_MBP	17-194
O35413	nArgBP2	RGpSLPDNSILHR	842-853	1 pS	S844		PKA (0.0687)	14-3-3 Mode 1 (0.0849)	-	
P13595	Neural cell adhesion molecule 1	AAFSKDEpSKEPIVEVR	767-782	1 pS	S774		CK2 (0.48), GSK3 (0.46) np		Mucin	730-915
Q810U3	Neurofascin precursor	DKEETEGNEpSpSEATSPVNAIV	1217-1237	3 pS/T	S1226		CK2 (0.1854)		-	
					S1227		CK2 (0.1839)		-	
						S1231	p38 MAPK (0.1295)		-	
P08553	Neurofilament triplet M protein	EEAEEKEEEPEAEKpSPVKpSPEAK	488-510	2 pS	S502 (known)		GSK-3 alpha (PMID: 7891113)		Myosyn_tail_1	114-840
					S506 (known)		GSK-3 alpha (PMID: 7891113)		Myosyn_tail_1	114-840
Q9Z0P4	Paralemmin	SETLVNAQQpTPLGpTPK	132-147	2 pT	T141		p38 MAPK (0.1077), ERK1 (0.4492)		Paralemmin	71-384
					T145		p38 MAPK (0.0960), Cdk5 (0.2001)		Paralemmin	71-384
		TpSTPVRpSPGGSTMMK	151-165	2 pS	S152		GSK3 (0.47) np		Paralemmin	71-384
					S157		p38 MAPK (0.1012), Cdk5 (0.2729)		Paralemmin	71-384
Q9Z1B3	Phospholipase C-beta-1	pTpYLpTAGQLAGNSSVEMYR	333-350	2 pT+1 pY	T333		CaMKII (0.48)		-	
					Y334		INSR (0.47) np		-	
					T336		GSK3 (0.44), CamKII (0.42) np		-	
Q9QYX7	Piccolo	SSPSSPSDLAK	1307-1317	2 pS		S1308	Cdk5 (0.64), GSK3 (0.51) np		Myosyn_tail_1	1187-2015
						S1311	p38 MAPK (0.1037)		Myosyn_tail_1	1187-2015
		EHKENPELVDDLpSPR	1397-1411	1 pS	S1409		p38 MAPK (0.53)		Myosyn_tail_1	1187-2015
		HSWHDEDDETFDEpSPELK	1532-1549	1 pS	S1545		p38 MAPK (0.55)		Myosyn_tail_1	1187-2015
		pTPPSNLpSPIEDApSPTEELRQAAEMEELHR	1730-1758	3 pS/T	T1730		PKB (0.87), Cdk5 (0.65)	14-3-3 mode 1 (0.3591)	DUF1388	1727-1754
					S1736		p38 MAPK (0.1282)		DUF1388	1727-1754
					S1742		p38 MAPK (0.1326)		DUF1388	1727-1754
		VYKLPTAVSLYpSPTDEQSVMQK	1788-1809	1 pS	S1799		p38 MAPK (0.0945)		Myosyn_tail_1	1187-2015
		SMpSDPKPLpSPTADE	3578-3590	2 pS	S3580		RSK (0.50) np	14-3-3 mode 1 (0.0907)	-	
					S3586		p38 MAPK (0.0933)		-	
		QpTpTLYLEPKITSTYEVIR	3863-3880	2pT	T3864		PKA (0.59) np		-	
					T3865		PKA (0.62) np		-	
		GRIPIVAQNpSEEEpSPLpSPVGQPMGMAR	4247-4273	3 pS	S4256		CK2 (0.65), CK1 (0.52) np		-	
					S4260		Cdk5 (0.55), GSK3 (0.51)		-	
					S4263		p38 MAPK (0.0917)		-	
		DQFGSSIIpSLPEVQQHMR	4288-4304	1 pS	S4295		Cdc2 (0.56), DNAPK (0.51) np		-	
Q99569	Plakophilin 4	TRpSPpSIDpSIQKDPR	508-521	3 pS	S510		CK1 (0.4501)		-	
					S512		CaMKII (0.4997), PKA (0.5056)		-	
					S515		CaMKII (0.45) np		-	
		STNpSpSPALLGIR	1045-1056	3 pS/T		S1047	Cdc2 (0.49), GSK3 (0.43) np		-	
					S1048		CaMKII (0.45), GSK3 (0.43)		-	
					S1049		p38 MAPK (0.1323)		-	
P70207	Plexin 2	YTGpSPDpSLR	1617-1625	2 pS	S1620		p38 MAPK (0.1223)		-	
					S1623		PKC (0.63) np		-	
Q91XM9	PSD-93/Chapsyn-110	AIpSLEGEPR	412-420	1 pS	S414		CK1 (0.59)		-	
Q62108	PSD-95/SAP90	EQLMNSpSLGpSGTASLRSNPKRGFpYIR	409-434	2 pS+1 pY	S415		PKA (0.78)		-	
					S418		PKC (0.80)		-	
					Y432		EGFR (0.55)		SH3_1	432-495
P35486	Pyruvate dehydrogenase E1 component alpha subunit	YHGHpSMSDPGVSYR	289-302	1 pS	S293 (known)		PDK (PMID: 11486000)	PDK1 binding (0.9566)	Transketolase_N	111-348
Q9JIR4	Rab3 effector, RIM-1	LQTHDESSLPLPQPpSPFMPR	881-900	1 pS	S895		Cdk5 (0.69)		-	
		SAKpSpSDpSDVSDVSAISR	1259-1275	3 pS	S1262		CK2 (0.1866)		-	
					S1263		CK2 (0.1966)		-	
					S1265		CaMKII (0.43), GSK (0.43) np		-	
Q8R361	Rab11 interacting protein Rip11	SPSHSSWLSTEGGR	278-291	1 pS		S280	Clk2 (0.7001)	14-3-3 mode 1 (0.4153)	-	
		TVpSDEASQLR	305-314	1 pS	S307		PKA (0.47), CaMKII (0.46) np		Bunya_G1	307-317
Q9QUH6	Ras GTPase-activating protein SynGAP	GpYMMRDLNSSIDLQSFMAR	742-760	1 pY	Y743		INSR (0.45), EGFR (0.42) np		-	
Q8K4S4	Reticulon 1A	pSVpSEDELIAAIK	350-361	2 pS	S350		PKA (0.1789), PKC epsilon (0.2617)		REV	286-251
					S352		DNAPK (0.3254)		-	
Q9ES28	Rho guanine nucleotide exchange factor 7 (beta-PIX)	KPpSDEEFAVR	514-523	1 pS	S516		PKA (0.4424)		-	
Q8VD37	RIKEN cDNA 3110007P09	NLpSSEEVARPR	167-177	1 pS	S169		RSK (0.54), PKA (0.51) np		-	
		RSpTPpTPELTSK	178-188	2 pT	T180		PKA (0.3958), Akt (0.5296)		-	
					T182		p38 MAPK (0.52), Cdk5 (0.51) np		-	
		TVPApTPPRpTGpSPLTVATGNDQAATEAK	255-281	3 pS/T	T259		p38 MAPK (0.1231), Cdk5 (0.2592) np		-	
					T263		GSK3 (0.47), CaMKII (0.43) np		-	
					S265		PKA (0.71), Cdk5 (0.62) np		-	
		IEKLPpSISDLDSIFGPVLpSPK	282-302	2 pS	S287		PKC epsilon (0.4601)		-	
					S300		p38 MAPK (0.55) np		-	
		WVHFpSDApSPEHVpTPELTPR	312-330	3 pS/T	S316		Cdc2 (0.50) np		-	
					S319		Cdk5 (0.2551)		-	
					T324		Cdk5 (0.2552)		-	

Appendix 3

Accession	Protein Name	Phosphopeptide sequences	Residues	Mod. Id	MS defined Phospho-sites	Predicted Phospho-sites	Scansite/NetPhosK Predicted cognate Kinase	Scansite Binding site	Pfam Domain	Domain Residues
Q8VD37	RIKEN cDNA 3110007P09	HVPpSPLNLEEVQK	368-380	1 pS	S371		Cdk5 (0.58), GSK3 (0.51) np		-	
Q8CID5	RIKEN cDNA 5730419I09	SQpSFSSDEVTELDLSHGK	459-476	2 pS	S461		Clk2 (0.6117)	14-3-3 mode 1 (0.2898)	-	
						S464	CK2 (0.5506)		-	
Q8C0E9	RIKEN full-length enriched library, clone:6030438I23	ISNpSSEFSAK	41-50	1 pS	S44		GSK3 (0.45), CK2 (0.44)		-	
Q9WV48	Shank1	HYTVGSVDpSFDAPSLIDGIDSGSDYIIKEK	633-662	1 pS	S641		CaMKII (0.46) np		-	
UPI0000193BD1	Similar to ankyrin 2, neuronal	LGYIpSVVDTLKVVTEEVTpTpT	796-815	3 pS/T	S800		CaMKII (0.44), GSK3 (0.440 np		Ank	786-818
					T814		CK1 (0.52) np		Ank	786-819
					T815		CK1 (0.55) np		Ank	786-820
		KGpSpSEEpSVDEDRGLVPEPLPTAK	1693-1715	3 pS	S1695		PKA (0.1484)		Chlorion_3	1598-18-42
					S1696		CK2 (0.55) np		Chlorion_3	1598-18-42
					S1699		CK2 (0.1338)		Chlorion_3	1598-18-42
		NEKIpSPVpSPSAK	1808-1819	2 pS	S1812		p38 MAPK (0.1322)		Chlorion_3	1598-18-42
					S1815		p38 MAPK (0.0957)		Chlorion_3	1598-18-42
		pSPQGLELPLPNR	2355-2366	1 pS	S2355		p38 MAPK (0.57)		-	
		KVpSpSpSpSpSEpSEPELTQLSK	2609-2626	6 pS	S2611		PKA (0.1434)		DUF572	2582-2659
					S2612		PKC (0.56) np		DUF572	2582-2659
					S2613		CK2 (0.62) np		DUF572	2582-2659
					S2614		CK1 (0.58), CK2 (0.51) np		DUF572	2582-2659
					S2615		CK2 (0.66) np		DUF572	2582-2659
					S2617		CK2 (0.64) np		DUF572	2582-2659
		VQPPSPLPSpSIDSNSSPEEATQFQPIVPK	2640-2668	4 pS		S2648	GSK3 (0.3317)		DUF572	2582-2659
					S2649		ATM (0.46), GSK3 (0.45) np		DUF572	2582-2659
						S2654	CK2 (0.1523)		DUF572	2582-2659
						S2655	p38 MAPK (0.1267)		DUF572	2582-2659
		SVYSGQDDEpSPEpSpSPEEQK	3208-3226	3 pS	S3217		CK2 (0.4404)		-	
					S3220		CK2 (0.4037)		-	
					S3221		CK1 (0.58), CK2 (0.58) np		-	
		ARpSYIETETESR	3351-3362	1 pS	S3353		CK2 (0.51), PKG (0.50) np		-	
XP_212807	Similar to Heat shock cognate 71 kDa protein	FSSSTQASIEIDSLYEGIDFYTSITR	273-298	1 pS		S280	CK1 (0.4348)		Hydantoinase_A	130-384
XP_220918	Similar to KIAA1531 protein	TSPpSpSPTFTFALSR	560-573	2 pS	S563		Cdc2 (0.49), GSK3 (0.45) np		-	
					S564		p38 MAPK (0.1074), Cdk5 (0.2754)		-	
XP_283703	Similar to SNIP-a protein	IGGKpSPPPPPPPRR	163-177	1 pS	S167		p38 MAPK (0.1311), Cdk5 (0.2403)		Drf_FH1	165-319
Q9QWI6	SNAP-25-interacting protein	RQpSDELTVPR	1052-1061	1 pS	S1054		PKA (0.3743)		-	
Q91YQ9	Solute carrier family 20, member 1	SpSPSEpSPLMEKK	268-279	2 pS	S269		ERK1 (0.4993), GSK3 (0.7197)		Na_H_Exchanger	4-358
					S273		Cdk5 (0.2501), Cdc2 (0.3049)		Na_H_Exchanger	4-358
Q9EPR4	Solute carrier family 23, member 2	SSLAETLDpSTGSLDPQR	69-85	2 pS		S70	PKC epsilon		EBH_binding	71-96
					S77		GSK3 (0.47), Cdc2 (0.46) np		EBH_binding	71-96
Q8BUN9	Solute carrier family 24	GQpSSApSLHNSLMR	332-344	2 pS	S334		Clk2 (0.4136)		-	
					S337		CK1 (0.2300)		-	
Q62261	Spectrin beta chain, brain 1	FpSMVRDLMLWMEDVIR	1917-1932	1 pS	S1918		PKA (0.57)		Spectrin	1912-2016
		VSEEAEpSQQWDTSKGDQVSQNGLPAEQGpSPR	2109-2139	2 pS	S2115		CK2 (0.5776)		DUF776	2114-2138
					S2137 (known)		Cdk5 (0.5389), (PMID: 15302935)		DUF776	2114-2138
		TSSKESSPVPpSPTLDRK	2158-2174	3 pS/T		S2159	PKC (0.78) np		-	
						S2164 (known)	ERK1 (0.5433) (PMID: 15302935)		-	
					S2168 (known)		GSK3 (0.5024), (PMID: 15302935)		-	
XP_129018	Spir-1 protein	APTLAELDpSpSDpSEEEK	388-403	3 pS	S396		CK2 (0.1254)		-	
					S397		CK2 (0.1167)		-	
					S399		CK2 (0.1668)		-	
Q64284	Voltage-gated potassium channel beta-2 subunit	QTGpSPGMIYSTR	17-28	1 pS	S20		p38 MAPK (0.1249), Cdk5 (0.2650)		-	
Q8R0S4	Voltage-dependent L-type calcium channel beta-4 subunit	ATHTpSSSTPMpTPLLGR	367-382	2 pS/T	S371		CK1 (0.56), GSK3 (0.48) np		Ca_channel_B	182-406
					T377		ERK1 (0.5229)		Ca_channel_B	182-406
Q8BWS5	Weakly similar to interacting protein 6	SVSTpSPSILPAVLK	323-336	2 pS		S325	PKA (0.51) np	14-3-3 mode 1 (0.1088)	-	
					S327		ERK1 (0.5791)		-	

Appendix 3 - Non-redundant phosphopeptide list from the synapse phosphoproteome with unique phosphorylation site assignation. Unambiguously assigned sites of phosphorylation are denoted in the sequence as pS/T/Y (red) and ambiguous sites are just coloured in red. Mod Id: number of phosphates detected. MS identified Phospho-sites: unambiguous identification of sites by MS. Predicted Phospho-sites: when sites could not be unambiguously assigned, Netphos and Scansite were used to predict the most probable site of phosphorylation and are indicated in the sequence in red. Predicted cognate kinase: Kinases predicted by Scansite and NetphosK are indicated for all sites. Scansite was used as the default program to predict potential kinases for phosphorylation sites. However, NetphosK was used in cases which Scansite did not produce any predictions and its use is denoted by "np" in the Predicted cognate kinase field. Prediction scores are included for both Scansite (0 = best prediction) and NetphosK (1 = best prediction). Also, in some cases Scansite predicted some phosphorylation sites to be involved in phosphorylation dependent protein interactions. Pfam Domains: when phosphorylation sites were found in a Pfam protein domain, the domain name and location in the protein sequence is indicated.

Appendix 4

IPI	Uniprot	Protein Name	Phosphopeptide Sequences	Mod ID	Residues	m/z	Mr (expt)
IPI00123624	Q8VDP4	2610301G19Rik protein	SVASNQSEMEYSSLQDMPK	2 pS	674-692	1145.89	2289.77
IPI00123624	Q8VDP4	2610301G19Rik protein	SVASNQSEMEYSSLQDMPK	3 pS	674-692	1185.89	2369.77
IPI00110523	Q9CXN5	3110050K21Rik protein	SKLSPSPSLR	2 pS	204-213	616.26	1230.51
IPI00021840	P62754	40S ribosomal protein S6 (RPS6)	RLSSLR	1 pS	233-238	406.12	810.22
IPI00177047	Q8CHU0	9430041O17Rik (ArgBP2)	KPLSVSPSTDGLPSPSPPPR	2 pS	179-198	745.67	2233.98
IPI00121136	Q9JIX8	Acin1	SLSPLSGTTDTKAESPAGR	2 pS	477-495	678.94/1017.92	2033.82
IPI00108780	P17426	Adapter-related protein complex 2 alpha 1 subunit	DTSSNDINGGVEPTPSTVSTPSPSADLLGLR	1 pS + 2 pS/T	634-664	1108.78	3323.30
IPI00202484	P62996	Arginine/serine-rich splicing factor 10 (Sfrs10)	RPSPSPYYSR	2 pS	262-271	476.86	1427.55
IPI00202484	P62996	Arginine/serine-rich splicing factor 10 (Sfrs10)	RRSPSPYYR	1 pY	262-269		
IPI00202484	P62996	Arginine/serine-rich splicing factor 10 (Sfrs10)	RRHSHSHSPMSTR	3 pS	92-104	605.85	1814.52
IPI00117229	O70305	Ataxin-2	TNSPSISPSMLSNAEHK	2 pS	830-846	653.91	1958.71
IPI00169477	Q8K019	Bcl-2-associated transcription factor 1 [Bclaf1]	NTPSQHSHSIQHSPER	2 pS	255-270	667.92	2000.73
IPI00169477	Q8K019	Bcl-2-associated transcription factor 1 [Bclaf1]	YSPSQNSPIHHIPSR	1 pS + 1 pS/Y	283-297	627.24/940.37	1878.71
IPI00169477	Q8K019	Bcl-2-associated transcription factor 1 [Bclaf1]	IDISPSALR	1 pS	653-661	526.24	1050.47
IPI00311962	Q9JLE4	Beta adducin delta	AGTKSPAVSPSKTSEDTKK	2 pS	610-628	693.62	2077.84
IPI00311962	Q9JLE4	Beta adducin delta	KLEQEQEGEKDIATEKPGSPVKSTPASPVQSPSK	3 pS + 1 pT	576-609	982.39	3925.53
IPI00360155	Q6DMN7	BR serine/threonine-protein kinase 2 (BRSK2)	SMEVLSVTDGGSPVPAR	2 pS	383-399	931.37	1860.72
IPI00360155	Q6DMN7	BR serine/threonine-protein kinase 2 (BRSK2)	SMEVLSVTDGGSPVPARR	2 pS	383-400	673.29/1009.41	2016.81
IPI00360155	Q6DMN7	BR serine/threonine-protein kinase 2 (BRSK2)	SRSISGASSGLSTSPLSSPR	1 pS + 1 pS	411-430	698.62	2092.82
IPI00326171	Q68EG2	CaMKII beta	GSLPPAALESSDSTNTTIEDEDAK	2 pS/T	345-368	870.01	2607.00
IPI00296678	Q5SQZ3	CaMKII gamma	RKSSSSVHLMEPQTTVVHNATDGIK	3 pS	353-377	988.09	2961.24
IPI00296678	Q5SQZ3	CaMKII gamma	SSSSVHLMEPQTTVVHNATDGIK	2 pS	355-377	866.69	2597.04
IPI00357503	Q7TQF8	Camsap1	QDSDSDVVDVEDTEQDFIGEDHPVVIPR	2 pS	809-820	1105.78	3314.31
IPI00385480	Q6P9K8	Caskin-1	YAASDSEPERDELLVPAAAGPYATVQR	2 pS	888-914	1012.77	3035.29
IPI00282957	Q8CIL3	CDNA sequence BC019977	SRGPTPTATGPR	1 pS + 1 pT	95-106	679.27	1356.52
IPI00282957	Q8CIL3	CDNA sequence BC019977	RSSQPSPTTVPASDSPPAK	2 pS	113-131	690.61/1035.42	2068.83
IPI00282957	Q8CIL3	CDNA sequence BC019977	RSSQPSPTTVPASDSPPAKQDVK	2 pS	113-135	847.36	2539.04
IPI00282957	Q8CIL3	CDNA sequence BC019977	RSSQPSPTTVPASDSPPAKQDVK	2 pS + 1 pS/T	113-135	874.02	2619.03
IPI00282957	Q8CIL3	CDNA sequence BC019977	AAEEKEPAAPASPAPSPVPSPTPAQPQK	3 pS	533-560	997.42	2989.23
IPI00261366	Q80UZ7	D430018P08 protein	KSHSSPSLNPDASPVTAK	1 pS + 1 pT	175-192	661.60	1981.79
IPI00261366	Q80UZ7	D430018P08 protein	KSHSSPSLNPDASPVTAK	3 pS	175-192	688.26	2061.75
IPI00261366	Q80UZ7	D430018P08 protein	SHSSPSLNPDASPVTAK	2 pS	176-192	618.90/927.86	1853.70
IPI00261366	Q80UZ7	D430018P08 protein	SHSSPSLNPDASPVTAK	3 pS	176-192	645.56/967.84	1933.67
IPI00190337	Q9WV69	Dematin	STSPPPSPEVWAESR	1 pS + 1 pS/T	90-104	893.84	1785.67
IPI00190337	Q9WV69	Dematin	HLSAEDFSR	1 pS	370-378	571.22	1140.42

Appendix 4

IPI	Uniprot	Protein Name	Phosphopeptide Sequences	Mod ID	Residues	m/z	Mr (expt)
IPI00327311	P97427	Dihydropyrimidinase related protein-1	SIPHITSDR	1 pS	8-16	553.24	1104.46
IPI00327311	P97427	Dihydropyrimidinase related protein-1	HAAPAPSAKSSPSKHQPPPIR	2 pS	512-532	774.34	2320.01
IPI00327311	P97427	Dihydropyrimidinase related protein-1	HAAPAPSAKSSPSKHQPPPIR	3 pS	512-532	801.01	2400.01
IPI00257508	O08553	Dihydropyrimidinase related protein-2	GLYDGPVCEVSVTPK	1 pT	497-511	822.34	1642.66
IPI00257508	O08553	Dihydropyrimidinase related protein-2	TVTPASSAKTSPAKQQAPPVR	3 pS + 2 pT	512-532	841.31	2520.92
IPI00257508	O08553	Dihydropyrimidinase related protein-2	GLYDGPVCEVSVTPKTVTPASSAKTSPAKQQAPPVR	7 pS/T	497-532	1057.43	4225.68
IPI00391486	Q7Z417	FMRP interacting protein	NDSWGSFDLR	1 pS	650-659	638.74	1275.46
IPI00009071	Q9R0U0	FUS interacting serine-arginine rich protein 1	SRSRSFDYNYR	3 pS	129-139	564.18	1689.51
IPI00009071	Q9R0U0	FUS interacting serine-arginine rich protein 1	SRSFDYNYR	2 pS	131-139	684.22	1366.43
IPI00009071	Q9R0U0	FUS interacting serine-arginine rich protein 1	SRSFDYNYRR	2 pS	131-140	508.52	1522.53
IPI00009071	Q9R0U0	FUS interacting serine-arginine rich protein 1	SFDYNYR	1 pS	133-139	522.67	1043.33
IPI00120424	Q68FF6	G protein-coupled receptor kinase-interactor 1	SLSSPTDNLELSAR	1 pS	368-381	785.33	1568.65
IPI00120424	Q68FF6	G protein-coupled receptor kinase-interactor 1	SQSELDDQHDYDSVASDEDTDQEPLPSAGATR	2 pS	382-413	1213.46	3637.35
IPI00120424	Q68FF6	G protein-coupled receptor kinase-interactor 1	SQSELDDQHDYDSVASDEDTDQEPLPSAGATR	2 pS + 1 pT	382-413	1240.09	3717.25
IPI00120424	Q68FF6	G protein-coupled receptor kinase-interactor 1	HGSGADSDYENTQSGDPLLGLEGKR	1 pS + 1 pS/T	599-623	921.69	2762.04
IPI00021595	Q80YT1	Glucocorticoid induced transcript 1 (Glcci1)	SASWGSADQLK	1 pS	25-35	615.24	1228.46
IPI00292228	Q9WV60	Glycogen synthase kinase-3 beta	GEPNVSYICSR	1 pY	210-220	652.75	1303.50
IPI00377522	Q6NXJ5	Gm1568	SPSLGDLSHSPQTAQHVER	2 pS	301-319	752.63	2254.87
IPI00137459	Q6PG39	HSHIN1 protein	LQRPKEESSEDENEVSNILR	2 pS	304-322	844.67	2530.98
IPI00006669	Q96D85	IQSEC1	MQFSFEGPEK	1 pS	55-64	640.23	1278.45
IPI00006669	Q96D85	IQSEC1	NSWDSPAFSNDVIR	1 pS	389-402	844.32	1686.63
IPI00298920	Q9UQ09	KARP-1-binding protein 1	TEEDSKSIKSDVPVYLK	2 pS	350-366	699.96	2096.84
IPI00298920	Q9UQ09	KARP-1-binding protein 1	TEEDSKSIKSDVPVYLKR	2 pS	350-367	751.98	2252.92
IPI00298920	Q9UQ09	KARP-1-binding protein 1	RFPTDYASTSEDEFGSNR	2 pS + 1 pT	1232-1249	773.58	2317.73
IPI00320643	P27546	MAP4	TSPSKPSSAPALKPGPK	1 pS/T	759-775	577.27	1728.79
IPI00352834	P78559	Microtubule-associated protein 1A	KVAELEEEQSQGSSYSDWVK	2 pS	550-570	849.32	2544.93
IPI00352834	XP_194040	Microtubule-associated protein 1A	HRADSKESLKAAPK	2 pS	760-773	566.57	1696.69
IPI00352834	XP_194040	Microtubule-associated protein 1A	ELALSSPEDLTQDFEELKR	2 pS	967-985	794.01/1190.51	2379.01
IPI00352834	XP_194040	Microtubule-associated protein 1A	ETSPTRGEPVPAWEGKSPEQEVR	1 pS + 1 pS/T	2077-2099	909.38	2725.11
IPI00352834	XP_194040	Microtubule-associated protein 1A	SPFEIISPPASPPEMTGQR	2 pS	2207-2225	1100.95/734.29	2199.86
IPI00352834	XP_194040	Microtubule-associated protein 1A	SPFEIISPPASPPEMTGQR	3 pS	2207-2225	760.93/1140.96	2279.90
IPI00352834	XP_194040	Microtubule-associated protein 1A	GKRSPTPGKGPVDR	1 pS + 1 pT	2983-2996	537.90	1610.66
IPI00130920	P14873	Microtubule-associated protein 1B	ELEAERSLMSSPEDLTKDFEELKAEEIDVAK	1 pS + 2 pS/T	819-849	1264.52	3790.55
IPI00130920	P14873	Microtubule-associated protein 1B	SLMSSPEDLTK	2 pS	825-835	684.24	1366.46
IPI00130920	P14873	Microtubule-associated protein 1B	SLMSSPEDLTKDFEELKAEEIDVAK	2 pS	825-849	995.44	2983.30
IPI00130920	P14873	Microtubule-associated protein 1B	LSPAKSPSLSPSPPSPIEK	1 pS + 2 pS	1246-1264	720.30	2157.87
IPI00130920	P14873	Microtubule-associated protein 1B	SPSLSPSPPSPIEK	2 pS	1251-1264	791.78	1581.54
IPI00130920	P14873	Microtubule-associated protein 1B	VLSPLRSPPLLGSESPYEDFLSADSK	2 pS	1389-1414	988.76	2963.27
IPI00130920	P14873	Microtubule-associated protein 1B	ESSPLYSPGFSDSTSAAK	1 pS + 1 pS	1787-1804	995.86	1989.71

Appendix 4

IPI	Uniprot	Protein Name	Phosphopeptide Sequences	Mod ID	Residues	m/z	Mr (expt)
IPI00118075	P20357	Microtubule-associated protein 2	EAQYKDQPAALPLAAEETANLPPSPPPSPASEQTATVEEDLLTASK	2 pS	113-158	1234.04	4932.14
IPI00118075	P20357	Microtubule-associated protein 2	DQPAALPLAAEETANLPPSPPPSPASEQTATVEEDLLTASK	2 pS	118-158	1079.21/1438.62	4312.83
IPI00118075	P20357	Microtubule-associated protein 2	DQPAALPLAAEETANLPPSPPPSPASEQTATVEEDLLTASK	3 pS	118-158	1465.31	4392.90
IPI00118075	P20357	Microtubule-associated protein 2	STELGSDYYELSDSR	1 pS/T	583-597	901.33	1800.65
IPI00118075	P20357	Microtubule-associated protein 2	GHDLSPLASDILTNTSGSMDEGDDYLPPTTPAVEK	2 pS	722-756	1268.53	3802.56
IPI00118075	P20357	Microtubule-associated protein 2	GHDLSPLASDILTNTSGSMDEGDDYLPPTTPAVEK	3 pS	722-756	1295.18	3882.53
IPI00118075	P20357	Microtubule-associated protein 2	ETSPETSLIQDEVALK	1 pS	1159-1174	920.40/920.43	1838.79/1838.84
IPI00118075	P20357	Microtubule-associated protein 2	DGSPDAPATPEKEEVAFSEYK	1 pS + 1 pT	1350-1370	809.62/1213.95	2425.84
IPI00118075	P20357	Microtubule-associated protein 2	SCVKRKTTAASGDLAQAPGAFK	1 pS + 2 pT	1505-1526	1224.28	2446.55
IPI00118075	P20357	Microtubule-associated protein 2	DKVTDGISKSPEKR	1 pS	1530-1543	547.25	1638.73
IPI00118075	P20357	Microtubule-associated protein 2	SGTSTPTTPGSTAITPGTPPSYSSR	2 pS/T/Y	1592-1616	856.68/1284.53	2567.04
IPI00118075	P20357	Microtubule-associated protein 2	SGTSTPTPGSTAITPGTPPSYSSR	2 pT + 1 pS/Y	1592-1616	1324.51	2647.00
IPI00118075	P20357	Microtubule-associated protein 2	TPGTPGTPSYPR	2 pT	1617-1628	695.75	1389.49
IPI00118075	P20357	Microtubule-associated protein 2	VAIIRTPPKSPATPK	1 pS + 1 pT	1645-1659	579.28	1734.81
IPI00118075	P20357	Microtubule-associated protein 2	VAIIRTPPKSPATPK	2 pS + 1 pT	1645-1659	908.40	1814.78
IPI00118075	P20357	Microtubule-associated protein 2	ARVDHGAEIITQSPSRSSVASPR	3 pS	1771-1793	887.69	2660.05
IPI00118075	P20357	Microtubule-associated protein 2	ARVDHGAEIITQSPSRSSVASPR	4 pS	1771-1793	914.35	2740.04
IPI00118075	P20357	Microtubule-associated protein 2	ARVDHGAEIITQSPSRSSVASPR	5 pS	1771-1793	941.01	2820.00
IPI00118075	P20357	Microtubule-associated protein 2	VDHGAEIITQSPSR	1 pS	1773-1786	530.55/795.34	1588.63/1588.66
IPI00118075	P20357	Microtubule-associated protein 2	VDHGAEIITQSPSRSSVASPR	2 pS	1773-1793	785.33	2352.96
IPI00118075	P20357	Microtubule-associated protein 2	VDHGAEIITQSPSRSSVASPR	2 pS + 1 pS	1773-1793	811.98/1217.48	2432.94
IPI00118075	P20357	Microtubule-associated protein 2	VDHGAEIITQSPSRSSVASPR	4 pS	1773-1793	838.64/1257.47	2512.91
IPI00118075	P20357	Microtubule-associated protein 2	VDHGAEIITQSPSRSSVASPR	3 pS + 2 pS/T	1773-1793	865.30	2592.87
IPI00118075	P20357	Microtubule-associated protein 2	RLSNVSSSGSINLLESPQLATLAEDVTAALAK	1 pS + 1 pS	1794-1825	1139.20	3414.58
IPI00118075	P20357	Microtubule-associated protein 2	RLSNVSSSGSINLLESPQLATLAEDVTAALAK	2 pS + 1 pS	1794-1825	1165.85	3494.52
IPI00118075	P20357	Microtubule-associated protein 2	RLSNVSSSGSINLLESPQLATLAEDVTAALAK	2 pS + 2 pS	1794-1825	894.64/1192.5	3574.49
IPI00118075	P20357	Microtubule-associated protein 2	RLSNVSSSGSINLLESPQLATLAEDVTAALAK	5 pS	1794-1825	1219.20	3654.59
IPI00118075	P20357	Microtubule-associated protein 2	LSNVSSSGSINLLESPQLATLAEDVTAALAK	2 pS	1795-1825	1087.14	3258.41
IPI00118075	P20357	Microtubule-associated protein 2	LSNVSSSGSINLLESPQLATLAEDVTAALAK	3 pS	1795-1825	1113.83	3338.48
IPI00124145.1	P10637	Microtubule-associated protein Tau	IPAKTTPSPKTPPGSGEPPK	1 pS/T + 1 pT	462-481	742.97	2225.88
IPI00124145.1	P10637	Microtubule-associated protein Tau	IPAKTTPSPKTPPGSGEPPK	2 pS/T + 1 pT	462-481	742.97	2225.88
IPI00124145.1	P10637	Microtubule-associated protein Tau	TTPSPKTPPGSGEPPK	1 pS + 1 pT	466-481	579.90	1736.67
IPI00124145.1	P10637	Microtubule-associated protein Tau	TTPSPKTPPGSGEPPKSGER	1 pS + 1 pT	466-485	722.96	2165.85
IPI00124145.1	P10637	Microtubule-associated protein Tau	SGYSSPGSPGTPGSR	2 pS	486-500	777.26	1552.51
IPI00124145.1	P10637	Microtubule-associated protein Tau	SGYSSPGSPGTPGSR	3 pS	486-500	817.25	1632.48
IPI00124145.1	P10637	Microtubule-associated protein Tau	SRTPSLPTPPTREPK	1 pS/T + 1 pT	501-515	608.59	1822.76
IPI00124145.1	P10637	Microtubule-associated protein Tau	KVAVVRTPPKSPSASK	1 pS + 1 pT	516-531	604.63/906.44	1810.88
IPI00124145.1	P10637	Microtubule-associated protein Tau	VAVVRTPPKSPSASK	1 pS + 1 pT	517-531	561.93/842.20	1682.78
IPI00124145.1	P10637	Microtubule-associated protein Tau	AKTDHGAEIVYKSPVVSGDTSPR	1 pS + 1 pT	675-697	858.70	2573.07
IPI00124145.1	P10637	Microtubule-associated protein Tau	AKTDHGAEIVYKSPVVSGDTSPR	3 pS	675-697	885.37/1327.53	2653.09
IPI00124145.1	P10637	Microtubule-associated protein Tau	TDHGAEIVYKSPVVSGDTSPR	1 pS + 1 pT	677-697	792.32/1187.99	2373.97
IPI00124145.1	P10637	Microtubule-associated protein Tau	TDHGAEIVYKSPVVSGDTSPR	2 pS + 1 pT	677-697	819.00	2453.98
IPI00124145.1	P10637	Microtubule-associated protein Tau	TDHGAEIVYKSPVVSGDTSPR	3 pS + 1 pT	677-697	845.63	2533.87
IPI00124145.1	P10637	Microtubule-associated protein Tau	SPVVSGDTSPR	1 pS	687-697	591.24	1180.46
IPI00124145.1	P10637	Microtubule-associated protein Tau	HLSNVSSTGSIDMVDSPQLATLADEVSASLAK	2 pS	698-729	1135.18	3402.52
IPI00124145.1	P10637	Microtubule-associated protein Tau	HLSNVSSTGSIDMVDSPQLATLADEVSASLAK	3 pS	698-729	1161.83	3482.47
IPI00124145.1	P10637	Microtubule-associated protein Tau	HLSNVSSTGSIDMVDSPQLATLADEVSASLAK	3 pS + 1 pS/T	698-729	1188.49	3562.44
IPI00124145.1	P10637	Microtubule-associated protein Tau	HLSNVSSTGSIDMVDSPQLATLADEVSASLAK	4 pS + 1 pT	698-729	1215.09	3642.26
IPI00276456	Q80TK8	MKIAA1078 protein	SESVEGFLSPSR	1 pS	635-646	687.77	1373.53
IPI00367847	Q9CYC6	mRNA decapping enzyme 2	RFGDSSDSDNGFSSAGSTPARPTVEK	3 pS	242-267	971.36	2911.06

Appendix 4

IPI	Uniprot	Protein Name	Phosphopeptide Sequences	Mod ID	Residues	m/z	Mr (expt)
IPI00108886	Q9D7P6	Nifun	AA**S**ALLLR	1 pS	13-20	447.73	893.44
IPI00108886	Q9D7P6	Nifun	AA**S**ALLL**R**SPR	2 pS	13-23	657.80	1313.58
IPI00324794	Q61120	N-SHC	AA**S**VECI**S**PVTPR	2 pS	352-364	745.28	1488.55
IPI00137943	Q9QZS3	Numb	TVVGPSVAPGNTAPSPSSPTSPTPDGTASSEMNNPHAIPR	3 pS/T	224-263	1031.41	4121.62
IPI00349306	Q13427	Peptidyl-prolyl cis-trans isomerase G (Cyclophilin G)	FRR**S**E**T**PPHWR	1 pS + 1 pT	353-363	543.55	1627.62
IPI00349306	Q13427	Peptidyl-prolyl cis-trans isomerase G (Cyclophilin G)	**S**E**T**PPHWR	1 pS + 1 pT	356-363	585.20	1168.38
IPI00330246	Q6PDH0	Phldb1	T**R**S**P**S**P**T**L**GESLAPR	2 pS	518-532	576.90	1727.67
IPI00099004	Q9QYX7	Piccolo protein	GRIPIVAQN**S**EEESPI**S**PVGQPMGMAR	3 pS	4247-4273	1030.74	3089.18
IPI00118438	O70495	Plenty-of-prolines-101 (Srrm1)	RI**S**PSA**S**PPR	2 pS	385-394	614.24	1226.47
IPI00118438	O70495	Plenty-of-prolines-101 (Srrm1)	HRPS**S**PA**T**PPPK	1 pS + 1 pT	397-408	477.76	1430.25
IPI00118438	O70495	Plenty-of-prolines-101 (Srrm1)	RR**S**P**S**PAPPPPPPPPPPR	2 pS	570-587	686.31	2055.90
IPI00118438	O70495	Plenty-of-prolines-101 (Srrm1)	**S**P**S**PAPPPPPPPPPPR	2 pS	572-587	582.24/872.87	1743.70
IPI00118438	O70495	Plenty-of-prolines-101 (Srrm1)	RY**S**PPIQR	1 pS	614-621	548.75	1095.48
IPI00118438	O70495	Plenty-of-prolines-101 (Srrm1)	RR**T**A**S**PPPPPK	1 pS + 1 pT	631-641	455.19	1362.54
IPI00118438	O70495	Plenty-of-prolines-101 (Srrm1)	RR**T**A**S**PPPPPKR	1 pS + 1 pT	631-642	507.23	1518.66
IPI00118438	O70495	Plenty-of-prolines-101 (Srrm1)	RR**T**A**S**PPPPPKR	1 pS + 1 pT	631-642	507.23	1518.66
IPI00118438	O70495	Plenty-of-prolines-101 (Srrm1)	RY**S**P**S**PPPK	2 pS	645-651	594.73	1187.44
IPI00118438	O70495	Plenty-of-prolines-101 (Srrm1)	RY**S**P**S**PPPKR	2 pS	645-652	672.78	1343.54
IPI00118438	O70495	Plenty-of-prolines-101 (Srrm1)	RGA**S**A**S**PQGRQ**S**P**S**P**S**TRPIR	2 pS + 2 pS/T	720-740	838.32	2511.94
IPI00118438	O70495	Plenty-of-prolines-101 (Srrm1)	KPPAPP**S**PVQ**S**Q**S**PSTNWSPAVPAK	2 pS	781-805	905.73	2714.15
IPI00118438	O70495	Plenty-of-prolines-101 (Srrm1)	KPPAPP**S**PVQ**S**Q**S**PSTNW**S**PAVPAK	3 pS	781-805	932.37	2794.09
IPI00118438	O70495	Plenty-of-prolines-101 (Srrm1)	KPPAPP**S**PVQ**S**Q**S**PSTNW**S**PAVPAK	4 pS	781-805	959.05	2874.12
IPI00118438	O70495	Plenty-of-prolines-101 (Srrm1)	KPPAPP**S**PVQ**S**Q**S**PSTNW**S**PAVPAKK	3 pS	781-806	975.08	2922.23
IPI00118438	O70495	Plenty-of-prolines-101 (Srrm1)	KPPAPP**S**PVQ**S**Q**S**PSTNW**S**PAVPAKK	4 pS	781-806	1001.74	3002.21
IPI00329791	O94894	Prp5-like DEAD-box protein	KGELMENDQDAMEY**SS**EEEEVDLQTALTGYQTK	2 pS	307-339	1314.53	3940.58
IPI00002534	Q8R4E7	Pur-gamma	HSAP**S**PPV**S**VG**S**EEHPHSVLK	3 pS	155-175	806.66	2416.95
IPI00144293	Q96MU7	Putative splicing factor YT521	GI**S**PIVFDR	1 pS	306-314	542.24	1082.47
IPI00200893	Q5J8K4	Rab3 effector, RIM-1	SQRI**S**D**S**DI**S**DYEVDDGIGVVPPVGYR	3 pS	736-762	1060.08	3177.23
IPI00200893	Q5J8K4	Rab3 effector, RIM-1	**S**LDEIHPTR	1 pS	812-820	574.23	1146.45
IPI00200893	Q5J8K4	Rab3 effector, RIM-1	QG**S**PTQ**S**PPADTSFGSR	2 pS	1021-1037	940.35	1878.68
IPI00200893	Q5J8K4	Rab3 effector, RIM-1	RA**S**Q**S**LESSSGPPCIR	2 pS	1446-1462	641.25/961.38	1920.74
IPI00200893	Q5J8K4	Rab3 effector, RIM-1	RA**S**Q**SS**LESSSGPPCIR	3 pS	1446-1462	1001.35	2000.69
IPI00005698	Q9EQZ7	Rab3-interacting molecule 2 (RIM 2)	RA**S**Q**SS**LESSTGPSYSR	1 pS + 1 pS	1513-1529	980.37	1958.74
IPI00370973	XP_290799	Rho GTPase activating protein 23	**S**AEALGPGALVSPR	1 pS	686-699	702.83	1403.65
IPI00370973	XP_290799	Rho GTPase activating protein 23	**S**AEALGPGALV**S**PR	2 pS	686-699	742.81	1483.60
IPI00225062	Q8BTI8	RNA binding protein homolog (Srrm2)	HA**SSS**PESLKPTPAPGSR	3 pS	335-352	682.58/1023.39	2044.76
IPI00225062	Q8BTI8	RNA binding protein homolog (Srrm2)	HA**SSS**PE**S**LKPTPAPGSR	4 pS	335-352	709.24/1063.37	2124.72
IPI00225062	Q8BTI8	RNA binding protein homolog (Srrm2)	**S**G**T**PPRPG**S**V**T**NMQADECTATPQR	2 pS/T	725-748	887.67	2660
IPI00225062	Q8BTI8	RNA binding protein homolog (Srrm2)	SR**SSS**PDSKMELGTPLR	3 pS	846-862	696.59	2086.74
IPI00225062	Q8BTI8	RNA binding protein homolog (Srrm2)	**SSS**PVTELTAR**S**PVK	1 pS + 1 pS	970-984	859.87	1717.73

Appendix 4

IPI	Uniprot	Protein Name	Phosphopeptide Sequences	Mod ID	Residues	m/z	Mr (expt)
IPI00225062	Q8BTI8	RNA binding protein homolog (Srrm2)	SSSPVTELTARSPVKQDK	2 pS	970-987	697.30/1045.45	2088.88
IPI00225062	Q8BTI8	RNA binding protein homolog (Srrm2)	EMPGSNIESSPEVEERPAVLSALDQSQSQPSK	2 pS	1125-1156	1196.14	3585.41
IPI00225062	Q8BTI8	RNA binding protein homolog (Srrm2)	SSSASPELKDGLPR	2 pS	1280-1293	802.32	1602.63
IPI00225062	Q8BTI8	RNA binding protein homolog (Srrm2)	AQSGTDSSPEHKIPAPR	3 pS	1474-1490	673.24/1009.38	2016.74
IPI00225062	Q8BTI8	RNA binding protein homolog (Srrm2)	AQSGTDSSPEHKIPAPR	3 pS + 1 pT	1474-1490	699.90/1049.36	2096.71
IPI00225062	Q8BTI8	RNA binding protein homolog (Srrm2)	SRSPLAIR	2 pS	1900-1907	530.23	1058.45
IPI00225062	Q8BTI8	RNA binding protein homolog (Srrm2)	RSLTRSPPAIR	2 pS + 1 pT	1922-1932	498.54	1492.60
IPI00225062	Q8BTI8	RNA binding protein homolog (Srrm2)	NHSGSRTPPVALSSSR	1 pS + 1 pT	1954-1969	604.91	1811.70
IPI00225062	Q8BTI8	RNA binding protein homolog (Srrm2)	NHSGSRTPPVALSSSR	2 pS + 1 pT	1954-1969	631.57/946.85	1891.68
IPI00225062	Q8BTI8	RNA binding protein homolog (Srrm2)	ARSRTPPSAPSQSR	1 pS + 1 pT	2262-2275	553.24/829.34	1656.69
IPI00225062	Q8BTI8	RNA binding protein homolog (Srrm2)	KPIDSLRDSRSLSYSPVER	3 pS + 1 pS	2538-2556	842.32	2523.94
IPI00329839	Q9CTM9	RNA Helicase-related protein homolog Ddx42	QQFHSKPADSDSDDDPLEAFMAEVEDQAAR	2 pS	100-129	1170.46	3508.37
IPI00515701	Q5RJI5	Serine/threonine kinase SADB	SRSVSGASTGLSSSPLSSPR	1 pS + 1 pS	430-449	693.95	2078.84
IPI00515701	Q5RJI5	Serine/threonine kinase SADB	STGLSSSPLSSPRSPVFSFSPEPGAGDEAR	8 pS	437-466	1216.15	3645.44
IPI00515701	Q5RJI5	Serine/threonine kinase SADB	SSGGTPLHSPLHTPR	1 pS + 1 pS	511-525	568.56	1702.66
IPI00515701	Q5RJI5	Serine/threonine kinase SADB	ASPTGTPGTTPPPSPGGGVGGAAWR	1 pS/T + 1 pT	526-550	1197.00	2391.99
IPI00515701	Q5RJI5	Serine/threonine kinase SADB	ASPTGTPGTTPPPSPGGGVGGAAWR	1 pS/T + 1 pT	526-550	824.97	2471.90
IPI00312081	Q9JLM8	Serine/threonine-protein kinase DCAMKL1	SGKSPSPSPTSPGSLRK	2 pS	327-343	610.59	1828.74
IPI00312081	Q9JLM8	Serine/threonine-protein kinase DCAMKL1	ISQHGGSSTSLSSTK	1 pS + 1 pS/T	346-360	818.81	1635.60
IPI00312081	Q9JLM8	Serine/threonine-protein kinase DCAMKL1	AQPAPPELNSESEDYSPSSSETVR	1 pS + 1 pS	727-750	913.00	2735.99
IPI00320690	Q61136	Serine/threonine-protein kinase PRP4 homolog	VQSGMGLILQGYESGSEEEGEIHEK	2 pS	130-141	956.07	2865.19
IPI00320690	Q61136	Serine/threonine-protein kinase PRP4 homolog	SKDASPINRWSPTR	2 pS	427-440	592.24/887.86	1773.71
IPI00204379	Q9WV48	Shank1	SMSVPDDAHFSMMVFR	2 pS	61-76	672.89/1008.85	2015.69
IPI00204379	Q9WV48	Shank1 (peptide is isoform 3 specific, site is not)	HYTVGSYDSFDAPSDYIIK	2 pS	633-651	779.96/1169.41	2336.86
IPI00125582	Q5RIT9	similar to PITPNM family member 3	GSPPLLDAPASPPQAPR	2 pS	494-510	915.88	1829.74
IPI00396130	Q9QWI6	SNAP-25-interacting protein (SNIP)	RGSDELTVPR	1 pS	1052-1061	605.26	1208.52
IPI00008513	Q8K004	Spata2	SVDAYDSYWESR	1 pS	247-258	779.27	1556.52
IPI00139364	Q8VE97	Splicing factor, arginine/serine-rich 4	ARSRSTSKSKPNVPAESR	3 pS/T	437-454	733.31	2196.92
IPI00139364	Q8VE97	Splicing factor, arginine/serine-rich 4	SRSTSKSKPNVPAESR	1 pS + 1 pS/T	439-454	630.93	1889.76
IPI00139364	Q8VE97	Splicing factor, arginine/serine-rich 4	SRSTSKSKPNVPAESR	2 pS + 1 pS/T	439-454	657.58	1969.71
IPI00139364	Q8VE97	Splicing factor, arginine/serine-rich 4	STSKSKPNVPAESR	1 pS + 1pS/T	441-454	549.89/824.33	1646.64
IPI00229604	Q8VCD2	SR-related CTD associated factor 6	SRSPTPPSAAGLGSNSAPPIPDSR	2 pS + 1 pS/T	822-845	853.67	2557.98
IPI00229604	Q8VCD2	SR-related CTD associated factor 6	SRSRSPTPPSAAGLGSNSAPPIPDSR		830-845	961.38	2881.11
IPI00136372	O88935	Synapsin I	GSHSQSSSPGALTLGR	2 pS	431-446	851.32	1700.63
IPI00136372	O88935	Synapsin I	PVAGGPGAPPAARPPASPSPQR	2 pS	535-556	732.32	2193.93
IPI00136372	O88935	Synapsin I	QSRPVAGGPGAPPAARPPASPSPQR	2 pS	532-556	856.04	2565.10
IPI00296573	Q8CC35	Synaptopodin	VASEEEEVPLVVYLK	1 pS	256-270	892.42	1782.82
IPI00128775	Q9R1L5	Syntrophin-associated serine-threonine protein kinase	VGSSHTRR	2 pS	1226-1233	530.24	1058.46
IPI00359917	Q99J36	THUMP domain containing protein 1	FIDKDQQPSGSEGEDDDAEAALKK	2 pS	78-101	918.35	2752.02

Appendix 4

IPI	Uniprot	Protein Name	Phosphopeptide Sequences	Mod ID	Residues	m/z	Mr (expt)
IPI00359323	Q5M7V8	Thyroid hormone receptor associated protein 3	ASVSDL**S**PR	1 pS	237-245	506.21	1010.41
IPI00359323	Q5M7V8	Thyroid hormone receptor associated protein 3	ER**S**PALK**S**PLQSVVVR	2 pS	246-261	963.44/642.63	1924.87
IPI00359323	Q5M7V8	Thyroid hormone receptor associated protein 3	IDI**S**PSTFR	1 pS	676-684	558.24	1114.47
IPI00359323	Q5M7V8	Thyroid hormone receptor associated protein 3	**S**REEEWDPEYTPK	1 pS	860-872	873.33	1744.65
IPI00372166	Q9UKE5	TRAF2 and NCK interacting kinase	VRAN**S**KSEG**S**PVLPHEPSK	2 pS	760-778	726.98	2177.92
IPI00372166	Q9UKE5	TRAF2 and NCK interacting kinase	AN**S**KSEG**S**PVLPHEPSK	2 pS	762-778	641.92	1922.73
IPI00119067	Q7TQD2	Tubulin polymerization-promoting protein (p25 alpha)	AANK**T**PPKSPGDPAR	1 pT	11-25	529.57	1585.69
IPI00119067	Q7TQD2	Tubulin polymerization-promoting protein (p25 alpha)	AANK**T**PPK**S**PGDPAR	1 pS + 1 pT	11-25	556.23/833.85	1665.67
IPI00119067	Q7TQD2	Tubulin polymerization-promoting protein (p25 alpha)	RL**S**LE**S**EGANEGATAAPELSALEEAFR	2 pS	29-55	993.42	2977.23
IPI00119067	Q7TQD2	Tubulin polymerization-promoting protein (p25 alpha)	RL**S**LE**S**EGANEGATAAPELSALEEAFRR	2 pS	29-56	1045.47	3133.39

Appendix 4 - Unique phosphopeptide sequences detected in double IMAC purifications from forebrain cytosolic fraction by LC-MS/MS analysis. Unambiguously assigned sites of phosphorylation are denoted in the sequence as S/T/Y (red) and ambiguous sites are just coloured in pink. Mod Id: number of phosphates detected.

Appendix 5

IPI	UniProt	Protein Name	Phosphopeptide Sequences	Mod ID	Residues	MS defined Phospho-sites	Predicted Phospho-sites	Scansite Kinase	Scansite Binding site	Order	Known kinase	Literature
IPI00108780	P17426	Adapter-related protein complex 2 alpha 1 subunit	DTpSNDINGOVEPTPSIVSTPpSPSADLLGLP	1 pS + 2 pS/T	634-664		S652	CK1 (0.5268)		Disord	uk	15572359
							T653	GSK3 (0.4195), ERK1 (0.5290)		Disord	uk	15572359
						S655		ERK1 (0.6120), CK1 (0.5251)		Ord	uk	15572359
IPI00121136	Q9JIX8	Apoptotic chromatin condensation-inducer in the nucleus (Acin1)	pSLSPLSQTTDTKASpSPAQR	2 pS	477-495	S477				Ord		
						S491				Disord		
IPI00202484	P62995	Arginine/serine-rich splicing factor 10 (Sfrs10)	RPRpSRpSRpSPMSTP	3 pS	92-104	S95		PKA (0.4018), Clk2 (0.5341)	14-3-3 m1 (0.4381)	Disord		
						S97		Clk2 (0.6270)		Disord		
						S99		Clk2 (0.5796), Cdk5 (0.5331), Cdc2 (0.5495)		Disord		
			RPpSPpSPpSYpSP	3 pS + 1pY	262-271	S264		PKA (0.3333)	14-3-3 m1 (0.4961)	Disord		
						S266		Akt (0.5820), PKA (0.5629), Cdk5 (0.5556)	14-3-3 m1 (0.5553)	Disord		
						Y268				Disord		
						S270				Disord		
IPI00117229	O70305	Ataxin-2 (Spinocerebellar ataxia type 2 homolog)	TNpSPSIpSPSMLSNAEHK	2 pS	830-846	S832		GSK3 (0.5069)		Disord		
						S836				Disord		
IPI00169477	Q8K019	Bcl-2-associated transcription factor 1 [Bclaf1]	NTPpSQNSHSIQNpSPER	2 pS	255-270	S258				Ord		
						S267		Cdk5 (0.4911)		Ord		
			IGPpSQNpSPIHHIPSR	1 pS + 1 pS/Y	283-297		S284			Disord	uk	14729942
						S289		Cdk5 (0.4127), Cdc2 (0.4751)		Disord	uk	14729942
			IDIpSPRLR	1 pS	653-661	S656				Disord		
IPI00311962	Q9QYB8	Beta adducin	AQTKpSPAVpSPSKTSEDTKK	2 pS	610-628	S614		GSK3 (0.5452), Cdk5 (0.5715)		Disord		
						S618		Cdk5 (0.4669), Cdc2 (0.5105), ERK1 (0.5580)		Disord		
			KLEQEQEGEKDIATEKPGSPVKpSpTPApSPVQpSPEK	3 pS + 1 pT	576-609	S598		GSK3 (0.4911)		Disord		
						T599				Disord		
						S602		GSK3 (0.5441), ERK1 (0.5441)		Disord		
						S606		Cdc2 (0.5393), Cdk5 (0.4986)		Disord		
IPI00326171	Q68EG2	CaMKII beta	GSLPPAALKpSpTNTTIEDEDAK	2 pS/T	345-368		S354	CK2 (0.5447)		Disord		
							T358	CK1 (0.4686)		Disord		
IPI00296678	Q5SQZ3	CaMKII gamma	RKpSpSPSHEVKLMSPQTTVVHNATDQIK	3 pS	353-377	S355		PKA (0.4521), PKC mu (0.5985), CaMKII (0.5667)	14-3-3 m1 (0.5965)	Disord		
						S358		PKA (0.5185), CaMKII (0.5543), Akt (0.6614)	14-3-3 m1 (0.5300)	Disord		
						S357				Disord		
IPI00357503	Q7TQF8	Camsap1	QDpSDpSDVVDVEDTEQDFIGEDHDVVIPR	2 pS	809-820	S811				Disord		
						S813				Disord		
IPI00276456	Q80TK8	Camsap1 (MKIAA1078 protein)	SEEVEQFLpSPSR	1 pS	635-646	S643		ERK1 (0.5377), Cdk5 (0.5333)		Disord		
IPI00385480	Q6P9K8	Caskin-1	YAApSDpSEPERDELLVPAAAGPYATVQR	2 pS	888-914	S891				Disord		
						S893				Disord		
IPI00349306	Q13427	Cyclophilin G	pSRpSPPHKR	1 pS + 1 pT	356-363	S356				Disord	uk	15302935
						T358		Cdk5 (0.4270), Cdc2 (0.4880)	14-3-3 m1 (0.3602)	Disord	uk	15302935
IPI00261366	Q80UZ7	D430018P08 protein	EKpSpSPSLNPDApSPVTAK	3 pS	176-192	S178			14-3-3 m1 (0.0952)	Disord		
						S179		Cdk5 (0.256), Cdc2 (0.3207), GSK3 (0.8025)		Disord		
						S187		Cdc2 (0.3394)		Disord		
IPI00190337	Q9WV69	Dematin	ATpSPPpSPEVWAESR	1 pS + 1 pS/T	90-104		S92	ERK1 (0.4387), GSK3 (0.5211), Cdk5 (0.5760)	14-3-3 m1 (0.4978)	Disord	uk	15572359
						S96		ERK1 (0.4657), GSK3 (0.5544), Cdk5 (0.5816)		Disord		
			HIpSAEDFER	1 pS	370-378	S372		ATM (0.4360), Clk2 (0.6472), PKC mu (0.5141)	14-3-3 m1 (0.5774)	Ord		
IPI00327311	P97427	Dihydropyrimidinase related protein-1	pSIPKITYDR	1 pS	8-16	S8		PKC epsilon (0.2933)		Disord		
			HAAPAPpSAPpSRpSPSKHQPPPIR	3 pS	512-532	S518				Ord		
						S521				Ord		
						S522				Ord		
IPI00257508	O08553	Dihydropyrimidinase related protein-2	GLYDGPVCEVSVpSPKTVpSPApSpSAKpSpSPAKQCAPPVR	1pS/T + 6pS/T	497-532	T509		Cdk5 (0.5741), Cdc2 (0.6580), ERK1 (0.6693)		Disord	uk	10757975
							T512	PKC epsilon (0.5223)		Disord		
						T514		Cdk5 (0.5964), Cdc2 (0.6333), ERK1 (0.6369)		Disord	GSK3B	15652488
						S517		PKC a/b/g (0.5315)		Disord		
						S518		GSK3 (0.5319)		Disord	uk	10757975
						T521				Disord		
						S522		Cdk5 (0.4986), Cdc2 (0.5843)		Disord	uk	10757975
IPI00391456	Q7Z417	FMRP (fragile X mental retardation protein)-interacting protein	NDpSWGSFDLR	1 pS	650-659	S652		PKC mu (0.5283), PKA (0.5788), Akt (0.5301)	14-3-3 m1 (0.4725)	Disord		
						S655				Disord		
IPI00120424	Q68FF6	G protein-coupled receptor kinase-interactor 1	SLSpSPTDNLELSAR	1 pS	368-381	S371				Disord		
			RQSELDDQMDYDpSVApSDEDpSDQEPLPSAGATR	2 pS + 1 pT	382-413	S394		CK2 (0.5255)		Ord		
						S397		CK1 (0.3974), CK2 (0.4778)		Disord		
						T401		CK2 (0.4302), CK1 (0.5017)		Disord		
			HQpSGADEDYEMTQKDDPLLGLEQKP	1 pS + 1 pS/T	599-623	S601		PKC epsilon (0.4830), PKC mu (0.4873)		Disord		
							S605	CK2 (0.5710)		Disord		
IPI00021596	Q80YT1	Glucocorticoid induced transcript 1 (Gloc1)	SApSWGSADQLK	1 pS	25-35	S27		Akt (0.5009), Clk2 (0.6182)	14-3-3 m1 (0.3003)	Ord		
IPI00292228	Q9WV60	Glycogen synthase kinase-3 beta	GEPHVpSYICSR	1 pY	210-220	Y216				Ord	MEK1	15020233
IPI00377522	Q9NXJ5	Gm1568	SPpSLGDLSMpSPQTAGHVER	2 pS	301-319	S303		Akt (0.4244), PKA (0.5261)	14-3-3 m1 (0.2867)	Ord		
						S310		Cdk5 (0.5798), Cdc2 (0.6183)		Disord		
IPI00137459	Q6PG39	HSHIN1 protein	LQRPKEEpSpSpSEDENEVSNILR	2 pS	304-322	S310		CK2 (0.4007)		Disord		
IPI00006669	Q96D85	IQSEC1	MQPpSPEQPEK	1 pS	55-64	S58		CaMKII (0.4867), ATM (0.5075)	PDK1 binding (0.5561)	Ord		
			HSMDpSPAFSMDVIR	1 pS	389-402	S393				Disord		
IPI00298920	Q9UQQ9	KARP-1-binding protein 1	TEEDSKpSIKpSDVPVYLK	2 pS	350-366	S356				Disord		
						S359				Ord		
			RPPTDYApSpTpSKDEPGSNR	2 pS + 1 pT	1232-1249	S1239				Disord		
						T1240		CK2 (0.4703)		Disord		
						S1241		CK2 (0.4713)		Disord		
IPI00320643	P27546	MAP4	IApSKPSSAPALKPGPK	1 pS/T	759-775		S760	Cdk5 (0.4799), PKA (0.4464)		Disord	Cdc2	9398320
IPI00352834	P78559	Microtubule-associated protein 1A	KVAELEEEQpSQpSpSPSYSDMVK	2 pS	550-570	S559				Disord		
						S562				Disord		
			GKPpSPpTPGKGPVDR	1 pS + 1 pT	2597-2610	S2600				Disord		
						T2602				Disord		
			HRADpSKEpSLKAAPK	2 pS	760-773	S764				Disord		
						S767				Disord		
			RLALEpSpSPEDLTQDFEELKE	2 pS	967-985	S971				Ord		
						S972				Ord		
			KTpSPIRGEPVPAWEGKpSPEQEVR	1 pS + 1 pS/T	2077-2099	S2093				Disord		
			pSPFEIIpSPPApSPFEMTGQF	3 pS	2145-2163	S2145				Ord		
						S2151		GSK3 (0.4759), ERK1 (0.4776)		Ord		
						S2155		ERK1 (0.4796)		Ord		

Appendix 5

IPI	UniProt	Protein Name	Phosphopeptide Sequences	Mod ID	Residues	MS defined Phospho-sites	Predicted Phospho-sites	Scansite Kinase	Scansite Binding site	Order	Known kinase	Literature
IPI00130920	P14873	Microtubule-associated protein 1B	KLEAKPpSLMpSpSPEDLTKDPEELKAEEIDVAK	3 pS	819-849	S825		PKC mu (0.4853)		Disord		
						S828				Disord		
						S829		ATM (0.4304)		Disord		
			LSPAKSPpSLpSPpSPPpSPIEK	2pS + 1 pS	1246-1264	S1253		GSK3 (0.6423)	14-3-3 m1 (0.1446)	Disord		
							S1255			Disord	uk	15345747
						S1260		p38 MAPK (0.0812), Cdk5 (0.2184), ERK1 (0.3835)		Disord	GSK3B	15731007
										Disord		
			VLpSPLPpSPPLLQHEPYEDPLEADEK	2 pS	1389-1414	S1391				Ord		
						S1395		GSK3 (0.4814)		Ord	uk	15345747
			KpSPLYpSPGPSDSTSAAK	1 pS + 2 pS	1787-1804		S1789			Ord	uk	15345747
						S1793				Ord	uk	15345747
IPI00118075	P20357	Microtubule-associated protein 2	DQPAALPLAAEEpTANLPPpSPPPpSPApSPpTATVEEELLTAEK	4 pS	118-158	T130				Ord		
						S136		p38 MAPK (0.0814), ERK1 (0.2797), ATM (0.452)		Ord	Cdk2, GSK3-alpha, ERK1, Cdk5	7525290
						S140		p38 MAPK (0.0865), ERK1 (0.3788)		Ord		
						S143				Ord		
			pTKLGSDTYELSDSR	1 pS/T	583-597		S583			m		
							T584			m		
										m		
			GNDLpPLApSDILpTMTpSpSMDKGDDYLPPTTPAVEK	3 pS	722-756		S726			Ord	uk	15345747
							S739	CK1 (0.3957)		Ord		
			ETpSPETSLIQDEVALK	1 pS	1159-1174	S1161				Disord		
			DSpSPDAPApSpTPEKEEVAFSEEYK	1 pS + 1 pT	1350-1370	S1352				Disord	uk	15345747
						T1358		p38 MAPK (0.0884), ATM (0.4478)		Disord	uk	15345747
			SCVKEKpSpTAApSDDLAQAPGAFK	1 pS + 2 pT	1505-1526	T1511		PKC delta (0.401)		Ord		
						T1512		PKC epsilon (0.4438)		Ord		
			DKVTDGISKpSPEKR	1 pS	1530-1543	S1539		Cdk5 (0.1394), Cdc2 (0.1765)		Disord		
			SGTpTPTTPGSpAIIPGTPTPpSISSR	2 pS/T/Y	1592-1616					Disord		
			SGTSpIPpYpTPGSTAITPGTPTPpSISSR	2 pT + 1 pS/Y	1592-1616	T1596		ERK1 (0.3621)		Disord		
						T1598		CK1 (0.232)		Disord	uk	15345747
			TPQpSPQpTPGYPR	2 pT	1617-1628	T1620		p38 MAPK (0.0757), Cdk5 (0.2102)		Ord	GSK3-alpha (1)	9759974
						T1623		p38 MAPK (0.0601), Cdk5 (0.1997)		Ord	GSK3-alpha (1)	9759974
			VAIIPpTPPKpSPApTPK	2 pS + 1 pT	1645-1659	T1650		Cdk5 (0.2169), GSK3 (0.661)		Disord		
						S1654		Cdc2 (0.2559), ERK1 (0.3871)		Disord		
						T1657		Cdk5 (0.2027), Cdc2 (0.2655)		Disord		
			ARVDHGAEIITQpSPpSPpSVApSPR	5 pS	1771-1793	S1783				Disord		
						S1785				Disord		
						S1787				Disord		
						S1788				Disord		
						S1791		Cdk5 (0.1872), Cdc2 (0.2366)		Disord		
			RLpSNVpSpSpSpSQpSINLLESPQLATLAEDVTAALAK	5 pS	1794-1825	S1796				Ord		
						S1799				Ord		
						S1800				Ord		
						S1801				Ord		
						S1803		CK1 (0.2123), Akt (0.4882)		Ord		
IPI00124145.1	P10637	Microtubule-associated protein Tau	IPAKTTPpSPKpSPPGSEPPK	1 pT + 2 pS/T	462-481		T467	ERK1 (0.5454)		Disord	Cdk5	14642273
						S469		Cdk5 (0.2198), Cdc2 (0.2998)		Disord	Cdk5	14642273
						T472		p38 MAPK (0.079), Cdk5 (0.1977), ERK1 (0.3493)		Disord	Cdk5	14642273
										Disord		
			SGYpSpSPGpSPGTPGSR	3 pS	486-500	S489		CK1 (0.4964)		Disord	Cdk5	14642273
						S490		Cdc2 (0.329)		Disord	Cdk5	14642273
						S493		Cdc2 (0.2576), ERK1 (0.3841)		Disord	Cdk5, PKA	14642273, 10460255
			SRpTpSLPTPPTREPK	1 pS/T + 1 pT	501-515	T503		CaMKII (0.3087), Akt (0.3681)	14-3-3 m1 (0.1093)	Disord		
							S505	Akt (0.2656)	14-3-3 m1 (0.2656)	Disord		
			KVAVVRpTPPKpSPSASK	1 pS + 1 pT	516-531	T522		Cdk5 (0.2169), Cdc2 (0.2806)		Disord	Cdk5, PKA	14642273, 10460255
						S526		Cdk5 (0.2196), Cdc2 (0.2506)		Disord	Cdk5	14642273
			TDHGAEIVYKpSPVVpSGDpTpSPR	3 pS + 1 pT	677-697	S687				Disord	Cdk5	14642273, 11152466
						S691		GSK3 (0.6882)		Disord	Cdk5	14642273, 8245007
						T694				Disord	Cdk5	14642273
						S695		Cdk5 (0.4257), Cdc2 (0.4751)		Disord		15345747, 11152466, 15302935
			RLpSNVpSpSpSpTQpSIDMVDSPQLATLADEVSASLAK	4 pS + 1 pT	698-729	S700				Ord	PKA	10460255
						S703				Ord		
						S704				Ord		
						T705				Ord		
						S707		PKC mu (0.5185), CK1 (0.4446)		Ord		15345747
IPI00367847	Q8CYC6	mRNA decapping enzyme 2	RPSDpSpSpSDpSDNSFSSAGSTPARPTVEK	3 pS	242-267	S246		Clk2 (0.4307)		Ord		
						S247		Clk2 (0.7770), CK2 (0.5804)	PDK1 binding (0.9375)	Ord		
						S249		CK1 (0.5251)		Disord		
IPI00177047	Q80TS1	nArgBP2	KPLSVSPSTDGLPpSPpSPPPR	2 pS	179-198	S192		ERK1 (0.3464)	14-3-3 m1 (0.3464)	Disord		
						S194		Cdk5 (0.4697), Cdc2 (0.5182)	14-3-3 m1 (0.3321)	Disord		
IPI00009071	Q9R0U0	Neural specific sr protein NSSR 1	ApSPpSPpSPDYNYR	3 pS	129-139	S129		Clk2 (0.4194), PKA (0.3803)	14-3-3 m1 (0.4404)	Disord		
						S131		Clk2 (0.4233), Akt (0.4451)	14-3-3 m1 (0.3403)	Disord		
						S133		Akt (0.2830), Clk2 (0.4154), PKC zeta (0.4174)	14-3-3 m1 (0.4602)	Disord	uk	15302935
IPI00108886	Q9D7P6	Nifun	AApSALLPpSPR	2 pS	13-23	S15		Clk2 (0.6553), PKA (0.5832), PKC zeta (0.5975)	14-3-3 m1 (0.5404)	Ord		
						S21		Cdk5 (0.5519), Cdc2 (0.5811)		Disord		
IPI00324794	Q61120	N-SHC	AApSVECIpSPVTPR	2 pS	352-364	S354		PKC a/b/g (0.5306), PKA (0.5737)		Ord		
						S359		ERK1 (0.6548)		Ord		
IPI00137943	Q9QZS3	Numb	TVVGPSVAPGNTAPSPSGPTSPTPDGTASSEMNNPHAIPR	3 pS/T	224-263		S240	GSK3 (0.5103)		m		
							S241	ERK1 (0.4603)		m		
							S244	GSK3 (0.5063), ERK1 (0.4611)		m		
IPI00282957	Q8CIL3	Parroc1	pSKGPpSPTATGPR	1 pS + 1 pT	95-106	S95		GSK3 (0.5387)		Disord		
						T99		Cdk5 (0.2768), Cdc2 (0.3363)		Disord		
			RRpSQPpSPTTVPASDpSPPAKQDVK	3 pS	113-135	S115		PKA (0.4574)	14-3-3 m1 (0.3768)	Disord		
						S118		PKC delta (0.4422)		Disord		
						S127		Cdk5 (0.5247)		Disord		
			AAEKEKPAAPApSPAPpSPVPpSPTPAQPQK	3 pS	533-560	S544		ERK1 (0.4929), GSK3 (0.4195)		Ord		
						S548		GSK3 (0.2766), ERK1 (0.4539)		Ord		
						S552		GSK3 (0.4979), ERK1 (0.4284)		Ord		

Appendix 5

IPI	UniProt	Protein Name	Phosphopeptide Sequences	Mod ID	Residues	MS defined Phospho-sites	Predicted Phospho-sites	Scansite Kinase	Scansite Binding site	Order	Known kinase	Literature
IPI00330246	Q6PDH0	Phldb1	TRpSPpTLGSSLAPR	2 pS	518-532		S522	Akt (0.4869)	14-3-3 m1 (0.5055)	Disord		
							T524	GSK (0.5279)		Disord		
							S520	PKA (0.5115)	14-3-3 m1 (0.4794)	Disord		
IPI00099004	Q9QYX7	Piccolo protein	GRIPIVAQNpSEEEpSPLpSPVGPMGMAR	3 pS	4247-4273	S4256				Ord	uk	15572359
						S4260				Ord	uk	15572359
						S4263		ERK1 (0.4257)		Ord	uk	15572359
IPI00118438	O70495	Plenty-of-prolines-101 (Srrm1)	RLpSPSApSPPR	2 pS	385-394	S387		PKA (0.0642), CK1 (0.2225)		Disord		
						S391		Cdk5 (0.1527), Cdc2 (0.197)		Disord		
			HRPSpSPApTPPPK	1 pS + 1 pT	397-408	S401				Disord		
						T404				Disord		
			RRpSPpSPAPPPPPPPPPR	2 pS	570-587	S572		PKA (0.0398)	14-3-3 m1 (0.1428)	Disord		
						S574		Akt (0.38)		Disord		
			RYpSPPIQR	1 pS	614-621	S616		Akt (0.3761), Cdc2 (0.2619)	14-3-3 m1 (0.1158)	Disord		
			RRpTApSPPPPK	1 pS + 1 pT	631-641	T633			14-3-3 m1 (0.1477)	Disord		
						S635		Akt (0.3493)	14-3-3 m1 (0.109)	Disord		
			RYpSPpSPPPK	2 pS	622-630	S624				Disord		
						S626				Disord		
			HOApSApSPQGRQPPPSTNPIR	2 pS + 2 pS/T	720-740	S723			14-3-3 m1 (0.1193)	Disord		
						S725				Disord		
			KPPAPPpSPVQpSPQpSPSTNMpSPAVPAK	4 pS	781-805	S787		GSK3 (0.4833)		Disord		
						S791		GSK3 (0.563)		Ord		
						S793				Ord		
						S799				Ord		
IPI00225062	Q8BT18	RNA binding protein homolog (S-rrm2)	HApSPpSYSPEpSLKPTPAPGSR	4 pS	335-352	S337				Disord		
						S338				Disord		
						S339				Disord		
						S342				Disord		
			QGIPPPRPQIVIMMQADECTATPQR	2 pS/T	725-748					Disord		
			SPpSPpSPpSPDSKMKLGTPLR	3 pS	846-862	S848			14-3-3 m1 (0.3365)	Disord		
						S849				Disord		
						S850				Disord		
			SSpSPVTRLTARpSPVKQDK	2 pS	970-987	S972				Disord		
						S981				Disord		
			HMPQpSMIRSpSPEVERRPAVLSALDQSQSQPSK	2 pS	1125-1156	S1129				Ord		
						S1134				Ord		
			SSpSApSPRLKDGLPR	2 pS	1280-1293	S1282		Akt (0.19)	14-3-3 m1 (0.0768)	Disord		
						S1284				Disord		
			AQpSQpSTDpSpSPEHKIPAPR	3 pS + 1 pT	1474-1490	S1476				Disord		
						T1478				Disord		
						S1480				Disord		
						S1481				Disord		
			pSPpSPLAIR	2 pS	1900-1907	S1900			14-3-3 m1 (0.0696)	Disord		
						S1902				Disord		
			RpSLpSTpSPPAIR	2 pS + 1 pT	1922-1932	S1923				Disord		
						T1925				Disord		
						S1927				Disord		
			KNpSQpSRpSTPVVALSSSR	2 pS + 1 pT	1954-1969	S1956				Disord		
						S1958				Disord		
						T1960				Disord		
			ARpSRpTPPSAPSQSR	1 pS + 1 pT	2262-2275	S2264				Disord		
						T2266			14-3-3 m1 (0.0631)	Disord		
			KPIDpSLRDRpSLpSTpSPVER	3 pS + 1 pS	2538-2556	S2548				Disord		
						S2550				Disord		
						S2552				Disord		
IPI00329791	O94894	Prp5-like DEAD-box protein	KGSLMRNDQDAMGYpSpSEKERVDLQTALTGYQTK	2 pS	307-339	S321				Ord		
						S322				Ord		
IPI00002534	Q8R4E7	Pur-gamma	HSAPpSPPVpSVQpSRSPSVLK	3 pS	155-175	S159				Disord		
						S163				Disord		
						S166				Disord		
IPI00144293	Q96MU7	Putative splicing factor YT521	GIpSPIVPR	1 pS	306-314	S308		Akt (0.5759), PKC mu (0.5542)		Disord	uk	15302935
IPI00200893	Q5JBK4	Rab3 effector, RIM-1	SQRIpSDpSDIpSDYEVDDGIGVVPPVGYR	3 pS	736-762	S740				Disord		
						S742				Disord		
						S745		CK2 (0.4927), CK1 (0.4430)		Disord		
			pSLDHINPTR	1 pS	812-820	S812				Disord		
			QQpSPTQpSPPADTSPGSR	2 pS	1021-1037	S1023				Disord		
						S1027		GSK3 (0.5226), ERK1 (0.5218)		Disord		
			RApSQpSpSLRSSSGPPCIR	3 pS	1446-1462	S1448		PKA (0.3781), ATM (0.4594)		Disord		
						S1450				Disord		
						S1451		PKC delta (0.3385), PKC a/b/g (0.4652)		Disord		
IPI00005698	Q9EQZ7	Rab3-interacting molecule 2 (RIM-2)	RApSQpSSLRSSTGPSYSR	1 pS + 1 pS	1513-1529	S1515		PKA (0.3862), ATM (0.4685)		Disord		
							S1518	PKC delta (0.3385), PKC a/b/g (0.4652)		Disord		
IPI00370973	XP_290706	Rho GTPase activating protein 23	pSAKALGPGALVpSPR	2 pS	686-699	S686				Ord		
						S697		Cdc2 (0.5413)		Disord		
IPI00329839	Q9CTM9	RNA Helicase-related protein homolog Ddx42	QQPHSKPADpSDpSDDDPLEAFMAEVEDQAAR	2 pS	100-129	S109		Clk2 (0.8090), CK2 (0.5444)		Ord		
						S111		DNA PK (0.43), CK2 (0.3395)		Ord		
IPI00515701	Q5RJI5	Serine/threonine kinase SADB (SRSK1)	pSPpTGLpSpSpSPLpSpSPRpSPVFSFSDEPGAGDEAR	6 pS/T	437-466	S437		CK1 (0.4740)		Disord		
						T438				Disord		
						S441				Disord		
						S442				Disord		
						S443				Disord		
						S446		GSK3 (0.4995)		Disord		
						S447		ERK1 (0.5137), Cdc2 (0.5341), Cdk5 (0.5208)		Disord		
						S450		PKC mu (0.4245), p38 MAPK (0.4042), GSK3 (0.4701)		Disord		
			ASGGTPLKpSPLHTPR	1 pS + 1 pS	511-525	S519		Cdc2 (0.4354), ERK1 (0.5125), Cdk5 (0.4184)		Ord		
			ASPTGTPGTpSPPPSPGGGVGGAAMR	1 pS/T + 1 pT	526-550		S527	ERK1 (0.4921)		Ord		
						T535		GSK3 (0.3606), ERK1 (0.3894)		Ord		

Appendix 5

IPI	UniProt	Protein Name	Phosphopeptide Sequences	Mod ID	Residues	MS defined Phospho-sites	Predicted Phospho-sites	Scansite Kinase	Scansite Binding site	Order	Known kinase	Literature
IPI00360155	Q6DMN7	BR serine/threonine-protein kinase 2 (BRSK2)	pSMEVLSVTDGGpSPVPAR	2 pS	383-399	S383		PKC delta, (0.4615), PKC epsilon (0.5097), PKA (0.5759)		Disord		
						S394		Cdk5 (0.5621)		Disord		
			SRSISGASpSQLpSTpSPLSSPR	1 pS + 1 pS	411-430	S424				Ord		
IPI00312081	Q9JLM8	Serine/threonine-protein kinase DCAMKL1	SGKpSPSPSPpYZpSPGSLRK	3 pS	327-343	S330		GSK3 (0.4607), Cdk5 (0.4995). ERK1 (0.5486)		Disord		
						T336		GSK3 (0.4719)		Disord		
						S337		p38 MAPK (0.3335), Cdk5 (0.4799), Cdc2 (0.4954)		Disord		
			ISQRGGpSTpSLSSTK	1 pS + 1 pS/T	346-360	S355				Disord		
			AQPAPPRLMRpSKGTYpSPSSSETVR	1 pS + 1 pS	727-750		S736	CK2 (0.5467)		Disord		
							S738	CK2 (0.5051)		Disord		
						S742		GSK3 (0.5431), CK1 (0.5274)		Disord	uk	15345747
IPI00320690	Q61136	Serine/threonine-protein kinase PRP4 homolog	VQSGMGLILQGYEpSQpSEEHGEIKEK	2 pS	130-141	S143		CK2 (0.5001)		Ord		
						S145		CK2 (0.4893)		Ord		
			SKDApSPINRWpSPTR	2 pS	427-440	S431				Ord	uk	15302935
						S437		Cdk5 (0.1606), Cdc2 (0.2149)		Ord	uk	15302935
IPI00204379	Q9WV48	Shank1	pSMpSVPDDAHYSMMVYR	2 pS	61-76	S61				Ord		
						S63		Akt (0.4455)	14-3-3 m1 (0.2317)	Ord		
			HYTVGpSYDpSFDAPSDYIIK	2 pS	633-651	S638				Disord		
						S641				Disord		
IPI00125582	Q5RIT9	Similar to PITPNM family member 3	GpSPPLLDAPApSPPQAPR	2 pS	494-510	S495		CK1 (0.5252), Cdk5 (0.5566)		Ord		
						S504		ERK1 (0.4447), p38 MAPK (0.4165)		Disord		
IPI00396130	Q9QWI6	SNAP-25-interacting protein (SNIP)	RGpSDELTVPR	1 pS	1052-1061	S1054				Disord	uk	15572359
IPI00008513	Q8K004	Spata2	pSVDATDSTWESR	1 pS	247-258	S247		CaMKII (0.5860)		Disord		
IPI00139364	Q8VE97	Splicing factor, arginine/serine-rich 4	pSRpSTpSKSPNVPARSR	3 pS	439-454	S439			14-3-3 m1 (0.4017)	Disord		
						S441			14-3-3 m1 (0.3151)	Disord		
						S443		Akt (0.4682)		Disord		
IPI00229604	Q8VCD2	SR-related CTD associated factor 6	SRpSPIPPpSAAGLGSNSAPPIPDSR	3 pS/T	822-845		S822	Clk2 (0.4208), Akt (0.4276)	14-3-3 m1 (0.3150)	Disord		
							S824	Clk2 (0.4799), Akt (0.4409)	14-3-3 m1 (0.3855)	Disord		
							S826	Akt (0.4835)	14-3-3 m1 (0.3246)	Disord		
							S829			Disord		
IPI00136372	O88935	Synapsin I	GSKSQSRPGALTLGR	2 pS	431-446					Ord		
			PVAGSPGAPPAARPPApSPpSPQR	2 pS	535-556	S551		ERK1 (0.3734), GSK3 (0.5927), Cdk5 (0.5305)		Ord	ERK1	8622996
						S553		Cdk5 (0.2065), Cdc2 (0.2899)		Ord	Cdk5	8702879
IPI00296573	Q8CC35	Synaptopodin	VApSREEVPLVVYLK	1 pS	256-270	S258		Clk2 (0.6714), CK2 (0.4809)		Ord		
IPI00128775	Q9R1L5	Syntrophin-associated serine-threonine-protein kinase	VGpSpSNTER	2 pS	1226-1233	S1228				Ord		
						S1229				Ord		
IPI00359917	Q99J36	THUMP domain containing protein 1	FIDKDQQPpSQpSEMGKDDDAEAALKK	2 pS	78-101	S86				Ord		
						S88		CK2 (0.5447)		Ord		
IPI00359323	Q5M7V8	Thyroid hormone receptor associated protein 3	ASVSDLpSPR	1 pS	237-245	S243		Cdk5 (0.4666)		Disord	uk	15302935
			SRpSPALKpSPLQSVVVR	2 pS	246-261	S248				Disord		
						S253				Disord		
			IDIpSPSTFR	1 pS	676-684	S679				Disord	uk	15302935
			pSRSSMWDPKTTPK	1 pS	860-872	S860				Disord		
IPI00372166	Q9UKE5	TRAF2 and NCK interacting kinase	VRANpSKSRQpSPVLPHPPSK	2 pS	760-778	S764		Akt (0.5328), Clk2 (0.6361)		Disord	uk	15144186
						S769				Disord	uk	15144186
IPI00119067	Q7TQD2	Tubulin polymerization-promoting protein	AAMKpTPPKpSPGDPAR	1 pS + 1 pT	11-25	T15		Cdk5 (0.4299), Cdc2 (0.4955), Akt (0.5155)		Disord	uk	15572359
						S19		Cdk5 (0.4912), ERK1 (0.4868). Cdc2 (0.4870)		Disord	uk	15572359
			RLpSLRpSRQSANRGATAAPRLSALRHAPR	2 pS	29-55	S31		PKA (0.3242), PKC delta (0.4079), CaMKII (0.5936)		Ord		
						S34				Ord		
IPI00021840	P62754	40S ribosomal protein S6 (RPS6)	RLSpSLR	1 pS	233-238	S236		PKC a/b/g (0.3915), PKC delta (0.3915), Akt (0.4385)	14-3-3 m1 (0.5382)	Disord	S6 kinase	1939282
IPI00123624	Q8VDP4	2610301G19Rik protein	pSVApSMQpSpSMMSYSSLQCMPK	3 pS	674-692	S674				Disord		
						S677				Disord		
						S680				Disord		
IPI00110623	Q9CXN5	3110050K21Rik protein	RKLpSPpSPSLR	2 pS	204-213	S207		ERK1 (0.5404)		Disord		
						S209		Cdk5 (0.4583), Cdc2 (0.5197)		Disord		

Appendix 5 - Non-redundant phosphopeptide list with unique phosphorylation site assignation for the cytosolic phosphoproteome. Unambiguously assigned sites of phosphorylation are denoted in the sequence as pS/T/Y (red) and ambiguous sites are coloured in rblue. Mod Id: number of phosphates detected. MS identified Phospho-sites: unambiguous identification of sites by MS. Predicted Phospho-sites: when sites could not be unambiguously assigned, Scansite was used to predict the most probable site of phosphorylation and are indicated in the sequence in red. Scansite kinase: Kinases predicted by Scansite are indicated for all sites. Also, in some cases Scansite predicted some phosphorylation sites to be involved in phosphorylation dependent protein interactions. Order: Prediction of intrinsic sequence disorder in regions surrounding phosphorylation sites was performed using DisEMBL. Known phosphorylation sites and cognate kinases are indicated and the relevant PMID number is in the “Literature” field, PMID’s coloured red correspond to our synapse phosphoproteome dataset.

Appendix 6

Protein Class	PSD #	MASC #	AMPA #	Swall Acc	PPID Acc	Abbreviated Name	Mr [kDa]	pI	1D SDS-PAGE LC-MS/MS analysis of PSD			1D SDS-PAGE LC-MS/MS analysis of AMPAR IP			1D SDS-PAGE LC-MS/MS analysis of NRC IP				1D SDS-PAGE LC-MS/MS analysis of MAGUK-pep		
									PSD highest occurance in gel slice #	PSD Peptides Occurrence by MS	PSD # of gel slices protein was present in	AMPAR highest occurance in gel slice #	AMPAR # of gel slices protein was present in	AMPA Peptides Occurrence by MS	Whole extract WB signal Extract	NRC IA WB	NRC IP WB	NRC - Peptides Occurrence by MS	NRC # IgG approach	MAGUK-pep WB result	MAGUK-pep MS-results	MAGUK-pep # peptide approach
Channels and Receptors																					
Glutamate Receptors	1		1	P23818	A0296	AMPA-R1	100	7.86	32	1	2	5	7	93	+++	-	-			-		
	2		2	P23819	A0297	AMPA-R2	96	6.85	32	26	3	11	8	139	+++	-	-			-		
	3		3	Q9Z2W9	A0299	AMPA-R3	100	8.48	32	9	1	2	1	13								
			4	P19493	A0300	AMPA-R4	98	7.59				2	1	12	+++	-	-			-		
		1		P39087	A0022	GLUR6	Frag	Frag							+++	+	+		1	+		1
	4	2		Q9EPU6	A0103	MGLUR1	131	6.45	42	1	1				+++	+++	+++		2	++		2
	5			Q9QYS2	A1353	MGLUR3	97	7.65	42	4	2											
	6	3		P31424	A0327	MGLUR5	130	8.09	42	3	1				+++		+		3	+		3
	7			P35400	A0453	MGLUR7	99	8.04	42	1	1											
	8	4		Q62683	A0001	NR1	106	8.99	33	13	2				+++	+++	+++	208	4	+++	9	4
	9	5		Q08948	A0002	NR2A	165	6.59	34	7	2				+++	+++	+++	36	5	+++	2	5
	10	6		Q62684	A0003	NR2B	163	6.33	37	17	5				+++	+++	+++	85	6	+++	13	6
Other ligand receptors	11			Q99917	A0039	CIRL-1	164	6.17	34	2	1											
	12			Q9Z173	A3463	CIRL-3	172	6.23	34	1	1											
	13			P18504	A0504	GABA-A-R1	49	9.21	6	3	1											
	14			Q9WV18	A0603	GABA-B-R1	106	8.42	34	4	3											
	15			O88871	A2000	GABA-B-R2	102	8.87	33	5	2											
	16			O55022	A3464	PGRMC	22	4.57	20	1	1											
Ca2+-ATPases	17			O55143	A3467	ATP2A2	115	5.23	32	39	9											
	18			P11505	A3465	ATP2B1	139	5.71	42	2	1											
	19			Q9R0K7	A0508	ATP2B2	133	5.7	39	10	4											
	20			Q64568	A3466	ATP2B3	139	5.67	40	2	1											
	21	7		Q64542	A0114	ATP2B4	133	6.17	35	8	5				+++	+	+		7	+		7
NA+/K+-ATPases	22	8	5	P06685	A0340	ATP1A1	113	5.27	35	26	5	7	1	2				1	8		1	8
	23	9		Q9Z1G8	A0342	ATP1A3	112	5.26	39	89	19							1	9		1	9
	24			Q9WV28	A0343	ATP1A4	115	5.52	35	29	10											
	25			P14094	A0344	ATP1B1	35	8.83	6	9	3											
Voltage- and Ligand-gated Ca2+-channel	26			P54282	A0593	CACNA1A	252	8.77	42	1	1											
	27			O55017	A0592	CACNA1B	261	8.85	42	1	1											
	28			Q01815	A0770	CACNA1C	240	6.45	42	1	1											
	29			Q61290	A3471	CACNA1E	257	8.47	42	7	1											
	30			O08532	A3474	CACNA2D1	17	4.6	35	29	3											
	31			Q9EQQ2	A3472	CACNA2D2	131	5.43	36	9	3											
	32			Q9Z1L5	A3473	CACNA2D3	123	5.5	34	10	1											
	33			P54285	A3476	CACNB3	55	5.93	2	1	1											
Voltage-gated K+-channel	34			Q9Z343	A3477	KCNQ2	84	9.72	31	1	1											
	35			P15385	A0764	KV1.2	57	4.77	30	11	3											
	36			Q64284	A3478	KV2.2	41	9.11	11	11	2											
Voltage-dependent anion channels	37	10		Q60932	A0426	VDAC1	32	8.55	15	75	6										3	10
	38	11		Q60930	A0426	VDAC2	32	7.44	15	49	7										2	11
	39			Q60931	A3479	VDAC3	31	8.96	17	30	3											
Na+-channel	40			Q99250	A0598	SCN2A2	226	5.56	41	5	2											
Ca2+-release channels	41			P11881	A0154	INSP3R	313	5.72	41	7	1											
	42			Q8ERN6	A3480	RYR2	565	5.75	42	10	2											
Other Channels and Receptors	43			P12023	A1169	APP	85	4.7	33	1	1											
	44			Q9Z1G4	A0414	ATP6N1	97	6.29	33	65	11											
	45			Q9JI12	A0409	ATP6V1H	56	6.07	5	2	1											
	46			Q08406	A3490	CNTFR	36	6.47	1	2	1											
	47			O08530	A3491	EDG-1	43	9.52	11	1	1											
	48			O35658	A1608	GC1Q-R	24	4.41	15	2	2											
	49			Q03555	A0637	GEPHYRIN	83	5.53	31	3	1											
	50			O89703	A3486	HAC-1	95	8.97	34	2	1											
	51			Q61291	A1175	LRP	505	5.17	42	3	1											
	52			P70206	A3482	PLEXIN-1	211	6.48	39	9	1											
	53			P70207	A3483	PLEXIN-2	210	6.05	38	3	1											
	54			Q91V14	A3489	SLC12A5	124	5.88	37	17	2											
	55			Q62634	A4034	SLC17A7	62	6.96	3	1	1											
	56	12		P43006	A0422	SLC1A2	62	6.24	1	5	3										1	12
	57			Q9ESI0	A3481	SLC4A10	123	5.95	38	1	1											
	58			P16283	A3488	SLC4A3	135	5.94	42	3	2											
	59			O88343	A3485	SLC4A4	121	6.31	34	2	1											
	60			P48768	A3484	SLC6A2	99	5	37	2	2											
MAGUKs / Adaptors / Scaffolders																						
PDZ-domain containing scaffolders	61			Q9QYX6	A0568	ACZONIN	548	6.07	42	26	3											
	62			O70589	A0126	CASK	105	6.16	32	4	1											
	63	13		Q63622	A0014	CHAPSYN-110	95	5.84	32	11	3				+++	+++	+++	108	10	+++	63	13
	64	14		Q9WV34	A0436	DLGH2	62	6.03	2	12	1										2	14
	65	15		O88910	A0435	DLGH3	64	5.81	27	1	1										4	15
	66	16		Q9Z216	A0104	HOMER	40	5.33	6	25	2				+++	++	++		11	++		16
	67			Q9R093	A0137	MAGUIN	102	6.69	34	2	1											
	68			Q9JLU4	A0074	PROSAP2	193	8.98	39	5	2											
	69	17		Q62108	A0013	PSD-95	80	5.56	32	36	4				+++	+++	+++	112	12	+++	122	17
	70	18		P70175	A0016	SAP102	93	6.32	32	12	2				+++	+	+	9	13	+++	75	18
	71	19		Q62402	A0015	SAP97	103	5.83	34	2	1							2	14		39	19
	72	20		Q9WU13	A0075	SHANK1	226	8.52	41	1	1				+++	+++	+++		15	+++		20
	73			Q9WV37	A0793	VAM-1	63	5.99	3	2	1											
	74	21		P39447	A0086	ZO-1	195	6.2	40	1	1							13	16			
Non-PDZ-domain containg scaffolders				Q9WTW1	A0322	AIP	112	6.29							+++	-	-			-		
	75			P25304	A3494	AGRIN	205	5.73	42	2	1											
	76	22		P24587	A0143	AKAP150	47	4.84	36	1	1				+++	+++	+++		17	+++		21
		23		Q8K3H0	A0188	APPL	79	5.28										8	18			
	77			Q60437	A1015	BAIAP2	58	8.88	3	7	2											
	78			Q8VHK2	A3990	CASKIN1	150	9.25	37	4	1											
	79			Q60902	A3493	EPS15R	99	4.89	34	8	1											
	80			O08917	A3496	FLOTILLIN-1	48	6.71	6	13	1											
	81			Q60634	A3497	FLOTILLIN-2	42	5.17	7	15	2											
		24		P97841	A0107	GKAP	110	6.86							+++	+++	++	4	19	+++	2	22
		25		Q60631	A0126	GRB2	25	5.89							+++		++		20			

Appendix 6

Protein Class	PSD #	MASC #	AMPA #	Swall Acc	PPID Acc	Abbreviated Name	Mr [kDa]	pI	1D SDS-PAGE LC-MS/MS analysis of PSD			1D SDS-PAGE LC-MS/MS analysis of AMPAR IP			Whole extract WB signal Extract	1D SDS-PAGE LC-MS/MS analysis of NRC IP				1D SDS-PAGE LC-MS/MS analysis of MAGUK-pep		
									PSD highest occurance in gel slice #	PSD Peptides Occurrence by MS	PSD # of gel slices protein was present in	AMPAR highest occurance in gel slice #	AMPAR # of gel slices protein was present in	AMPA Peptides Occurrence by MS		NRC IA WB	NRC IP WB	NRC - Peptides Occurrence by MS	NRC # IgG approach	MAGUK-pep WB result	MAGUK-pep MS-results	MAGUK-pep # peptide approach
non-PDZ-domain containg scaffolders				Q925T6	A0301	GRIP	122	6.1							+++	-	-			-		
		24		O35145	A0162	JIP-1	77	4.85							++++		+++		21	+		23
	82			Q9EQR5	A3495	LIME	30	9.53	38	1	1											
	83			Q9Y2A7	A0545	NAP-1	129	6.18	33	16	2											
	84			Q9EPM9	A3492	NEUROBEACHIN	327	5.8	41	4	1											
		27		P25388	A0124	RACK1	35	7.6							++	cl	cl			+		24
	85			O94938	A4033	RC3	340	5.93	41	8	1											
		28		Q99P24	A0009	YOTIAO	454	4.95							+++	+++	+++		22	+++		25
14-3-3	86			P35213	A0467	14-3-3 beta	28	4.81	18	7	1											
	87	29		P42655	A0280	14-3-3 epsilon	29	4.63	17	15	1				+++	cl	cl			+	5	26
	88	30		P11576	A0633	14-3-3 eta	28	4.81	18	13	3										3	27
	89	31		P35214	A0362	14-3-3 gamma	28	4.8	18	17	2				+++	+	+		23	+	3	28
	90			P35216	A0907	14-3-3 theta	28	4.69	18	3	1											
	91	32		P35215	A0361	14-3-3 zeta	28	4.73	18	17	2										3	29
Ser/Thr Kinases																						
	92			Q9UPV4	A3502	AAK1	94	8.72	34	1	1											
	93	33		Q60823	A0466	AKT2	56	5.98	5	2	1				+++		ch			+		30
	94	34		P11798	A0011	CAMKII alpha	54	6.88	14	94	20				+++	+++	+++	4	24	+++	2	31
	95	35		P28652	A1851	CAMKII beta	60	6.87	2	48	6				+++	+++	+++	1	25	+++	1	32
	96			P15791	A1853	CAMKII delta	60	6.84	3	2	1											
	97			P11730	A1852	CAMKII gamma	59	7.63	26	2	2											
	98			O54875	A3504	CDC42BPB	194	6.38	37	4	1											
	99			P97633	A0896	CK-1	37	9.55	13	1	1											
	100			O60737	A0753	CK-2 alpha	45	7.79	9	4	1											
	101			Q99PS2	A4036	CSNK1E	57	9.42	8	2	1											
	102			Q9JLM8	A3506	DCAMKL1	84	9	30	13	2											
	103			Q05512	A3505	EMK	86	9.7	29	6	2											
		36		Q91YW5	A0265	ERK1	43	6.28							+++	++	++		26	++		33
		37		P27703	A0181	ERK2	41	6.5							+++	++	++		27	++		34
				Q61532	A0321	ERK3	62	5.05							+++	-	-			-		
	104			P26817	A0813	GRK2	80	6.54	29	1	1											
		38		Q9WV60	A0222	GSK3 BETA	47	8.98							+++	++	++		28	++		35
				P47809	A0322	JNKK1	44	8.28							+++	-	-			-		
		39		Q09110	A0491	MAP2K3	39	7.05							+++		+		29	+		36
				Q61084	A0490	MAP3K3	71	9.02							+++		-			-		
	105			Q9JM92	A3503	MAP4K6	147	7.07	37	1	1											
				P47811	A0650	MAPK14	41	5.55							+++		-			-		
		40		Q61831	A0489	MAPKP49	53	6.51							+++		+		30	+		37
		41		P31938	A0266	MEK1	43	6.26							+++	++	++		31	+		38
		42		Q63932	A0267	MEK2	44	6.56							+++	++	++		32	+		39
		43		O35406	A0164	MKK7	39	8.85							+++		ch			+	3	40
		44		Q62767	A0287	MKP2	43	7.51							+++	+++	+++		33	++		41
	106			Q63092	A3499	NOVEL CAMK	54	5.37	27	9	3											
				Q9CST0	A0337	P70-S6K	56	6.02							+++		-			-		
		45		Q9Z2A0	A0260	PDK-1	64	6.95							+++		ch			+		42
				P12849	A0303	PKA-R1 ALPHA/BETA	43	5.71							+++	ch	ch			-		
				Q9DBC7	A0304	PKA-R1a-ALPHA	43	5.27							+++	ch	ch			-		
				P12367	A0305	PKA-R2 ALPHA	45	4.78							+++	ch	ch			-		
	107	46		P12369	A0306	PKA-R2B	46	4.9	5	3	1				+++	++	++		34	++		43
				P20444	A0307	PKC ALPHA	77	6.73							+++	-	-			-		
	108	47		P04411	A0168	PKC BETA	77	6.57	29	8	1				+++	++	++		35	++		44
				P28867	A0308	PKC DELTA	78	7.2							+++	-	-			-		
	109	48		P16054	A0144	PKC EPSILON	84	6.62	31	2	1				+++	++	++		36	++		45
	110	49		P05697	A0259	PKC GAMMA	78	7.27	29	13	1				+++	+++	+++		37	+++		46
				Q62074	A0311	PKC IOTA	67	5.53							+++	-	-			-		
				O94806	A0309	PKC NU	100	6.42							+++	-	-			-		
				O02111	A0310	PKC THETA	82	7.88							+++	-	-			-		
	111	50		P05206	A4037	PRKACB	41	8.66	10	4	2				+++	++	++		38	++		47
	112			Q61410	A3500	PRKG2	87	8.44	30	1	1											
		51		Q99N57	A0138	RAF1	73	9.3							+++	++	++		39	+		48
	113			P70336	A3501	ROCK-2	161	5.73	36	5	1											
		52		P18654	A0268	RSK	84	6.41							+++	++	+++		40	++		49
				Q9Z1B7	A0475	SAPK4	42	8.67							+++		-			-		
	114			Q9WUN2	A3507	TBK1	83	6.4	29	1	1											
Tyr Kinase																						
				P34152	A0313	FAK	124	6.25							+++	-	-			-		
	115			P16879	A3508	FES	94	6.35	11	2	1											
	116			P39688	A0019	FYN	60	5.78	3	6	1				+++	-	-			-		
		53		Q9QVP9	A0030	PYK2	116	5.99							+++	+	+		41	+		50
	117	54		P05480	A0018	SRC	60	7.99	3	7	1				+++	+	+		42	+		51
Protein Phosphatases																						
	118			Q8VEL2	A4016	FLJ90311	59	6.18	26	1	1											
	119	55		P20652	A0123	PP2B	59	5.58	2	6	1				+++	+	+	3	43	+		52
		56		O35299	A0289	PP5	57	5.83							+++	++	++		44	++		53
	120			Q9Z1Q2	A0147	PPP1CA	38	5.94	11	4	1											
	121			P37140	A4043	PPP1CB	37	5.84	11	1	1											
	122	57		P37139	A4042	PPP1CC	37	6.12	12	3	1				+++	+++	+++		45	+++		54
	123	58		P13353	A0132	PPP2CA	36	5.3	12	4	1				+++	+++	+++		46	+		55
	124	59		Q96DH3	A0295	PPP2R1A	65	5	2	13	1							1	47			
				P35821	A0278	PTP1B	50	5.78							+++	ch	ch			-		
				P29351	A0315	PTP1C	68	7.66							+++	-	-			-		
		60		P35235	A0020	PTP1D	67	6.51							+++	++	++			++		56
		61		P54830	A0118	PTPN5	61	5.11							++	(p61)	(p61)		48	(p61)		57
	125			Q64487	A1198	PTPRD	210	6.11	31	2	1											
	126			Q9EQ17	A0223	PTPRF	212	6.16	30	1	1											
	127			Q9QW00	A2624	PTPRS	168	6.44	29	37	3											
	128			Q9WUT8	A1202	PTPZETA	172	4.5	42	16	2											
	129			O60228	A3509	SBF-1	208	6.75	38	7	1											

Appendix 6

Protein Class	PSD #	MASC #	AMPA #	Swall Acc	PPID Acc	Abbreviated Name	Mr (kDa)	pI	1D SDS-PAGE LC-MS/MS analysis of PSD: PSD highest occurance in gel slice #	PSD Peptides Occurrence by MS	PSD # of gel slices protein was present in	1D SDS-PAGE LC-MS/MS analysis of AMPAR IP: AMPAR highest occurance in gel slice #	AMPAR # of gel slices protein was present in	AMPA Peptides Occurrence by MS	Whole extract WB signal Extract	1D SDS-PAGE LC-MS/MS analysis of NRC IP: NRC IA WB	NRC IP WB	NRC - Peptides Occurrence by MS	NRC # IgG approach	1D SDS-PAGE LC-MS/MS analysis of MAGUK-pep: MAGUK-pep WB result	MAGUK-pep MS-results	MAGUK-pep # peptide approach
G-proteins and Modulators																						
G-proteins	130			Q9JHW8	A3514	CENTG1	78	10.49	35	10	2											
	131			Q96HU8	A4017	DI-RAS2	22	8.98	20	1	1											
	132			P08752	A0055	GNAI2	40	5.28	10	31	3											
	133			P18872	A0043	GNAO1	40	5.34	10	52	3											
	134			P21279	A0042	GNAQ	41	5.66	10	4	2											
	135	82		P04894	A0051	GNAS1	46	5.69	6	6	2										1	58
	136			O70443	A0052	GNAZ	41	7.66	10	15	1											
		83		Q9QWG8	A0069	GNB1	37	5.47													1	59
	137	84		P54312	A0070	GNB2	37	5.6	13	32	4										2	60
		85		P29387	A0072	GNB4	37	5.59													1	61
	138			P54311	A0073	GNB5	39	5.67	9	2	1											
		86		Q61411	A0200	H-RAS	21	5.16							+++	cl	cl			+		62
	139			Q9H0I7	A3974	MIRO1	67	5.56	29	7	2											
	140			O08989	A3512	M-RAS	24	8.81	20	2	1											
				Q9JLX1	A0619	RAB11	24	6.12							++	cl	cl			-		
	141			P35284	A3513	RAB12	Frag	Frag	18	1	1											
		87		P53994	A0363	RAB2	24	6.08													1	63
				Q6E0I2	A0360	RAB27	25	5.21							+	cl	cl			-		
	142	88		P05713	A0037	RAB3	25	4.85	20	12	3				+++	cl	cl			++		64
		89		Q9JKM7	A0365	RAB37	25	5.53													1	65
	143			O88565	A0359	RAB-5	24	8.32	20	3	1				+	cl	cl			-		
		70		Q9JJD4	A0364	RAB6A	24	5.42													1	66
	144			P55258	A0620	RAB-8	24	9.26	19	2	1											
		71		P15154	A0194	RAC1	21	8.77							+++	++	+		49	+		67
	145	72		P05810	A0203	RAL-A	24	6.66	20	5	1				+++	cl	cl			+		68
		73		P28746	A0152	RAN	24	7.01							+++	cl	cl			++	3	69
				P10113	A0323	RAP1	21	6.39							+++	-	-			-		
		74		Q9D3D5	A0033	RAP2A	12	5.4							+++	++	++		50	+		70
	146			Q8NVA2	A3524	SEPT	49	6.36	5	11	1											
	147			P42208	A3515	SEPT2	42	6.1	9	2	2											
	148			Q9Z1S5	A3516	SEPT3	53	7.45	9	6	1											
	149			P28661	A3517	SEPT4	55	5.58	5	2	1											
	150			Q9Z2Q6	A3518	SEPT5	43	6.21	9	22	3											
	151			Q9R1T4	A3519	SEPT6	50	6.03	4	4	2											
	152			O55131	A3520	SEPT7	51	8.73	6	21	3											
	153			Q9ESF7	A3525	SEPT8	Frag	Frag	2	17	3											
	154			Q9QZR6	A3521	SEPT9	64	8.65	27	1	1											
Modulators	155			Q8ESG7	A0843	ARHGEF2	112	7.06	33	5	1											
	156			O08757	A0646	ARHGEF7	80	6.45	30	1	1											
	157			Q9Z1P0	A3528	CGEF2	115	6.45	32	10	1											
	158	75		O88938	A0276	CITRON	235	6.1	37	14	1				+++	+++	+++		51	++		71
	159			Q9JL26	A4046	FMN2	Frag	Frag	35	1	1											
		76		Q9TZT5	A0470	FRABIN	87	5.65							+++		++		52	+		72
	160	77		P06837	A0215	GAP43	24	4.63	6	10	3				+++	++	+		53	+	1	73
	161			Q9Z272	A3532	GIT1	85	6.47	31	2	1											
	162	78		Q9JIF2	A0214	KALIRIN	271	5.7	37	7	1				+++		+		54	+		74
	163			Q9D0B5	A2428	KIAA0783	92	7.31	34	1	1											
	164			Q9C0H5	A3530	KIAA1686	121	7.3	35	1	1											
	165	79		Q04690	A0196	NF1	320	7.1	40	1	1				+++	+	+		55	+		75
	166			O94764	A4047	RANBP9	78	6.31	30	1	1											
				Q91YX7	A0025	RASGAP	94	6.72							+++	-	-			-		
	167			O54829	A0932	RGS7	55	7.68	3	6	1											
	168	80		Q9ESK6	A0024	SYNGAP	143	9.09	34	44	5				+++	++	++	9	56	++		76
Signalling molecules and Enzymes																						
Heat shock / Chaperonins / Chaperonins	169			P20029	A0451	BiP	70	5.01	28	22	3											
	170			P11983	A3541	CCT1	60	5.82	2	16	1											
	171			P80314	A3540	CCT2	57	5.97	4	6	2											
	172			P80318	A3544	CCT3	61	6.28	1	34	2											
	173			P80315	A3538	CCT4	58	8.24	3	8	1											
	174			P80316	A3543	CCT5	60	5.72	2	14	1											
	175			P80317	A3545	CCT6A	58	6.63	2	14	2											
	176			P80313	A3539	CCT7	60	7.95	3	8	1											
	177			P42932	A3542	CCT8	60	5.44	2	15	1											
	178			P54102	A3536	DNAJA1	46	6.65	6	2	1											
	179			Q9QYJ0	A3535	DNAJA2	46	6.09	6	5	1											
	180			P08113	A2004	ENDOPLASMIN	90	4.72	31	6	1											
	181			P08109	A0146	HSC-71	71	5.37	27	41	4											
	182			Q61699	A3533	HSP105	96	5.42	33	1	1											
	183			P19326	A3550	HSP90	58	5.35	2	5	1											
	184			P38647	A2010	HSP74	69	5.5	28	11	2											
	185			P11499	A2003	HSP90	83	4.97	30	11	2											
	186	81		P17879	A2007	HSPA1	70	5.46	30	1	1				+++	++	++	24	57	+++	7	77
	187			O43301	A3534	KIAA0417	75	6.32	28	12	3											
	188			O54946	A1578	MRJ	27	6.99	19	3	1											
	189			Q9NVH1	A3537	DNAJC11	63	8.54	1	11	1											
	190			P38660	A3548	PDIA6	46	4.99	5	1	1											
	191			Q99MP2	A3549	SSR1	32	4.36	12	2	1											
Phosphodiesterases	192			P16330	A3553	CNP-1	47	9.08	7	69	4											
	193			Q9QYJ5	A3552	PDE10A	89	6.06	29	3	1											
	194			Q01062	A3551	PDE2A	105	5.28	31	11	4											
Adenylate/Guanylate cyclase	195			Q9WVI4	A1979	GUCY1A2	82	7.78	29	5	1											
Secreted / secretory / signalling	196			Q61361	A3556	BCAN	94	4.83	42	3	1											
	197			Q62059	A3555	CSPG-2	365	4.56	42	6	1											
	198			O88898	A3557	OLFM1	54	6.33	26	2	2											
	199			O35282	A1304	SEC-8	111	6.07	32	3	2											
ATP synthases	200	82		Q03265	A0377	ATP5A1	55	8.26	5	22	3										11	78
	201			P56480	A0378	ATP5B	52	4.99	5	22	3											
	202	83		P35435	A0379	ATP5C	30	8.87	16	9	1										3	79
	203	84		P50516	A0397	ATP6A1	68	5.62	27	31	4										4	80

Appendix 6

Protein Class	PSD #	MASC #	AMPA #	Swall Acc	PPID Acc	Abbreviated Name	Mr [kDa]	pI	1D SDS-PAGE LC-MS/MS analysis of PSD			1D SDS-PAGE LC-MS/MS analysis of AMPAR IP			1D SDS-PAGE LC-MS/MS analysis of NRC IP					1D SDS-PAGE LC-MS/MS analysis of MAGUK-pep		
									PSD highest occurance in gel slice #	PSD Peptides Occurrence by MS	PSD # of gel slices protein was present in	AMPAR highest occurance in gel slice #	AMPAR # of gel slices protein was present in	AMPA Peptides Occurrence by MS	Whole extract WB signal Extract	NRC IA WB	NRC IP WB	NRC - Peptides Occurrence by MS	NRC # IgG approach	MAGUK-pep WB result	MAGUK-pep MS-results	MAGUK-pep # peptide approach
ATP synthases	204			P50517	A0400	ATP6B2	57	5.57	4	15	2											
	206			Q9Z1G3	A0401	ATP6C	44	6.36	9	12	2											
	206	86		P51863	A0403	ATP6D	40	4.89	11	14	2										1	81
	207			P50516	A0404	ATP6E	27	9.28	16	21	2											
	208			P57746	A0402	ATP6M	28	9.44	15	9	2											
	209			Q92501	A3559	TO83	67	9.32	1	5	1											
NADH-Ubiquinone Oxidoreductase	210			Q99LC3	A3560	NDUFA10	37	5.96	9	5	1											
	211			Q16795	A3563	NDUFA9	39	9.68	13	9	2											
	212			P28331	A3561	NDUFS1	77	5.36	29	24	1											
	213			Q91WD5	A3562	NDUFS2	53	6.52	7	9	1											
	214			Q9DC72	A3564	NDUFS3	26	5.45	19	9	2											
	215	86		Q9D6J6	A0431	NDUFV2	24	5.31	20	1	1										2	82
Polyubiquitin	216			Q62317	A1257	UBC	14	9.33	40	8	7											
Proteases	217			O88543	A3570	COPS-3	46	6.19	9	1	1											
	218			P46101	A3565	DPP-6	97	5.95	30	12	3											
	219			Q9EQF2	A3571	KELL	81	5.89	35	1	1											
	220			O35593	A3566	PAD-1	35	6.18	13	2	1											
	221			Q9R1P0	A3567	PSMA-4	29	7.58	18	1	1											
	222			O88761	A3568	PSMD-1	106	5.25	32	2	1											
	223			Q99J85	A3569	PSMD-8	30	6.03	16	1	1											
	224			O88399	A1731	SYNJ2	164	8.05	32	1	1											
Development	226			Q05175	A3574	BASP1	22	4.5	2	7	2											
	226			P18572	A3573	BSG	28	5.36	6	5	2											
	227			Q8CYN5	A3575	CBCP1	39	6.76	11	5	2											
	228			Q9JIQ6	A3998	CNNM1	66	5.41	34	1	1											
	229			O88058	A3579	CYFIP1	145	6.53	34	17	1											
	230			Q6CQQ2	A4001	DAAM1	123	7.12	33	3	1											
	231			Q9QXL0	A3578	DBCCR-1	89	9.22	31	2	1											
	232			Q9D2N4	A3577	DTNA	84	6.29	29	1	1											
	233			O88741	A3576	GDAP1	41	8.51	11	5	2											
	234			Q8VE33	A3980	GDAPTL1	42	6.13	9	1	1											
	235			Q9QZY2	A3581	GPRIN1	85	8.63	34	3	1											
	236			Q62813	A3572	LSAMP	32	5.53	1	29	6											
	237			O08919	A2535	NUMBL	64	9.07	28	4	2											
Other signalling molecules				P34971	A0115	ADRB1	50	9.12							+++		-			-		
		89		Q9WV31	A0277	ARC	45	5.3							+++	+++	+++		58	+++		84
		87		P16587	A0421	ARF3	20	7.04													1	83
		88		Q61337	A0476	BAD	22	9.15							+++	cl	cl			+		85
		90		P02593	A0008	CALM	17	4.09							+++	+++	+++		59	+++		86
				O35350	A0317	CALPAIN	82	5.62							+++	-	-			-		
		91		P47728	A0437	CALRETININ	31	4.94													1	87
	238			Q99KW9	A4005	CDA08	67	5.4	30	1	1											
				P11440	A0448	CDC2	34	8.4							++	-	-			-		
				Q05769	A0485	COX-2	67	6.86							+++		-			-		
		92		P47713	A0106	CPLA2	85	5.28							+++	+++	++		60	++		88
				Q60829	A0472	DARPP-32	22	4.6							+++		-			-		
	239	93		Q61358	A0095	DNM1	98	7.61	31	55	4				+++	++	++	4	61	+++		89
				P70313	A0316	E.NOS	133	6.62							+++	-	-			-		
		94		Q00731	A0434	EGF-1ti4	22	9.25													1	90
		95		P35637	A0217	FUS	53	9.4										1	62			
		96		P53762	A0488	HF-1A	87	6.19							+++		+		63	+		91
				O70262	A0487	HO-2	36	5.61							+++		-			-		
				P15208	A0195	INSR	83	5.79							+++	-	-			-		
		98		P35569	A0484	IRS-1	131	8.97							+++		+		64	+		93
				Q99M51	A0128	NCK	43	6.06							+++	-	-			-		
		99		Q9Z0J4	A0262	NNOS	160	6.73							+++	++	++		65	++		94
	240			O70465	A4013	OOZ4	313	6.62	41	1	1											
		100		Q12888	A0286	P53BP1	214	4.82										2	66			
		101		Q62077	A0012	PLC G	149	5.73							+++	++	++		67	++		95
		102		Q9Z1B3	A0330	PLCB-1	138	5.67							+++		+		68	++		96
	241			P04925	A4012	PRNP	23	9.56	15	6	5											
				P98083	A0171	SHC	63	6.09							+++	-	-			-		
		103		Q9WQ36	A4014	SPINK5															1	97
		104		O88602	A0473	STARGAZIN	36	9.13							+++		+		69	-		
				Q9Z108	A0358	STAUFEN	54	9.55							+	-	-			-		
		97		Q9Z1W9	A0166	STK39	60	5.92													1	92
	242			Q9WUD1	A3978	STUB1	35	5.71	13	3	1											
				P27870	A0173	VAV	98	6.11							+++	-	-			-		
Other enzymes	243			P17764	A3694	ACAT1	41	8.43	9	3	1											
	244			Q925C4	A3582	ACC1	265	6.1	39	2	2											
	245			Q91V92	A3586	ACLY	120	7.13	33	10	1											
	246	106		Q99KI0	A0429	ACO2	85	8.08	30	10	1										1	98
	247			Q9D1E8	A3630	AGPAT5	42	9.41	13	3	1											
	248			P24549	A3585	ALDH1A1	54	7.91	33	1	1											
	249			P05064	A1923	ALDOA	39	8.4	9	28	2											
	250	106		P05063	A0428	ALDOC	39	6.79	10	2	1											
		107		Q9QWW0	A0432	AOP2	25	5.72													1	100
	251			O75061	A0685	AUXILIN	100	6.82	33	2	1											
	252			P29147	A3702	BDH	33	8.41	16	3	1											
	253			Q99PU5	A3593	BG1	80	5.67	29	2	1											
	254			P30275	A3690	CKMT1	43	7.31	8	21	3											
	255			P17427	A3605	CRMP1	62	6.64	1	5	2											
	256			Q9CZU6	A3688	CS	52	8.72	8	3	1											
	257			Q9D0M3	A3698	CYC1	27	8.5	16	12	2											
	258			P16381	A3652	D1PAS1	73	6.74	28	2	1											
	259			P53395	A3675	DBT	46	6.17	5	1	1											
	260			O54734	A3619	DDOST	46	5.34	6	3	1											
	261			Q91VR5	A3656	DDX1	83	6.8	30	5	1											
	262			Q62167	A3653	DDX3	73	6.73	29	4	1											
	263			Q61656	A3814	DDX5	69	9.06	26	2	1											

Appendix 6

Protein Class	PSD #	MASC #	AMPA #	Swall Acc	PPID Acc	Abbreviated Name	Mr [kDa]	pI	1D SDS-PAGE LC-MS/MS analysis of PSD: PSD highest occurance in gel slice #	PSD Peptides Occurrence by MS	PSD # of gel slices protein was present in	1D SDS-PAGE LC-MS/MS analysis of AMPAR IP: AMPAR highest occurance in gel slice #	AMPAR # of gel slices protein was present in	AMPA Peptides Occurrence by MS	Whole extract WB signal Extract	1D SDS-PAGE LC-MS/MS analysis of NRC IP: NRC IA /WB	NRC IP WB	NRC - Peptides Occurrence by MS	NRC # IgG approach	1D SDS-PAGE LC-MS/MS analysis of MAGUK-pep: MAGUK-pep WB result	MAGUK-pep MS-results	MAGUK-pep # peptide approach
Other enzymes	264			Q64591	A3699	DECR1	32	8.94	16	3	1											
	265			Q99J47	A3651	DKFZP566O084	35	9.66	15	1	1											
	266			Q9BQF9	A3641	DKFZP781F069	79	5.7	30	1	1											
	267	106		P08461	A0605	DLAT	68	8.81	1	15	2							1	70			
	268			Q08749	A3674	DLD	50	6.43	3	7	1											
	269			Q01205	A3682	DLST	41	5.82	5	6	1											
	270			Q9WU83	A3638	DPM1	30	9.64	19	2	2											
	271	110		O08553	A0442	DPYSL2	62	5.95	1	46	3										7	102
	272			O35098	A3606	DPYSL4	62	6.51	1	1	1											
	273			Q9JMG8	A3610	DPYSL5	61	6.5	1	1	1											
	274			O08600	A3783	ENDOG	28	8.84	19	1	1											
	275			P17182	A3615	ENO1	47	6.36	8	1	1											
	276			Q9DCT9	A3779	ETFDH	68	7.34	1	6	1											
	277			Q63835	A3659	FACL6	78	8.69	28	4	1											
	278			Q9EQR9	A3584	FASN	272	6.13	39	3	1											
	279			Q8ESW4	A3620	FLJ10842	47	8.65	7	1	1											
	280			Q8TE79	A3634	FLJ23841	29	8.7	18	1	1											
	281			P26037	A3594	FTHFD	99	5.79	29	1	1											
	282	111		Q9Z148	A0444	G9A	138	5.57	27	2	1										1	103
	283			P28492	A3670	GA	55	6.13	1	1	1											
	284	112		P16858	A0331	GAPDH	36	8.45	12	33	3										3	104
	285			Q9Z2W5	A3592	GCS1	92	9.06	30	5	1											
	286	113		P15105	A0424	GLNS	42	6.47	8	9	2										9	105
	287			P26443	A3680	GLUD	56	6.71	5	20	3											
		108		P05202	A0423	GOT2	45	8.97													1	101
	288			Q64521	A3667	GPD2	76	5.82	27	9	2											
	289			Q64428	A3683	HADHA	79	8.95	29	18	2											
	290			Q60587	A3684	HADHB	48	9.35	6	13	2											
	291			P17710	A3663	HK1	108	6.44	33	67	7											
	292			O70503	A3633	HSD17B12	35	9.55	17	1	1											
	293			Q9EQK1	A3660	IDH2	51	8.88	37	2	2											
	294			Q9D6R2	A3629	IDH3A	40	6.27	10	6	1											
	295			Q91VA7	A3658	IDH3B	42	8.76	9	5	1											
	296			P70404	A3669	IDH3G	39	8.66	10	5	2											
	297			Q8R2M5	A3780	ILE5	Frag	Frag	32	2	1											
	298			Q62784	A3587	INPP4A	106	6.64	32	1	1											
	299			O75043	A3640	KIAA0455	Frag	Frag	30	4	1											
	300			O60275	A2429	KIAA0522	Frag	Frag	36	2	1											
		114		P06151	A0370	LDH1	36	7.76													7	106
		115		P16125	A0369	LDHB	36	5.7													3	107
	301			P21396	A3672	MAOA	60	8.12	1	1	1											
	302			Q99MR8	A3604	MCCC1	74	6.68	28	2	1											
	303			P08249	A3696	MDH2	33	8.37	12	4	1											
	304			P48318	A0635	MGAD	67	8.74	1	2	1											
	305			P00405	A3778	MTCO2	26	4.6	20	1	1											
	306			P46935	A1245	NEDD4	103	5.12	34	1	1											
	307			Q8ECJ4	A3635	NRDR	28	9.08	19	1	1											
	308			Q60597	A3664	OGDH	116	6.36	32	6	1											
	309			Q8JLD0	A3665	OGDHL1	114	6.23	32	9	1											
	310			Q05920	A3662	PC	127	6.05	34	12	1											
	311			Q91ZA3	A3601	PCCA	80	7.02	28	2	1											
	312			Q920X1	A3781	PDCD8	56	7.28	26	1	1											
	313			P35486	A3695	PDHA1	40	6.78	8	10	2											
	314			P49432	A3667	PDHB	36	5.29	12	21	2											
	315			Q9WUA3	A1926	PFK-C	85	6.73	30	18	2											
	316	116		P12382	A1925	PFK-L	85	6.62	29	4	1							7	71			
	317			P47858	A0261	PFK-M	85	8.23	30	7	2											
	318	117		P42337	A0153	PI3-K	51	7.65	5	3	1				+++	+	+		72	+		108
	319			O08952	A3656	PI4-K	231	6.57	38	20	2											
	320			Q8CXY9	A3622	PIGK	45	5.32	9	1	1											
	321			O70161	A3591	PIP5K1C	72	5.41	31	1	1											
		118		P53657	A0427	PKLR	62	6.63													1	109
	322			Q99JY8	A3632	PPAP2B	35	9.25	15	3	1											
	323			Q80730	A3631	PRPS1	35	6.51	14	1	1											
	324			Q922W5	A3657	PYCR1	32	6.36	15	1	1											
	325			P53534	A3588	PYGB	97	6.4	31	1	1											
	326			Q9EPK2	A3625	RP2	39	5.11	10	2	1											
	327			P07153	A3639	RPN1	66	6.12	26	7	1											
	328			P25235	A3611	RPN2	67	5.73	1	7	1											
	329			P97700	A3731	SLC25A11	34	9.89	18	14	2											
	330			Q9Z2I9	A3621	SUCLA2	44	5.33	8	6	1											
	331			Q9WUM5	A3697	SUCLG1	32	9.23	12	5	2											
		119		P35700	A0371	TDPX2	22	8.26													3	110
	332	120		P17751	A0433	TPI1	27	7.09	20	3	1										3	111
		121		Q61171	A0372	TPX	22	8.2													2	112
	333			Q9CZ13	A3676	UQCRC1	49	5.28	6	13	1											
	334			P52551	A3677	UQCRC2	47	8.83	8	9	2											
Transcription and Translation																						
Transcription Elements	335			Q9JK31	A3789	ATF7IP	138	4.66	30	1	1											
	336			Q924D3	A3580	CYFIP2	146	6.64	34	11	2											
	337	122		Q61336	A0439	BCAP37	33	9.83	15	12	2										3	113
	338			Q9JI19	A3798	FBP	41	6.27	10	2	1											
		123		P58501	A0296	GCFC	105	5.52										2	73			
	339			Q9D0E1	A3792	HNRPM	78	8.8	27	1	1											
	340			Q99L39	A3794	NRBF1	40	9.17	11	1	1											
	341			Q9Z1N2	A3790	ORC1L	95	9.23	30	2	1											
	342			P24142	A3793	PHB	30	5.57	17	18	2											
	343			P42669	A3795	PURA	35	6.07	9	3	1											
	344			O35295	A3796	PURB	34	5.35	10	3	1											
	345			Q9R0S1	A3788	SND1	99	6.64	32	14	1											
	346			Q8K055	A3799	TRIM32	72	6.49	29	1	1											

Appendix 6

Protein Class	PSD #	MASC #	AMPA #	Swall Acc	PPID Acc	Abbreviated Name	Mr [kDa]	pI	1D SDS-PAGE LC-MS/MS analysis of PSD: PSD highest occurance in gel slice #	PSD Peptides Occurrence by MS	PSD # of gel slices protein was present in	1D SDS-PAGE LC-MS/MS analysis of AMPAR IP: AMPAR highest occurance in gel slice #	AMPAR # of gel slices protein was present in	AMPA Peptides Occurrence by MS	1D SDS-PAGE LC-MS/MS analysis of NRC IP: Whole extract WB signal Extract	NRC IA WB	NRC IP WB	NRC - Peptides Occurrence by MS	NRC # IgG approach	1D SDS-PAGE LC-MS/MS analysis of MAGUK-pep: MAGUK-pep WB result	MAGUK-pep MS-results	MAGUK-pep # peptide approach
Elongation Factors	347			P27706	A3801	EEF1A2	50	9.11	13	1	1											
	348			O70251	A3802	EEF1B	25	4.47	15	4	2											
	349			P58252	A3804	EEF2	95	6.42	31	1	1											
	350			P10126	A1820	EF1A	50	9.1	6	10	1											
	351			Q9D8N0	A1968	EF1G	50	6.31	6	6	1											
	352			O70194	A3805	EIF3S7	64	5.78	26	1	1											
	353			Q922K6	A3806	EIF4B	41	6.17	29	1	1											
	354			P29341	A2508	PABPC1	71	9.48	27	1	1											
Ribosomal Proteins	355			Q63742	A3286	RIBP34	35	9.57	14	5	1											
	356			P53026	A3813	RPL10A	25	9.98	19	4	1											
	357	124		P47963	A0367	RPL13	24	11.54	19	9	2									1	114	
		125		P19253	A0368	RPL13A	23	11.03												1	115	
	358			Q63507	A3814	RPL14	23	11.11	20	1	1											
	359			P50878	A3808	RPL4	47	10.92	6	4	1											
	360			P47962	A1222	RPL5	34	9.78	13	1	1											
	361			P47911	A3810	RPL6	33	10.77	13	6	2											
	362	126		P14148	A0366	RPL7	31	10.89	17	7	1									2	116	
	363			P12970	A3811	RPL7A	30	10.56	17	6	1											
	364			Q6ZQ37	A3812	RPL8	28	11.04	16	1	1											
	365			P14869	A3807	RPLP0	34	5.91	12	10	2											
	366			P25444	A3815	RPS2	31	10.25	16	4	1											
	367			P17073	A3816	RPS3	27	9.68	16	11	1											
	368			P97351	A3817	RPS3A	30	9.75	15	8	2											
	369			P47961	A3818	RPS4X	29	10.16	18	6	1											
	370			P09058	A3823	RPS8	24	10.32	19	6	1											
Cytoskeletal and Cell Adhesion																						
Actin / ARP	371	127		Q35247	A0091	ACTIN	42	5.37	17	179	12				+++	+++	+++		74	+++	142	117
	372			P53488	A0235	ACTR2	45	6.29	9	2	1											
	373			Q99JY9	A0234	ACTR3	47	5.61	6	3	1											
	374			Q9R0Q6	A3287	ARPC1A	42	8.46	11	4	1											
		128		O15144	A0233	ARPC2	34	6.84												5	118	
		129		O15145	A0232	ARPC3	21	8.78												4	119	
		130		O15509	A0231	ARPC4	20	8.53												2	120	
Tubulin	376	131	6	P02551	A0017	TUBULIN	50	4.94	4	317	45	7	6	26	+++	+++	+++	6	75	+++	49	121
Actinin																						
	376	132		Q9JI91	A0007	ACTN	104	5.36	31	13	2				+++	+++	+++	25	76	+++	1	122
	377			Q921P2	A2341	ACTN1	103	5.23	32	9	3											
		133		Q88990	A2343	ACTN3	103	5.31												2	123	
	378	134		P57780	A2344	ACTN4	105	5.26	32	7	1									73	124	
Spectrin	379	135		Q62261	A0010	SPECTRIN	274	5.66	39	875	55				+++	++	++		77	++	2	125
MAPs	380			Q9QXZ0	A2342	ACF7	608	5.36	42	30	2											
	381			P34926	A0085	MAP1A	33	5.46	42	16	2											
	382			P14873	A0625	MAP1B	Precursor	Precursor	42	38	3											
	383			P10638	A0135	MAPT	76	6.34	6	2	1											
	384	136		P20357	A0134	MTAP2	199	4.8	41	32	5				+++	++	+		78	+	1	126
	385			O55129	A3376	MTAP6	96	9.51	34	17	3											
Catenins	386			Q61301	A1914	CATNA	105	5.66	32	37	5				+++	-	-			-		
	387	137		Q02248	A0112	CATNB	85	5.53	31	17	3				+++	++	++		79	+		127
	388			P30999	A0218	CTNND1	102	6.49	32	2	1											
	389			O35927	A0469	CTNND2	135	7.57	35	1	1				+++	-	-			-		
Myelin	390			P35802	A3425	GPM6A	31	5.17	14	1	1											
		138		P04370	A0350	MBP	27	9.58												6	128	
	391	139		Q61885	A0351	MOG	25	8.38	20	3	1									4	129	
	392			Q63912	A3590	OMG	47	8.79	31	15	5											
	393	140	7	Q62079	A0352	PLP1	27	8.03	38	13	9	1	1	3						18	130	
Cell recognition molecules	394			P97685	A2663	NFASC	Frag	Frag	38	22	3											
	395			P97547	A3607	SOFR1	41	8.61	1	5	3											
	396			Q05546	A1074	TENASCIN-R	149	4.82	42	52	7											
Keratins	397			Q9CWC9	A3741	2410039E07RIK	26	4.83	37	3	1											
	398			Q920F2	A1669	KRT5	62	7.59	37	3	1											
	399			Q62332	A3846	KRT6A	59	8.04	20	2	1											
Other Cytoskeletal Proteins	400	141		Q9EPW6	A3976	ABLIM1	88	8.91	28	2	1									1	131	
	401			P42024	A0660	ACTR1A	43	6.19	7	6	1											
	402			Q9R1V4	A3880	ADAM11	59	6.23	27	1	1											
	403	142		Q9R1V6	A0373	ADAM22	70	7.76	30	12	3									2	132	
	404			Q9R1V7	A3881	ADAM23	92	7.95	29	9	1											
	405			Q9QYC0	A3876	ADD1	81	5.62	33	16	3											
	406			Q9QYB6	A0567	ADD2	81	5.87	33	12	1											
	407			Q9QYB5	A3877	ADD3	79	5.65	31	7	2											
	408			Q01484	A3873	ANK2	430	5.03	39	95	10											
	409			Q61307	A3874	ANK3	212	6.18	37	23	9											
	410	143		Q68737	A0285	BASSOON	419	7.62	42	60	4							11	80			
				P55284	A0320	CADHERIN 5	83	5.03							+++	-	-			-		
	411	144		P47754	A0417	CAPZ ALPHA-2	33	5.58	12	16	3									2	133	
	412	145		P47757	A0418	CAPZ BETA	31	5.47	16	13	2									1	134	
	413			O15083	A3879	CAST	111	6.62	33	16	3											
	414			Q8R3F2	A3211	CKAP4	66	5.63	1	8	1											
	415			Q8C5W0	A3870	CLMN	117	4.93	36	1	1											
		146		Q60598	A0083	CORTACTIN	61	5.24							+++	+	+		81	++		135
	416	147		Q9QX74	A0081	CORTBP1	159	6.76	37	4	1				+++	+	++		82	+++		136
	417			Q8R555	A3975	CRTAC1	70	4.95	29	5	2											
		148		Q9JP98	A0446	CSE1L	110	5.59							+++	++	++		83	+	2	137
	418	149		Q9QXS6	A0438	DBN1	77	4.45	35	13	3										1	138
		150		Q61495	A0291	DESMOGLEIN	Frag	Frag							+++	++	++		84	+		139
	419			Q99KH4	A1111	E391	52	6.57	1	2	1											
				P09503	A0318	E-CADHERIN	80	4.29							+++	-	-			-		
	421			Q99MZ2	A3878	ELKS	112	6.11	33	9	2											
	423			Q9JMB3	A3865	EPM1L3	103	5.2	35	16	3											
	422			Q9Z2H5	A3867	EPM4L1	98	5.46	32	3	1											
				P26040	A0326	EZRIN	69	5.83							+++	-	-			-		

Appendix 6

Protein Class	PSD #	MASC #	AMPA #	Swall Acc	PPID Acc	Abbreviated Name	Mr [kDa]	pI	1D SDS-PAGE LC-MS/MS analysis of PSD			1D SDS-PAGE LC-MS/MS analysis of AMPAR IP			1D SDS-PAGE LC-MS/MS analysis of NRC IP					1D SDS-PAGE LC-MS/MS analysis of MAGUK-pep		
									PSD highest occurance in gel slice #	PSD Peptides Occurrence by MS	PSD # of gel slices protein was present in	AMPAR highest occurance in gel slice #	AMPAR # of gel slices protein was present in	AMPA Peptides Occurrence by MS	Whole extract WB signal Extract	NRC IA WB	NRC IP WB	NRC - Peptides Occurrence by MS	NRC # IgG approach	MAGUK-pep WB result	MAGUK-pep MS-results	MAGUK-pep # peptide approach
Other Cytoskeletal Proteins		151		Q5JJ36	A0204	FILAMIN	281	5.73							+++	+	-			++		140
	424			Q96I05	A3979	FLJ40935	52	6.48	13	1	1											
	426	162		P13020	A0419	GELSOLIN	83	5.72	31	1	1										3	141
				P11688	A0474	INTEGRIN ALPHA 5	110	5.41							+++		-			-		
	428	163		P46660	A0292	INTERNEXIN	56	5.16	2	31	1							2	85			
	427			Q14008	A4009	KIAA0097	226	8.03	38	1	1											
	426			Q9ERD4	A3875	KIDINS220	191	6.43	35	2	1											
	429			Q61792	A3662	LASP1	30	6.61	11	5	2											
		164		P14733	A0430	LMNB1	67	5.11													1	142
	430			Q9D052	A3991	LYRIC	21	9.87	29	1	1											
	420			Q15555	A0654	MAPRE2	37	5.36	15	2	1											
	432	166		P12839	A0376	NEF3	96	4.76	35	15	2										2	144
	431	165		P08551	A0375	NF-L	61	4.63	27	21	3										6	143
	433			O35449	A3989	NQ5	31	7.26	10	2	1											
	434			Q62034	A3863	PARALEMMIN	42	4.79	1	8	3											
				Q99LE7	A0248	PAXILLIN	65	5.73							+++	-	-			-		
				P10287	A0319	P-CADHERIN	80	4.5							+++	-	-			-		
	435			P30427	A1941	PLECTIN	534	5.71	41	37	2											
	436			O75145	A3869	PPFIA3	Frag	Frag	35	10	1											
		167		Q28523	A0216	SLMAP	74	5.05										3	86			
	437			O35274	A0273	SPINOPHILIN	90	4.66	34	3	1											
	438			Q6Z327	A3871	SYNPO	74	9.31	32	6	1											
				Q9D8T6	A0325	TENSIN	65	9.56							+++	-	-			-		
	439			P01831	A1218	THY1	13	9.16	19	31	4											
				Q64727	A0098	VINCULIN	117	5.72							+++	-	-			-		
	440			Q9ERQ9	A0148	WASF1	62	5.94	29	3	1											
	441			Q9NSL0	A3872	WBP3	Frag	Frag	36	1	1											
	442			Q9ESJ4	A1027	WISH	79	6.62	30	2	1											
Other Cell Adhesion Molecules	443			Q06890	A1276	APOJ	49	5.46	15	1	1											
	444			Q61137	A3896	ASTN	Frag	Frag	34	5	1											
	445			Q9JHW6	A3897	ASTN2	100	5.17	33	1	1											
	446			O54991	A0758	CASPR	154	6.64	37	33	4											
	447			P23242	A0493	CONNEXIN-43	43	8.97	10	6	1											
	448			P12960	A0757	CONTACTIN-1	109	5.76	34	109	12											
	449			Q61330	A1073	CONTACTIN-2	Frag	Frag	35	7	2											
	450			Q9QUP5	A3893	CRTL1	39	7.64	7	3	1											
	451			P55066	A3892	CSPG3	135	5.45	42	8	2											
	452			Q9QY32	A3895	CSPG5	57	4.32	37	3	2											
	453			P70587	A0150	DENSIN-180	167	6.13	37	9	2											
			8	P15924	A0454	DESMOPLAKIN	332	6.44				3	2	9								
	454			Q62718	A3887	HNT	32	5.81	2	3	2											
	455			Q9Z0J8	A3888	KILON	35	6.13	5	3	1											
	456	168		P11627	A0290	L1CAM	139	5.62	36	8	3				+++	+++	+++		87	++		145
	457	169		Q9JIA1	A0440	LGI1	60	8.16	2	41	2										8	146
	458			P20917	A3898	MAG	67	4.99	32	4	2											
	459	170		P15116	A0219	N-CADHERIN	82	4.42	34	6	2				+++	++	++		88	+++		147
	460			P13595	A3884	NCAM1	117	4.72	34	31	7											
	461			O35136	A3885	NCAM2	91	5.71	34	2	1											
	462			P97798	A3894	NEO1	159	5.95	38	2	1											
	463			Q62765	A0129	NLGN	89	5.96	31	1	1											
	464			Q62889	A3889	NLGN2	90	5.78	32	4	2											
	465			Q62889	A3890	NLGN3	90	5.71	33	5	2											
	466			Q9Z0E0	A1163	NORBIN	79	5.38	28	2	1											
	467			Q07314	A3891	NRXN3	171	6.12	37	7	2											
	468			P32736	A3886	OPCML	32	5.35	3	3	2											
	469	161		Q62251	A0141	VEL1	20	9.32	19	5	1										10	148
	470			Q9Z252	A3899	VEL2	23	8.71	18	1	1											
Synaptic Vesicles / Protein Transport																						
Clathrin	471			P08081	A0122	CLTA	27	4.41	11	6	3											
	472			P08082	A3901	CLTB	25	4.56	12	3	2											
	473	182		P11442	A0119	CLTC	192	5.5	36	163	11				+++	++	++	1	89	+++	2	149
Synaptic vesicle	474			O08838	A0121	AMPH	75	4.57	34	4	1											
	475			O08839	A1109	AMPHIPHYSIN2	65	4.95	30	1	1											
	476			Q61548	A0584	AP180	92	4.72	37	14	2											
	477			P52303	A0686	AP1B1	105	5	29	19	4											
	478			P17426	A0251	AP2A1	108	6.63	33	60	4											
	479			P17427	A0252	AP2A2	104	6.51	31	38	2											
	480			Q9DBG3	A0253	AP2B1	105	5.22	32	64	5											
	481			P20172	A0254	AP2M1	50	9.57	5	13	1											
	482			O54774	A1324	AP3D	135	7.04	36	1	1											
	483			P53678	A3907	AP3M2	47	7.15	7	2	1											
	484			P35564	A0452	CALNEXIN	65	4.5	31	4	1											
	485			P53621	A2102	COPA	138	7.7	36	1	1											
	486			Q55029	A0237	COPB2	102	5.17	31	1	1											
	487			Q9EP81	A2346	CYLN2	116	6.07	34	5	1											
	488	183		Q01469	A0441	E-FABP	15	6.6													1	150
	489			Q9QXY6	A1331	EHD3	61	6.03	2	1	1											
	490			Q9Z1Z3	A3926	EPN2	62	7.67	28	1	1											
	491			O54922	A3921	EXO70	75	6.27	28	1	1											
	492			O54924	A3923	EXO84	81	5.3	29	2	1											
	493			O55126	A3913	GBAS	33	9.31	20	6	1											
	494			O88630	A3935	GOSR1	26	9.34	19	2	1											
	495			Q9JKY5	A3906	HIP1R	119	6.14	33	2	1											
	496			Q61595	A3531	KTN1	153	5.67	35	4	1											
	497			O88994	A4008	MGB7	38	8.95	12	2	1											
	498			Q9Y3A3	A3935	MOB3	26	5.5	19	2	1											
	499			Q62769	A0560	MUNC-13-1	196	5.1	38	2	2											
	500			P54921	A0596	NAPA	33	5.3	14	4	1											
	501			P28663	A3910	NAPB	34	5.32	14	4	1											
	502			Q9CWZ7	A3911	NAPG	35	5.31	11	7	2											

Appendix 6

Protein Class	PSD #	MASC #	AMPA #	Swall Acc	PPID Acc	Abbreviated Name	Mr [kDa]	pI	1D SDS-PAGE LC-MS/MS analysis of PSD			1D SDS-PAGE LC-MS/MS analysis of AMPAR IP			1D SDS-PAGE LC-MS/MS analysis of NRC IP				1D SDS-PAGE LC-MS/MS analysis of MAGUK-pep		
									PSD highest occurance in gel slice #	PSD Peptides Occurrence by MS	PSD # of gel slices protein was present in	AMPAR highest occurance in gel slice #	AMPAR # of gel slices protein was present in	AMPA Peptides Occurrance by MS	Whole extract WB signal Extract	NRC IA WB	NRC IP WB	NRC - Peptides Occurrence by MS	NRC # IgG approach	MAGUK-pep WB result	MAGUK-pep MS-results	MAGUK-pep # peptide approach
Synaptic vesicle	503			O55125	A3914	NIPSNAP1	33	9.48	19	13	2											
	504	164		P46460	A0117	NSF	83	6.52	28	75	9				+++	++	++	1	90	+++		151
	505			O55011	A3927	PICALM	65	9.02	28	1	1											
	506			Q9JM89	A3932	PXR2	69	5.16	26	1	1											
	507			Q62991	A3928	RAB10	72	6.09	26	1	1											
	508			O35168	A0035	RIM-1	180	9.62	36	1	1											
	509			P47708	A0212	RPH3A	75	8.56	30	8	2											
	510			P97878	A3925	SEC10L1	82	6.27	27	4	1											
	511			Q9D4R7	A3924	SEC15L2	17	8.48	31	1	1											
	512			Q9D4H1	A3920	SEC5L1	104	6.78	31	1	1											
	513			Q62825	A3922	SEC8L1	86	5.79	29	7	1											
	514			Q99JB2	A3942	SLP2	38	8.95	10	1	1											
	515	165		P70586	A0333	SNAP25	23	4.66	20	44	2										4	152
	516			Q9QXY3	A3903	SNIP	127	9.39	35	57	10											
	517	166		Q64548	A0445	S-REX	83	4.53	35	1	1										2	153
	518	167		Q9QXG3	A0040	STX	33	5.2	14	67	6				+++	++	++		91	++	4	154
	519			Q9ER00	A3939	STX12	31	5.33	13	3	1											
	520			O70439	A3940	STX7	30	5.6	10	2	2											
	521	168	9	O08599	A0238	STXBP1	68	6.62	26	45	8	9	10	173							8	155
	522			O02563	A3902	SV2	83	5.38	32	1	1											
	523			Q15431	A3863	SYCP1	114	5.85	35	1	1											
	524			O88935	A0161	SYN1	70	9.84	29	29	4											
	526			Q9QWV7	A3904	SYN2	52	7.61	4	22	3											
	526	169		O55100	A0288	SYNAPTOGYRIN	26	4.49	17	3	2				+	+	+		92	+	1	156
	527			Q9WVG8	A3933	SYNGR3	26	8.42	20	3	1											
	528			Q62910	A0120	SYNJ1	173	6.43	36	6	2											
	529			Q62277	A3941	SYP	33	4.91	10	3	1											
	530	170		P46096	A0039	SYT1	47	8.68	2	77	9										2	157
	531			Q9R0N7	A3905	SYT7	45	9.33	28	3	2											
	532			Q9DCD5	A4027	TJP4	59	5.74	11	1	1											
	533			Q9Z1S2	A0562	TOM35YN	124	6.62	35	2	2											
	534			Q9Z270	A3937	VAPA	27	8.58	17	15	2											
	535			Q9QY76	A3938	VAPB	27	7.64	18	4	1											
	536			Q01853	A3918	VCP	89	5.14	31	2	1											
	537			Q920Q4	A3908	VPS16	95	6.48	31	1	1											
	538			Q63615	A3930	VPS33A	68	6.6	26	2	1											
	539			Q61123	A3909	VPS35	92	5.28	29	4	1											
	540			Q9H3I8	A1326	VPS41	90	6.18	31	1	1											
	541			P97390	A3929	VPS45A	65	8.42	1	9	1											
	542			O88384	A3934	VTI1L1	27	8.73	19	10	2											
Transporters	543			Q9CXJ4	A3947	ABCB8	78	9.25	1	5	1											
	544			P55096	A3946	ABCD3	75	9.36	26	4	1											
	545	171		Q9H936	A3773	GC1	34	9.42	15	2	1										2	158
	546			Q9QZP4	A3948	MTCH1	40	9.54	9	5	2											
	547			Q99LZ6	A3949	MTCH2	Frag	Frag	15	15	3											
	548			P47802	A3956	MTX1	36	5.82	14	4	1											
	549			O88441	A3957	MTX2	30	5.44	20	6	1											
	550			Q01405	A3945	SEC23A	86	6.64	29	1	1											
	551			Q63965	A3951	SFXN1	36	9.21	13	1	1											
	552			Q9JHY2	A3952	SFXN3	35	9.79	14	5	1											
	553			Q925N0	A3954	SFXN5	37	9.49	11	2	1											
	554			P32089	A3750	SLC25A1	33	9.77	17	5	2											
	555	172		O75746	A0443	SLC25A12	75	8.58	26	7	1										1	159
	556			Q9QXX4	A3747	SLC25A13	74	8.77	27	1	1											
	557			P16036	A3776	SLC25A3	35	9.18	17	2	1											
	558	173		P48962	A0395	SLC25A4	33	9.73	18	24	2										16	160
	559	174		Q09073	A0396	SLC25A5	33	9.74	18	26	2										9	161
	560			P97441	A3943	SLC30A3	42	5.91	10	2	1											
	561			Q8CZW5	A3950	TOMM70A	68	7.08	27	6	1											
	562			P40336	A3944	VPS26	38	6.13	11	1	1											
Motor Proteins	563			Q8R1R2	A1178	BERP	81	8.01	29	3	1											
	564			Q99J03	A3970	CLASP2	68	6.24	35	4	1											
	565			Q9QUR2	A3969	DCTN4	53	7.84	2	2	1											
	566			P38650	A0133	DNCH1	532	6.06	41	109	8											
	567			Q99KJ8	A0656	DYNAMITIN	44	5.14	5	5	1											
	568			O88487	A2127	DNCI2	68	5.16	29	3	1											
	569			O88485	A0661	DYNEINIC	71	5.02	28	6	2											
	570			Q9HBB2	A3962	FLJ13881	63	5.16	38	1	1											
	571			Q9D9F6	A3973	MMT	26	6	30	29	4											
	572			Q9Z0R4	A0571	INTERSECTIN-1	194	7.91	38	5	1											
	573			P28740	A3966	KIF2	81	5.92	32	8	2											
	574			Q61768	A3967	KIF5B	110	5.96	33	5	1											
	575	175		O88448	A3968	KLC2	67	6.77	27	1	1							2	93			
		176		Q63791	A4048	MRLC2	20	4.77													1	162
	576	177		Q9JLT0	A3961	MYH10	229	5.49	39	46	4				+++	+++	+++	27	94	++	194	163
	577	178		O08638	A2331	MYH11	227	5.37	39	19	2				+++	+++	+++	27	95	++	9	164
		179		O61355	A3202	MYH7	223	5.63													1	165
	578	180		Q8VDD5	A3166	MYH9	226	5.54	39	15	3				+++	+++	+++	27	96	++	20	166
		181		Q60605	A4049	MYL6	17	4.56													1	167
	579	182		P46735	A3958	MYO1B	129	9.29	34	4	1				+++	+++	+++	27	97	++		168
	580			Q63357	A3959	MYO1D	116	9.45	32	1	1											
	581			Q64331	A3960	MYO6	146	9.01	36	6	1											
	582	183		Q9QYF3	A0149	MYOSIN (V)	212	8.87	38	85	5				+++	+++	+++	27	98	++	3	169
	583			O60313	A3971	OPA1	108	6.86	31	35	5											
	584			O08788	A0658	P150GLUED	142	5.72	35	5	2											
	585			P30622	A2345	RESTIN	161	5.31	29	1	1											
	586			Q9JMH9	A3964	TIAF1	231	5.84	39	9	1											
	587			P70566	A3972	TMOD2	39	5.34	9	9	1											
	588			Q63608	A3965	TPM1	33	4.71	18	1	1											
	589			P21107	A3205	TPM3	33	4.68	15	1	1											

Appendix 6

Protein Class	PSD #	MASC #	AMPA #	Swall Acc	PPID Acc	Abbreviated Name	Mr [kDa]	pI	1D SDS-PAGE LC-MS/MS analysis of PSD			1D SDS-PAGE LC-MS/MS analysis of AMPAR IP			Whole extract WB signal Extract	1D SDS-PAGE LC-MS/MS analysis of NRC IP				1D SDS-PAGE LC-MS/MS analysis of MAGUK-pep		
									PSD highest occurance in gel slice #	PSD Peptides Occurrence by MS	PSD # of gel slices protein was present in	AMPAR highest occurance in gel slice #	AMPAR # of gel slices protein was present in	AMPA Peptides Occurrence by MS		NRC IA WB	NRC IP WB	NRC - Peptides Occurrence by MS	NRC # IgG approach	MAGUK-pep WB result	MAGUK-pep MS-results	MAGUK-pep # peptide approach
Novel	590			Q9D5I5	A3994	0710001P09RIK	30	8.67	15	2	1											
	591			Q8Q2T2	A3995	1700054N08RIK	28	9.3	11	1	1											
	592			Q9CZU2	A3995	2610528A17RIK	29	9.85	17	6	1											
	593			Q8Z2D2	A3510	4833411B01RIK	Frag	Frag	38	1	1											
	594			Q9D217	A4003	A330102H22RIK	30	4.82	32	1	1											
		186		Q9DB63	A2084	APG16L	53	5.66										1	100			
		184		Q9D5J9	A4015	BA342L8.1	48	9.12													1	170
	595			Q8JLH7	A4007	CDK5RAP3	57	4.68	26	2	1											
	596			Q9CU45	A3900	DJ403A15.3	Frag	Frag	29	4	1											
		185		Q8K2R1	A0281	DKFZP564A026	Frag	Frag										3	99			
	597			Q8VD37	A3987	DKFZP761D221	35	6.36	34	1	1											
	610			Q8R1A4	A3986	DOCK7	85	5.73	39	1	1											
	598			Q9CRB9	A3558	FLJ20420	26	8.56	19	8	1											
	599			Q12914	A4002	G2	Frag	Frag	32	1	1											
	600			Q9HCQ3	A2325	HCMOGT1	85	5.73	40	1	1											
	616			Q96TC7	A3511	HUCEP-10	52	5.02	1	3	1											
	601			Q14700	A4000	KIAA0090	112	7.38	33	2	1											
	602			Q9CRD2	A3977	KIAA0103	95	5.88	15	1	1											
	603			Q922I2	A4006	KIAA0143	93	6.36	30	4	1											
	604			Q60302	A3912	KIAA0555	95	5.88	31	1	1											
	605			Q9Y2F0	A4011	KIAA0941	58	9.33	2	2	1											
	606			Q91ZG0	A4035	KIAA0953	Frag	Frag	30	1	1											
	607			Q9ULT3	A3992	KIAA1136	Frag	Frag	42	1	1											
	608			Q9ULQ0	A3983	KIAA1170	Frag	Frag	6	1	1											
	609			Q8HCM3	A3554	KIAA1549	Frag	Frag	31	1	1											
	611			Q9Z2I0	A4004	LETM1	83	6.16	30	2	1											
	612			Q86TV6	A3468	LOC145567	83	6.29	30	1	1											
	613			Q8N3B6	A4023	LOC90550	40	8.83	16	6	1											
	614			Q8K4V0	A3999	LRPPRC	147	5.92	34	2	1											
	615			O35606	A4045	MIC1	75	7.89	27	1	1											
	617			Q9CUN3	A3993	SVH	38	6.16	17	1	1											
	618			Q9JJK2	A3988	TASP	51	6.92	7	6	2											
	619			Q8EQN7	A3997	TTYH1	49	4.96	5	1	1											
	620			Q9Z0I9	A4010	WDR7	163	6.58	36	3	1											

Appendix 6 - Overview of G2C PSP datasets. Proteins found in the PSP (MASC (combined NRC, MAGUK-pep proteins) and PSD) are listed with running numbers for each dataset. UniProt and PPID accession numbers and gene names are included for each protein. In the PSD (yellow fill) and AMPAR (light green fill) datasets, gel slices in which a given protein was detected highest and the numbers of gel slices a given protein was detected in is indicated. The NRC (dark green fill) dataset is composed of previously published western blotting analysis of the antibody-based purification, in which (+)/(-) indicates the semi-quantatitive presence of the protein and of new LC-MS/MS analysis of antibody-based purification for which peptide counts are included. Similarly the MAGUK-pep (gold fill) dataset is also derived from a combination of western blotting data and LC-MS/MS analysis, where the presence/absence of a protein/peptide is indicated in the manner described above. In the western blotting section, "whole" denotes presence of a given protein in whole forebrain extract. The MASC, which is a combination of the NRC and MAGUK-pep datasets, has a running number which includes proteins detected in both datasets by both western blotting and LC-MS/MS.

Appendix 7

Protein Class	UniProt Acc	PPID Acc	Unigene cluster ID	OMIM	Gene name	Abbreviated name	Confidence (# of times identified)	Postsynaptic density datasets								Complexes			Other structures			# TM helices
								G2C #	Yosh #	Jordan #	Peng #	LI #	Satoh #	Walik #	PSD from Literature	MASC #	AMPA #	mGluR5 #	Synaptosome	Clathrin Vesicle	Mitochondria	
Channels and Receptors																						
Glutamate Receptors	P23818	A0296	Mm.4920	138248	Gria1	AMPA-R1	5	1	1	1	1				8260822		1					3
	P23819	A0297	Mm.220224	138247	Gria2	AMPA-R2	6	2	2	2	2	1			8260822		2					4
	Q9Z2W9	A0299	Mm.327681	305915	Gria3	AMPA-R3	5	3	3	3	3				12470934		3					3
	P19493	A0300	Mm.209263	138246	Gria4	AMPA-R4	3		4		4				11416187		4					4
	P39087	A0022	Mm.332838	138244	Grik2	GLUR6	2		5		4					1						3
	Q61625	A0328	Mm.321227	602368	Grid2	GRID2	1			5												3
	Q9EPV6	A0103	Mm.370186	604473	Grm1	MGLUR1	3	4			5				10433288	2						7
	Q9QYS2	A1353	Mm.318966	601115	Grm3	MGLUR3	3	5			6				11891482							6
	P31424	A0327	Mm.235018	604102	Grm5	MGLUR5	2	6							10670165	3						6
	P35400	A0453	Mm.240881	604101	Grm7	MGLUR7	2	7	6													6
	Q62683	A0001	Mm.278672	138249	Grin1	NR1	7	8	7	6	7	2		1	9186306	4						3
	O08948	A0002	Mm.2953	138253	Grin2a	NR2A	7	9	8	7	8	3		2	9186306	5		1				4
	Q62684	A0003	Mm.322010	138252	Grin2b	NR2B	7	10	9	8	9	4		3	9186306	6						4
	Q03391	A0005	Mm.152584	602717	Grin2d	NR2D	1		10													4
Other ligand receptors	O88917	A0038	Mm.260733	--	Lphn1	CIRL-1	4	11	11		10				12196118							8
	Q9Z173	A3463	Mm.273631	--	Lphn3	CIRL-3	2	12			11											7
	P18504	A0503	Mm.6834	602376	Gabra1	GABA-A-R1	3	13	12				1									4
	Q9WV18	A0603	Mm.32191	603540	Gabbr1	GABA-B-R1	5	14	13	9	12				14688901							7
	O88871	A2000	Mm.101909	607340	Gpr51	GABA-B-R2	3	15			13				14688901							7
	O55022	A3464	Mm.9052	300435	Pgrmc1	PGRMG	1	16														1
Ca2+-ATPases	O55143	A3467	Mm.227583	108740	Atp2a2	ATP2A2	1	17														8
	P11505	A3465	Mm.166944	108731	Atp2b1	ATP2B1	1	18										2	1			7
	Q6R0K7	A0508	Mm.321755	108733	Atp2b2	ATP2B2	1	19											2	1		8
	Q64568	A3466	Mm.210095	300014	Atp2b3	ATP2B3	1	20											3			6
	Q64542	A0114	Mm.188617	108732	Atp2b4	ATP2B4	1	21								7			4			7
NA+/K+-ATPases	P06685	A0340	Mm.280103	182310	Atp1a1	ATP1A1	2	22			14					8	6	5	5			10
	P06686	A0341	Mm.207432	182340	Atp1a2	ATP1A2	1			10								4	6	2		8
	Q9Z1G6	A0342	Mm.44101	182350	Atp1a3	ATP1A3	3	23	14		15					9		5	7	3		8
	Q9WV28	A0343	Mm.337950	607321	Atp1a4	ATP1A4	1	24											8			10
	P14094	A0344	Mm.4550	182330	Atp1b1	ATP1B1	1	25										6	9	4		1
Voltage- and Ligand-gated Ca2+-channel	P54282	A0593	Mm.334658	601011	Cacna1a	CACNA1A	2	26	15									7				17
	O55017	A0592	Mm.4424	601012	Cacna1b	CACNA1B	1	27														21
	Q01815	A0770	Mm.236039	114205	Cacna1c	CACNA1C	2	28	16													19
	Q61290	A3471	Mm.298091	601013	Cacna1e	CACNA1E	4	29		11	16	8										19
	O08532	A3474	Mm.173392	114204	Cacna2d1	CACNA2D1	2	30			17											
	Q8EQQ2	A3472	Mm.273084	607082	Cacna2d2	CACNA2D2	1	31														
	Q9Z1L5	A3473	Mm.370172	606399	Cacna2d3	CACNA2D3	1	32														
	Q7TPF2		Mm.41252	114207	Cacnb1	CACNB1	2			12	18											
	P54285	A3476	Mm.3544	601958	Cacnb3	CACNB3	3	33	17	13												
	Q9JJV5	--	Mm.241121	606403	Cacng3	CACNG3	1		18													4
Voltage-gated K+-channel	P52190	A1867	Mm.140760	600504	Kcnj4	KCNJ4	1		19													3
	Q8VHF1	--	Mm.343607	600150	Kcnma1	KCNMA1	2		20		19											6
	Q9Z343	A3477	Mm.40615	602235	Kcnq2	KCNQ2	1	34														5
	P63142	A0764	Mm.56930	176262	Kcna2	KV1.2	2	35							11487820							5
	Q64284	A3478	Mm.302496	601142	Kcnab2	KV2.2	1	36														
Voltage-dependent anion channels	Q60932	A0425	Mm.3555	604492	Vdac1	VDAC1	5	37		14	20	6			10817960	10			10		1	
	Q60930	A0426	Mm.262327	193245	Vdac2	VDAC2	4	38		15	21	7				11			11		2	
	Q60931	A3479	Mm.227704	--	Vdac3	VDAC3	4	39	21	16	22								12		3	
Na+-channel	Q99250	A0598	Hs.93485	601219	SCN2A	SCN2A2	1	40														17
	P15390	--	Mm.218743	603967	Scn4a	SCN4A	2		22			9										20
Ca2+-release channels	P11881	A0154	Mm.227912	147265	Itpr1	INSP3R	2	41							11416187							6
	Q9ERN6	A3480	Mm.239871	180902	Ryr2	RYR2	1	42														6
Other Channels and Receptors	P12023	A1169	Mm.277585	104760	App	APP	2	43							7494461							1
	Q9Z1G4	A0414	Mm.340818	192130	Atp6v0a1	ATP6N1	2	44			23								13			6
	Q9UI12	A0409	Mm.27082	608861	ATP6V1H	ATP6V1H	2	45			24									5		
	O14514	A1017	Mm.43133	602682	BAI1	BAI1	2		23	17												7
	O60242		Mm.336569	602684	BAI3	BAI3	2			18	25											8
	Q08406	A3490	Mm.272210	118946	Cntfr	CNTFR	1	46														
	O08530	A3491	Mm.982	601974	Edg1	EDG-1	1	47														7
	Q80VZ2	--	Mm.3249	602188	Epha4	EPHA4	1		24													1
	P54762	A0542	Mm.22897	600600	EPHB1	EPHB1	1								9883737							1
	Q63645	--	Mm.1614	600933	F2rl1	F2RL1	1		25													7
	Q920E0	--	Rn.81079	--	F2rl3	F2RL3	1		26													7
	O35658	A1608	Mm.30049	601269	C1qbp	GC1Q-R	1	48											14		4	
	Q03555	A0637	Mm.341742	603930	Glphn	GEPHYRIN	4	49		19	26				12967996							
	P60894	--	Mm.167625	605188	GPR85	GPR85	1		27													7
	O88703	A3486	Mm.12956	602781	Hcn2	HAC-1	1	50														3
	Q9NZN1	--	Mm.267669	300206	IL1RAPL1	IL1RAPL1	1		28													1
	Q61291	A1175	Mm.271854	107770	Lrp1	LRP	1	51														1
	P15209	A0506	Mm.130054	600456	Ntrk2	NTRK2	1								9037644							2
	P70206	A3482	Mm.3789	601055	Plxna1	PLEXIN-1	1	52														1
	P70207	A3483	Mm.2251	601054	Plxna2	PLEXIN-2	1	53														1
	O63495	A1094	Mm.3383	600214	Ager	RAGE	1				27											1
	Q91V14	A3489	Mm.252987	606726	Slc12a5	SLC12A5	1	54										9				12
	O62634	A4034	Mm.255831	605208	Slc17a7	SLC17A7	1	55														10
	P43006	A0422	Mm.267547	600300	Slc1a2	SLC1A2	2	56	29							12			15	6		9
	Q9ES70	A3481	Mm.314497	605556	Slc4a10	SLC4A10	1	57														11
	P16283	A3488	Mm.5053	106195	Slc4a3	SLC4A3	1	58														12
	O88343	A3485	Mm.41044	603345	Slc4a4	SLC4A4	1	59										10				9
	P48768	A3484	Mm.241147	601901	Slc8a2	SLC8A2	1	60														9
MAGUKs / Adaptors / Scaffolders																						
PDZ-domain containing scaffolders	Q9QYX6	A0568	Mm.332219	604918	Polo	ACZONIN	3	61	30	20												
	Q70589	A0125	Mm.327591	300172	Cask	CASK	2	62		21												
	Q63622	A0014	Mm.257035	603583	Dlg2	CHAP/GYN-110	6	63	31	22		9	2		8754483	13			18			
	Q9WV34	A0436	Mm.36242	600723	Mpp2	DLGH2	2	64		23						14						
	O88910	A0435	Mm.20449	601114	Mpp3	DLGH3	2	65		24						15						
	Q925T6	A0301	Mm.196692	604597	Grip1	GRIP	2		32						9089286			11				
	Q9C0E4	A0302	Mm.333764		GRIP2	GRIP2	1								9788844							
	Q92216	A0104	Mm.37533	604796	Homer1	HOMER1	8	66	33	25	29	10	3	4	9069287	16		12				
	Q99JP6	--	Mm.10022	604800	Homer3	HOMER3	1			26												

Appendix 7

Protein Class	UniProt Acc	PPID Acc	Unigene cluster ID	OMIM	Gene name	Abbreviated name	Confidence (# of times identified)	Postsynaptic density datasets								Complexes			Other structures			# TM helices
								G2C #	Yosh #	Jordan #	Peng #	Li #	Satoh #	Walikonis #	PSD from Literature	MASC #	AMPA #	mGluR5 #	Synaptosome	Clathrin Vesicle	Mitochondria	
	Q9R093	A0137	Mm.197074	--	Cnksr2	MAGUIN	5	87	34	27	29	11										
	Q9WV37	A0793	Mm.41288	606959	Mpp6	MPP6	1	69														
	Q62083	A0449	Mm.259464	605926	Prkcabp	PICK1	1								10027300							
	Q9JLU4	A0074	Mm.146855	606230	Shank3	PROSAP2	4	69	35	28	30											
	Q62108	A0013	Mm.27256	602887	Dlg4	PSD-95	6	70	38	29	31	12	4	5	1419001	17			17			
	P70175	A0016	Mm.4615	300189	Dlg3	SAP102	6	71	37	30	32	13			8780648	18						
	Q62402	A0015	Mm.382	601014	Dlg1	SAP97	5	72	38	31		14			8427241	19						
	Q9WU13	A0075	Mm.358922	604999	Shank1	SHANK1	6	73	39	32	33	15			10433288	20		13				
	Q9H0F6	A0602	Mm.41463		DKFZp434N192	SHARPIN	1								11178675							
	P39447	A0086	Mm.4342	601009	Tjp1	ZO-1	3	74	40		34					21						
	P70625	A1881	Mm.104744	607709	Tjp2	ZO-2	1				35								15			
non-PDZ-domain containg scaffolders	P25304	A3494	Mm.273098	103320	Agrn	AGRIN	1	75														1
	P24587	A0143	Mm.311452	604688	Akap6	AKAP150	3	76			36				1512224	22						
	O54931	A0796	Mm.348266	604582	Akap2	AKAP2	1		41													
	Q8K3H9	A0188	Mm.202322	604299	Appl1	APPL	0									23						
	O35413	A0109	Mm.211096	--	Arpp2	ARGBP2	3		42	33	37											
	Q60437	A1015	Mm.197534	605475	Baiap2	BAIAP2	6	77	43	34	38	16	5	6	10460238							
	O88881	A0228	Mm.342085	--	BEGAIN	BEGAIN	3		44	35					9766850							
	Q8VHK2	A3690	Mm.70989	--	Caskin1	CASKIN1	3	78		36	39											
	O70333	A0275	Mm.109329	604594	1200020A08Rik	CRIPT	1								8681782							
	Q9Y2H0	--	Mm.22094	--	DLGAP4	DAP4	1		45													
	P97837	A0447	Mm.323861	605438	Dlgap2	DLGAP2	2		46		40								18			
	Q60902	A3493	Mm.268894	--	Eps15l1	EPS15R	1	79														
	O08917	A3496	Mm.2931	606998	Flot1	FLOTILLIN-1	3	80	47		41											
	Q60634	A3497	Mm.130227	131560	Flot2	FLOTILLIN-2	1	81														
	P97841	A0107	Mm.311840	605445	Dlgap1	GKAP	5		48	37	42	17			9024896	24			20			
	Q60760	A0270	Mm.273117	601523	Grb10	GRB10	1		49													
	Q60631	A0126	Mm.6900	108355	Grb2	GRB2	1						6			25						
	O35145	A0162	Mm.2720	604641	Mapk8ip1	JIP-1	0									26						
	Q5EQR5	A3495	Mm.272712	--	2810038K19Rik	LIME	1	82														1
	Q9Y2A7	A0545	Mm.26203	604891	NCKAP1	NAP-1	4	83	50	38	43								21			
	Q9EPM9	A3492	Mm.256536	604889	Nbea	NBEA	1	84														
	P68040	A0124	Mm.5305	176981	GNB2L1	RACK1	0									27						
	O94936	A4033	Hs.511386	--	KIAA0856	RC3	3	85		39	44											
	Q8R550	--	Mm.296495	300374	Sh3kbp1	SH3KBP1	2		51	40												
	P98083	A0171	Mm.86595	600560	Shc1	SHC	1								10617960							
	Q62417	A0100	Mm.210815	605264	Sorbs1	SORBS1	2		52		45											
	Q62245	A0159	Mm.60975	182530	Sos1	SOS1	1								10617960							
	Q60803	A0862	Mm.27431	601896	Traf3	TRAF-3	3			41	46				11760165							
	Q99ME2	--	Mm.193395	--	Wdr6	WDR6	1		53													
	Q99P24	A0009	Mm.46044	604001	Akap9	YOTIAO	1								9482788	28						
14-3-3	P35213	A0467	Mm.34319	601289	Ywhab	14-3-3 beta	4	86	54	42	47							14	22			
	P62260	A0280	Mm.234700	605066	YWHAE	14-3-3 epsilon	3	87	55	43						29						
	P68510	A0633	Mm.332314	113508	YWHAH	14-3-3 eta	3	88	56	44						30		15				
	P61982	A0362	Mm.233813	605356	YWHAG	14-3-3 gamma	3	89	57	45						31		16			5	
	Q9JJ20	A0901	Mm.44482	601290	Sfn	14-3-3 sigma	1			46												
	P35216	A0907	Mm.289530	--	Ywhaq	14-3-3 theta	3	90	58	47								17				
	P35215	A0361	Mm.332735	601288	Ywhaz	14-3-3 zeta	4	91	59	48	48					32		18			6	
Ser/Thr Kinases																						
	Q9JPV4	A3502	Mm.116833	--	AAK1	AAK1	1	92														
	Q60823	A0496	Mm.177194	164731	Akt2	AKT2	1	93								33						
	P11798	A0011	Mm.131530	114078	Camk2a	CAMKII alpha	6	94	60	49	49	18	7	7	8209822	34			23			
	P28652	A1851	Mm.4857	607707	Camk2b	CAMKII beta	7	95	61	50	50	19		8	8209822	35			24			
	P15791	A1853	Mm.255822	607708	Camk2d	CAMKII delta	5	96	62	51	51	20							25			
	P11730	A1852	Mm.235182	602123	Camk2g	CAMKII gamma	5	97	63	52	52	21							26			
	O54875	A3504	Mm.27397		Cdc42bpb	CDC42BPB	1	98														
	P49615	A0667	Mm.298798	123831	Cdk5	CDK5	1			53												
	Q92772	--	Mm.44963	603442	CDKL2	CDKL2	1		64													
	O76039	--	Mm.336310	300203	CDKL5	CDKL5	1		65													
	P97633	A0896	Mm.26908	600505	Csnk1a1	CK1	1	99														
	Q60737	A0753	Mm.23692	115440	Csnk2a1	CK2 alpha	2	100	66									19				
	P67871		Mm.7405	115441	CSNK2B	CK2 beta	1				53									7		
	Q99PS2	A4036	Mm.30199	600863	Csnk1e	CSNK1E	2	101	67													
	Q9JLM8	A3506	Mm.295263	604742	Dcamkl1	DCAMKL1	4	102	68	54	54											
	Q05512	A3505	Mm.258986	600526	Mark2	EMK	3	103	69	55												
	Q91YW5	A0265	Mm.6385	601795	Mapk3	ERK1	0									36						
	P63086	A0161	Mm.196581	176948	Mapk1	ERK2	1								10617960	37						
	P26817	A0813	Mm.254144	109535	Adrbk1	GRK2	3	104		56					1403099							
	Q9WV60	A0222	Mm.200770	605004	Gsk3b	GSK3 BETA	1			57						38						
	O09110	A0491	Mm.18494	602315	Map2k3	MAP2K3	0									39						
	Q60700	A1730	Mm.172897	600447	Map3k12	MAP3K12	1		70													
	P97820	A0177	Mm.19073	604656	Map4k4	MAP4K4	1		71													
	Q9JM92	A3503	Mm.42967	--	Map4k6	MAP4K6	3	105		58	58											
	Q61831	A0489	Mm.39253	602897	Mapk10	MAPKP49	0									40						
	P31938	A0266	Mm.248907	176872	Map2k1	MEK1	1		72							41						
	Q63932	A0267	Mm.275436	601263	Map2k2	MEK2	0									42						
	O35406	A0164	Mm.3906	603014	Map2k7	MKK7	0									43						
	Q62767	A0287	Mm.170276	602747	Dusp4	MKP2	0									44						
	Q63092	A3499	Mm.274540	--	1G5	NOVEL CAMK	2	106		59												
	Q9Z2A0	A0260	Mm.10504	605213	Pdpk1	PDK-1	0									45						
	P12367	A0305	Mm.253102	176910	Prkar2a	PKA-R2A	1								1512224							
	P12369	A0306	Mm.25594	176912	Prkar2b	PKA-R2B	4	107	73			22			1512224	46						
	P31750	A0189	Mm.6645	164730	Akt1	PKB	1								12711090							
	P20444	A0307	Mm.222178	176960	Prkca	PKCALPHA	3		74	60					11196693							
	P04411	A0168	Mm.207496	176970	Prkcb1	PKCBETA	4	108	75	61					3778477	47						
	P16054	A0144	Mm.287660	176975	Prkce	PKCEPSILON	3	109		62					12577062	48						
	P63319	A0259	Mm.7980	176980	Prkcg	PKCGAMMA	6	110	76	63	59	23			3778477	49						
	P68181	A4037	Mm.16766	176892	PRKACB	PRKACB	2	111							1512234	50						
	Q63787	A0304	Mm.30039	189830	Prkar1a	PRKAR1A	1			64												
	Q63433	A0871	Mm.213000	601032	Pkn1	PRKCL1	1		77													
	Q61410	A3500	Mm.263002	601591	Prkg2	PRKG2	1	112														
	Q99N57	A0138	Mm.184163	164760	Raf1	RAF1	0									51						
	P70336	A3501	Mm.276024	604002	Rock2	ROCK-2	2	113	78													
	Q9CSY0	A0337	Mm.374825	608938	Rps6kb1	RPS6KB1	1								12711090							

Appendix 7

Protein Class	UniProt Acc	PPID Acc	Unigene cluster ID	OMIM	Gene name	Abbreviated name	Confidence (# of times identified)	Postsynaptic density datasets								Complexes			Other structures			# TM helices
								G2C #	Yosh #	Jordan #	Peng #	LI #	Satoh #	Walik #	PSD from Literature	MASC #	AMPA #	mGluRS #	Synaptosome	Clathrin Vesicle	Mitochondria	
	P18664	A0268	Mm.328476	300075	Rps6ka3	RSK	1								10617950	82						
	Q9Z1W9	A0166	Mm.198414	607848	Stk39	STK39	1		79							53						
	Q9WUN2	A3507	Mm.34580	604834	Tbk1	TBK1	2	114		65												
	Q9UKE5		Mm.126193	–	TNIK	TNIK	2			66	67											
	Q8TBQ5	A3529	Mm.260934	608128	DUET	TRAD	1		80													
Tyr Kinase																						
	P25911	A0522	Mm.317331	165120	Lyn	LYN	1								14629711							
	P06240	A0463	Mm.293753	153390	Lck	LCK	1								14629711							
	P16879	A3508	Mm.48757	190030	Fes	FES	1	115														
	P39688	A0019	Mm.4848	137025	Fyn	FYN	2	116							7488181							
	Q9QVP9	A0030	Mm.21613	601212	Ptk2b	PYK2	1								12576483	64						
	P05480	A0018	Mm.22845	190090	Src	SRC	2	117							14629711	55						
	Q04736	A0553	Mm.4558	164880	Yes	YES	1								14629711							
Protein Phosphatases																						
	Q9D7X3	–	Mm.196295	600183	Dusp3	DUSP3	1			67												
	Q8R3L3	–	Mm.266191	608867	Dusp10	DUSP10	1		81													
	Q8VEL2	A4016	Mm.197816	–	1110061O04Rik	FLJ90311	1	118														
	P63329	A0123	Mm.331389	114105	Ppp3ca	PP2B	3	119	82						9427241	66						
	O35299	A0289	Mm.3294	600658	Ppp5c	PP5	0									57						
	Q8CB81	–	Mm.341988	–	Ppm1e	PPM1E	1			69												
	Q9Z1Q2	A0147	Mm.277629	176875	Ppp1ca	PPP1CA	7	120	83	69	59	24	9		8393087							
	P62143	A4043	Mm.241931	600590	PPP1CB	PPP1CB	4	121	84	70					8393087							
	P63088	A4042	Mm.280784	176914	Ppp1cc	PPP1CC	5	122	85	71	69				8393087	68						
	O35867	A0510	Mm.332901	602468	Ppp1r9a	PPP1R9A	2			72	60											
	O35274	A0273	Mm.229087	603325	Ppp1r9b	PPP1R9B	6	123	86	73	61	25			12711090							
	P13353	A0132	Mm.260288	176915	Ppp2ca	PPP2CA	2	124		74						59						
	Q96CH3	A0295	Mm.294138	605983	PPP2R1A	PPP2R1A	3	125	87	75						60						
	P35235	A0020	Mm.8681	176876	Ptpn11	PTP1D	1								10381839	61						
	P54830	A0118	Mm.4654	176879	Ptpn5	PTPN5	0									62						
	Q64487	A1198	Mm.184021	601598	Ptprd	PTPRD	2	126	88													1
	Q9EQ17	A0223	Mm.29855	179590	Ptprf	PTPRF	1	127														1
	Q9QW00	A2624	Mm.258771	601576	Ptpsigma	PTPRS	3	128			62		9									2
	Q9WUT8	A1202	Mm.41539	176891	Ptprz1	PTPZETA	2	129			63											
	O60228	A3509	Mm.35483	603560	SBF1	SBF-1	1	130														
G-proteins and Modulators																						
G-proteins	Q91Z61	–	Mm.44995	607862	Diras1	DIRAS1	1			76												
	Q96HU8	A4017	Mm.25648	–	DIRAS2	DIRAS2	1	131														
	P08752	A0055	Mm.196464	139360	Gnai2	GNAI2	2	132							1326762						8	
	P18872	A0043	Mm.291445	139311	Gnao1	GNAO1	5	133	89	77	64		10					20	27			
	P21279	A0042	Mm.195898	600998	Gnaq	GNAQ	1	134										21			9	
	P63095	A0051	Mm.125770	139320	GNAS	GNAS1	1	135								63						
	O70443	A0052	Mm.32595	139160	Gnaz	GNAZ	2	136			65											
	Q9QWG8	A0069	Mm.2344	139380	Gnb1	GNB1	4		90		66	26	11			64			28	10		
	P62880	A0070	Mm.30141	139390	GNB2	GNB2	2	137			67					65			29	11		
	P29387	A0072	Mm.139192	–	Gnb4	GNB4	0									66						
	P54314	A0073	Mm.17604	604447	Gnb5	GNB5	2	138					12									
	Q61411	A0200	Mm.334313	190020	Hras1	HRAS	1		81							67						
	P01118	–	Hs.505033	190070	KRAS	KRAS2	3		92	78	68											
	Q08969	A3512	Mm.2045	608435	Miras	MIRAS	1	139														
	Q9NYS0	–	Mm.29648	604496	NKIRAS1	NKIRAS1	1					27										
	P08556	A0199	Mm.256975	164790	Nras	NRAS	1		83													
	P53994	A0363	Mm.240224	179509	Rab2	RAB2	0									68			30	12		
	P63012	A0037	Mm.5083	179490	Rab3a	RAB3	2	140		79						69		22	31			
	O88565	A0359	Mm.329123	179512	Rab5a	RAB5	1	141														
	Q9JJD4	A0364	Mm.28650	179513	Rab6a	RAB6A	0									70						
	P55258	A0620	Mm.162811	165040	Rab8a	RAB8	1	142														
	P61027	A0621	Mm.74596	–	Rab10	RAB10	1			80								23	32	13		
	P35284	A3513	Mm.248313	–	Rab12	RAB12	1	143														
	Q8JKM7	A0365	Mm.143789	–	Rab37	RAB37	0									71						
	P15154	A0194	Mm.292510	602048	RAC1	RAC1	1			81						72				14		
	Q9TU25		Mm.1972	602049	RAC2	RAC2	1				69											
	P05810	A0203	Mm.2348	179550	Rala	RAL-A	2	144			70					73						
	P62826	A0152	Mm.297440	601179	RAN	RAN	0									74						
	P62835	A0323	Mm.333868	179520	RAP1A	RAP1	1				71											
	Q8QZJ1	A0033	Mm.261448	179540	5830461H18Rik	RAP2A	0									75						
	P84096		Mm.259795	179505	RHOG	RHOG	2			82	72											
	Q9H0I7	A3974	Mm.261491	–	DKFZp586D022	RHOT1	1	145													7	
	P42208	A3515	Mm.242324	601506	SEPT2	SEPT2	3	146	94				13									
	Q9Z1S5	A3516	Mm.309707	608314	SEPT3	SEPT3	3	147			73		14									3
	P28661	A3517	Mm.2214	603696	SEPT4	SEPT4	1	148														
	Q9Z2Q6	A3518	Mm.20365	602724	SEPT5	SEPT5	6	149	95	83	74	28	15					24				
	Q9R1T4	A3519	Mm.260036	–	SEPT6	SEPT6	4	150			75	29	16									
	O55131	A3520	Mm.270256	603151	SEPT7	SEPT7	5	151	96	84	76			9				25	33			
	Q9ESF7	A3529	Mm.274399	608418	SEPT8	SEPT8	3	152			77		17									
	Q9QZR6	A3521	Mm.38450	604061	SEPT9	SEPT9	2	153			78											
	Q9NVA2	A3524	Mm.236587	–	SEPT11	SEPT11	2	154				30										
Modulators	Q9D853	–	Mm.258910	–	Arfgap3	ARFGAP3	1			85												
	Q9ESQ7	A0843	Mm.239329	607560	Arhgef2	ARHGEF2	2	155		86												
	Q08757	A0646	Mm.244068	605477	Arhgef7	ARHGEF7	2	156	97													
	Q9QUI3	A0707	Mm.44841	300429	Arhgef9	ARHGEF9	1		98													
	O75689	–	Hs.135183	608114	CENTA1	CENTA1	1		99													
	Q9JHW8	A3514	Mm.336761	605476	Centg1	CENTG1	3	157		87	79											
	Q9UPQ3	–	Mm.291135	608651	CENTG2	CENTG2	1		100													
	Q99P47	–	Mm.250703	–	CENTG3	CENTG3	2		101	88												
	Q9Z1P0	A3528	Mm.196153	606058	Rapgef4	CGEF2	3	158		89	80											
	Q88938	A0276	Mm.8321	605629	Cit	CITRON	7	159	102	90	81	31		10	9870942	78		26				
	Q8TDL2	–	Mm.29629	–	Dab2ip	DAB2IP	1		103													
	Q8IZD9		Mm.150259	603123	DOCK3	DOCK3	1				82											
	Q96JJ3		Mm.35064	606421	ELMO2	ELMO2	2			91	83											
	Q9JLJ8	A4046	Mm.138913	604656	Fmnl1	FMNL	1	160														
	Q91ZT5	A0470	Mm.256131	–	Fgd4	FGD4	0									77						
	P97379		Mm.290530	–	G3bp2	G3BP2	1				84											
	P06837	A0215	Mm.1222	162060	Gap43	GAP43	3	161	104		85					78			34	15		

Appendix 7

Protein Class	UniProt Acc	PPID Acc	Unigene cluster ID	OMIM	Gene name	Abbreviated name	Confidence (# of times identified)	Postsynaptic density datasets								Complexes			Other structures			# TM helices
								G2C #	Yosh #	Jordan #	Peng #	Li #	Satoh #	Walik #	PSD from Literature	MASC #	AMPA #	mGluR5 #	Synaptosome	Clathrin Vesicle	Mitochondria	
	P50396	A0567	Mm.205830	300104	Gdi1	GDI1	1		106									27	35	18		
	Q92272	A3532	Mm.290182	606434	Git1	GIT1	3	162		92	86											
	Q9DA59	A0067	Mm.234342	--	Gnqt2	GNG12	1		106													
	Q921M4	A3009	Mm.106376	602580	Golga2	GOLGA2	1		107													
	Q9JIF2	A0214	Mm.260934	604605	Kalirin	KALIRIN	5	163	106	93	87				10692441	79						
	Q96D85	A2428	Mm.196943	--	IQSEC1	IQSEC1	3	164	108		88											
	Q8C0H5	A3530	Mm.291713	--	KIAA1688	KIAA1688	3	165		94	89											
	O35550	A0589	Mm.7087	603616	Rabep1	RABEP1	1					32										
	Q04690	A0196	Mm.255596	162200	Nf1	NF1	2	166				33				90						
	Q923H2	A1069	Mm.143717	605991	Ngef	NGEF	1			96												
	Q8ESQ7	A2431	Hs.154658	602327	Psd	PSD	1		110													
	Q9NYI0	A2433	Mm.32525	--	PSD3	PSD3	2		111	95												
	Q7L804	--	Mm.24167	608599	RAB11FIP2	RAB11FIP2	3	167		97	90											
	O94784	A4047	Mm.148781	603854	RANBP9	RANBP9	1	168														
	Q8K2L6	--	Mm.180763	600278	Rap1ga1	RAP1GA1	1			98												
	Q8CHG7	A0034	Mm.31220	--	Rapgef2	RAPGEF2	1			99												
	Q8BU92	--	Hs.167371	606136	A330066M24Rik	RASAL2	1			100												
	P70392	--	Mm.248530	606614	Rasgrf2	RASGRF2	1		112													
	P49801		Mm.153013	603894	Rgs6	RGS6	1				91											
	O54829	A0932	Mm.7956	602517	Rgs7	RGS7	1	169														
	P49804	--	Hs.458417	607189	Rgs8	RGS8	1					34										
	Q9QZD0	--	Mm.44606	607191	Rgs17	RGS17	1		113													
	Q9ZQG0	A1159	Mm.20945	605072	Rgs19ip1	RGS19IP1	1					35										
	Q8ESK6	A0024	Mm.291291	603384	Syngap1	SYNGAP	7	170		101	92	36	18	11	9620684	81			36			
	Q60992	--	Mm.179011	600428	Vav2	VAV2	1		114													
Signalling molecules and Enzymes																						
Heat shock / Chaperones / Chaperonins	Q91YN9	--	Hs.50220	603882	Bag2	BAG2	1					37										
	P20029	A0451	Mm.330160	138120	Hspa5	BIP	4	171		102	93	38							37			
	P11983	A3541	Mm.32019	186980	Cct1	CCT1	2	172					19									
	P80314	A3540	Mm.247788	605139	Cct2	CCT2	2	173				39										
	P80318	A3544	Mm.256034	600114	Cct3	CCT3	1	174														
	P80315	A3538	Mm.296985	605142	Cct4	CCT4	2	175			94											
	P80316	A3543	Mm.282158	--	Cct5	CCT5	2	176					20									
	P80317	A3545	Mm.153159	104613	Cct6a	CCT6A	1	177														
	P80313	A3539	Mm.289900	605140	Cct7	CCT7	1	178														
	P42932	A3542	Mm.328673	--	Cct8	CCT8	1	179														
	P23927	--	Mm.178	123990	Cryab	CRYAB	1			103												
	P54102	A3536	Mm.27897	602837	Dnaja1	DNAJA1	2	180		104												
	Q9QYJ0	A3535	Mm.279692	--	Dnaja2	DNAJA2	3	181			95	40										
	Q99M87	--	Mm.246337	--	Dnaja3	DNAJA3	1		115												8	
	O54946	A1578	Mm.290110	--	Dnajb6	DNAJB6	3	182	116			41										
	Q9QYJ6	--	Mm.248776	604136	Dnajb10	DNAJB10	1		117													
	O75061	A0685	Mm.76494	606375	DNAJC6	DNAJC6	1	183														
	P08109	A0146	Mm.290774	600816	Hspa8	HSC-71	6	184	118	105	96	42			1626737				38		9	
	Q61699	A3533	Mm.270681	--	Hsph1	HSP105	1	185														
	P19226	A3550	Mm.1777	118190	Hspd1	HSP60	2	186					21						39		10	
	P38647	A2010	Mm.209419	600548	Hspa9a	HSP74	3	187		106		43							40		11	
	P11499	A2003	Mm.2180	140572	Hspcb	HSP90	3	188	119	107									41		12	
	P17879	A2007	Mm.6368	140550	Hspa1b	HSPA1	6	189	120	108		44	22	12		82						
	O43301	A3534	Mm.39739	--	HSPA12A	HSPA12A	2	190		109												
	P04792	--	Mm.13849	602195	HSPB1	HSPB1	1			110												
	Q8NVH1	A3537	Mm.21353	--	DNAJC11	DNAJC11	1	191														
	O55042	A0905	Mm.17464	163890	Snca	SNCA	1			111									42	17		
	Q6CY30	A3549	Mm.138726	600868	Ssr1	SSR1	1	192														1
	P08113	A2004	Mm.87773	191175	Tra1	TRA1	1	193														
	Q922R8	A3548	Mm.222825	--	Txndc7	TXNDC7	1	194														
Phosphodiesterases	P16330	A3553	Mm.15711	123830	Cnp	CNP-1	5	195	121	112	97	45							43			
	Q01063	A3551	Mm.247564	602056	Pde2a	PDE2A	3	196		113	98											
	Q9VBJ5	--	Mm.20181	600127	Pde4b	PDE4B	1			114												
	Q01063	--	Mm.343333	600129	Pde4d	PDE4D	1			115												
	Q9QYJ5	A3552	Mm.87161	--	Pde10a	PDE10A	1	197														
	Q8VID8	--	Mm.246613	604961	Pde11a_v1	PDE11A	1					46										
Adenylate/Guanylate cyclase	Q9WVI4	A1979	Mm.371695	601244	Gucy1a2	GUCY1A2	1	198														
Secreted / secretory / signalling	Q61361	A3556	Mm.4598	--	Bcan	BCAN	4	199	122		99	47										
	Q62059	A3555	Mm.158700	118661	Casq2	CSRQ-2	3	200	123		100											
	O88998	A3557	Mm.43278	605366	Olfm1	OLFM1	2	201			101											
	O35382	A1304	Mm.265512	608185	Sec8l1	SEC8L1	1	202														
ATP synthases	Q03265	A0377	Mm.276137	164360	Atp5a1	ATP5A1	3	203		116	102					83			44		13	
	P56480	A0378	Mm.238973	102910	Atp5b	ATP5B	4	204	124	117	103								45		14	
	P35435	A0379	Mm.12677	108729	Atp5c1	ATP5C	1	205								84			46		15	
	P56516	A0397	Mm.217787	607027	Atp6v1a	ATP6A1	4	206	125	118	104					85						
	P62815	A0400	Mm.249096	606939	Atp6v1b2	ATP6B2	4	207	126	119	105											
	Q9Z1G3	A0401	Mm.276818	603097	Atp6v1c1	ATP6C	1	208														
	P51863	A0403	Mm.17708	607028	Atp6v0d1	ATP6D	3	209			106	48				86						
	P50518	A0404	Mm.29045	108746	Atp6v1e1	ATP6E	3	210			107	49										
	P57746	A0402	Mm.311549	--	Atp6v1d	ATP6V1D	1	211														
	Q925I1	A3559	Mm.241152	--	Atic3a	ATAO3A	3	212	127			50										
NADH-Ubiquinone Oxidoreductase	Q62425	--	Mm.41926	603833	Ndufa4	NDUFA4	1			120												1
	Q9DCJ5		Mm.19834	603359	Ndufa8	NDUFA8	1				108											
	Q16795	A3563	Mm.29939	603834	NDUFA9	NDUFA9	3	213		121	109										16	
	Q99LC3	A3560	Mm.248778	603835	Ndufa10	NDUFA10	1	214													17	
	P28331	A3561	Mm.290791	157655	NDUFS1	NDUFS1	3	215		122	110										18	
	Q91WD5	A3562	Mm.21569	602985	Ndufs2	NDUFS2	1	216														
	Q9DCT2	A3564	Mm.30113	603846	Ndufs3	NDUFS3	3	217		123	111										19	
	Q9D6J6	A0431	Mm.2206	600532	Ndufv2	NDUFV2	2	218			112					87			47		20	
Polyubiquitin	P62986	--	Mm.43005	191321	Uba52	UBA52	1		128													
	Q63446	A1256	Mm.282093	191339	Rps27a	UBB	1		129													
	Q62317	A1257	Mm.331	191340	Rps27a	UBC	3	219	130		113								48		21	
Proteases / protease inhibitors	O88543	A3570	Mm.40	604995	Copa3	COPS-3	1	220														
	P46101	A3565	Mm.42078	126141	Dpp6	DPP-6	1	221														1
	Q9EQE2	A3571	Mm.19958	110900	Kel	KELL	1	222														1
	O35593	A3566	Mm.218198	607173	Pamd14	PAD-1	1	223														
	Q9R1P0	A3567	Mm.30270	--	Psma4	PSMA4	1	224														
	Q63569	--	Mm.289832	186852	Psmc3	PSMC3	1					51										

Appendix 7

Protein Class	UniProt Acc	PPID Acc	Unigene cluster ID	OMIM	Gene name	Abbreviated name	Confidence (# of times identified)	Postsynaptic density datasets: G2C #	Yosh #	Jordan #	Peng #	Li #	Satoh #	Walk #	PSD from Literature	Complexes: MASC #	AMPA #	mGluR5 #	Other structures: Synaptosome	Clathrin Vesicle	Mitochondria	# TM helices
	Q88761	A3568	Mm.280013	--	Psmd1	PSMD1	1	225														
	Q13200	--	Mm.243234	606223	PSMD2	PSMD2	1		131													
	Q99JI85	A3569	Mm.273152	--	Psmd8	PSMD8	1	226														
	Q9Z2X3	--	Mm.17640	603480	Psmd10	PSMD10	1		132													
	P48594	--	Mm.283677	600518	SERPINB4	SERPINB4	1			124												
	Q9JJI7	--	Mm.243926	606797	St14	ST14	1		133													1
	O88399	A1731	Mm.236068	--	Synj2	SYNJ2	1	227														
Development	Q05175	A3574	Mm.29586	605940	Basp1	BASP1	1	228											49			
	P18572	A3573	Mm.726	109480	Bsg	BSG	1	229														1
	Q9CYN5	A3575	Mm.86523	--	5730405I09Rik	CBCP1	1	230														
	Q9JIQ6	A3998	Mm.39388	607802	Cnnm1	CNNM1	1	231														
	O88558	A3579	Mm.37249	606322	Cyfip1	CYFIP1	3	232	134		114											
	Q8CQQ2	A4001	Mm.87417	606626	Daam1	DAAM1	1	233														
	Q9QXL0	A3578	Mm.248788	602865	Dbccr1	DBCCR-1	1	234														
	Q9D2N4	A3577	Mm.94371	601239	Dtna	DTNA	2	235							8631824							
	Q03173	--	Mm.87759	--	Enah	ENAH	1		135													
	O88741	A3576	Mm.18218	606598	Gdap1	GDAP1	3	236	136	125												2
	Q8VE33	A3980	Mm.102080	--	Gdap1l1	GDAP1L1	2	237		126												1
	Q9QZY2	A3581	Mm.41812	--	Gprin1	GPRIN1	1	238														
	Q62813	A3572	Mm.310524	603241	Lsamp	LSAMP	1	239												18		1
	Q8K377	--	Mm.292568	--	Lrrtm1	LRRTM1	1			127												1
	O43300	--	Mm.39900	--	LRRN2	LRRTM2	1		137													1
	Q86VH4	--	Mm.94135	--	LRRTM4	LRRTM4	1		138													2
	O08919	A2635	Mm.255487	604018	Numbl	NUMBL	1	240														
	Q924Y6	--	Mm.32831	605186	Wif1	WIF1	1					62										
Other signalling molecules	O35639	--	Mm.7214	106490	Anxa3	ANXA3	1		139													
	Q91VR9	--	Mm.5061	--	Arf2	ARF2	1			128												
	P61206	A0421	Mm.221298	103190	ARF3	ARF3	0									88						
	Q9WV31	A0277	Mm.25405	--	Arc	ARC	2		140						7857651	89						
	O35143	--	Mm.2171	--	Atpif1	ATPI	1		141												22	
	Q61337	A0476	Mm.4387	603167	Bad	BAD	0									90						
	O75498	A1114	Mm.9749	604515	BLNK	BLNK	1		142													
	P63030	--	Mm.288510	--	Brp44l	BRP44L	1		143												23	
	P62157	A0008	Mm.285993	114180	CaM	CALM	4		144		115	63			9427241	91		28	60			
	Q9NZT1		Mm.21075	605183	CALML5	CALML5	2			129	116											
	P47728	A0437	Mm.2755	114051	Calb2	CALRETININ	0									92		29				
	O08918	--	Mm.3527	603203	Ccng2	CCNG2	1		145													
	P27274	A1630	Mm.247265	107271	Cd59	CD59	1				117											2
	Q99KW9	A4005	Mm.334685	--	Tip	CDA08	1	241														1
	P61810	A0666	Mm.142275	603460	Cdk5r1	CDK5R1	1								14749431							
	Q9JLH7	A4007	Mm.28297	608202	Cdk5rap3	CDK5RAP3	1	242														
	Q9Z140		Mm.5249	605688	Cpne6	CPNE6	1				118											
	P47713	A0106	Mm.4186	600522	Pla2g4a	CPLA2	0									93						
	P49621		Mm.242576	604070	Dgkb	DGKB	2			130	119											
	Q61358	A0095	Mm.44736	602377	Dnm1	DNM1	5	243	146	131	120				1629737	94		30	61			
	P39052	A1927	Mm.39292	602378	Dnm2	DNM2	3			132	121				11583995			31	62			
	Q8R3T9	A1928	Mm.29567	--	Dnm3	DYN3	2			133					12646135				63			
	Q9NRX7	--	Mm.335310	607815	Eb-1	EB-1	3		147	134	122											
	Q00731	A0434	Mm.282184	192240	Vegf	EGF-164	0									95						
	P35637	A0217	Mm.277680	137070	FUS	FUS	0									96						
	P53762	A0488	Mm.250265	126110	Arnt	HIF-1	0									97						
	P35569	A0484	Mm.4952	147545	Irs1	IRS-1	0									98						
	Q9Z0J4	A0262	Mm.44249	163731	Nos1	NNOS	1								9427241	99						
	O70465	A4013	Mm.254610	--	Odz4	ODZ4	2	244			123											1
	O08957	--	Mm.232930	607409	Nrn1	NRN	2		148		124											
	Q12888	A0286	Mm.215389	605230	TP53BP1	P53BP1	0									100						
	O88377		Mm.39700	603261	Pip5k2b	PIP5K2B	2			135	125											
	Q62077	A0012	Mm.44463	172420	Plcg1	PLC G	0									101						
	Q9Z1B3	A0330	Mm.330607	607120	Plcb1	PLCB1	1		149							102		32				
	Q9JHK5	A0919	Mm.98232	173570	Plek	PLEK	1		150													
	P04925	A4012	Mm.648	176640	Prnp	PRNP	3	245	151		126											2
	O94811	--	Mm.39752	608773	Tppp	P25	2		152	136												
	Q9ES97	--	Mm.246990	604249	Rtn3	RTN3	1			137										19		2
	Q9JK11	--	Mm.192580	604475	Rtn4	RTN4	1				127											2
	P05942	--	Mm.3925	114210	S100a4	S100A4	1		153													
	P06702	A1551	Mm.2128	123886	S100A9	S100A9	1				128								54			
	P50114	A1342	Mm.235998	176990	S100b	S100B	1		154										55			
	P07092	A1547	Mm.3093	177010	Serpine2	SERPINE2	1				129											
	Q9NQ38	A4014	Mm.35369	605010	SPINK5	SPINK5	0									103						
	O88602	A0473	Mm.277338	602911	Cacng2	STARGAZIN	1								11806122	104						4
	O55106	A2120	Mm.311915	--	Strn	STRN	1			138												
	P58404	--	Mm.21612	--	Strn4	STRN4	1			139												
	Q9WUD1	A3978	Mm.277599	607207	Stub1	STUB1	2	246				64										
	O75962	A0739	Hs.130031	601893	TRIO	TRIO	1		155													
	P20156	A2164	Mm.297818	602186	Vgf	VGF	1		156													
Other enzymes	Q925C4	A3582	Mm.31374	200350	Acac	ACC1	1	247											56			
	Q91V92	A3586	Mm.282039	108728	Acly	ACLY	2	248		140									57			
	P55266	--	Mm.316628	601059	Adar	ADAR	1		157													
	Q9D1E8	A3630	Mm.24117	--	Agpat5	AGPAT5	1	249														2
	P51635	--	Mm.30085	103830	Akr1a1	AKR1A1	1					65							58			
	P24549	A3585	Mm.250866	100640	Aldh1a1	ALDH1A1	1	250														
	P05064	A1923	Mm.275831	103850	Aldoa	ALDOA	5	251	158	141	130	66						33	59			
	P05063	A0428	Mm.7729	103870	Aldoc	ALDOC	3	252		142		67				105		34	60			
	P11274	A0902	Mm.333722	151410	BCR	BCR	1			143												
	Q99PU5	A3593	Mm.20592	--	Lpd	BG1	1	253														
	O88738	--	Mm.290908	605638	Birc6	BIRC6	1		159													
	P24270	--	Mm.4215	115500	Cat	CAT	1		160													
	Q8VDP6	--	Mm.28219	605893	Cdipt	CDIPT	1			144												3
	P07335	A3693	Mm.16831	123280	Ckb	CKB	3			145	131				1629737				61	20		
	Q8TE79	A3634	Mm.11827	--	Comtd1	COMTD1	1	254														1
	P52825	--	Mm.307620	600650	Cpt2	CPT2	1		161													
	P97427	A3605	Mm.290995	602462	Crmp1	CRMP1	4	255	162	146	132								62			
	Q9D5V5	--	Mm.218910	601741	Cul5	CUL5	1			147												
	Q9NQC7	--	Mm.24282	605018	CYLD	CYLD	3		163	148	133											

Appendix 7

Protein Class	UniProt Acc	PPID Acc	Unigene cluster ID	OMIM	Gene name	Abbreviated name	Confidence (# of times identified)	Postsynaptic density datasets								Complexes			Other structures			# TM helices
								G2C #	Yosh #	Jordan #	Feng #	Li #	Satoh #	Walik #	PSD from Literature	MASC #	AMPA #	mGluR5 #	Synaptosome	Clathrin Vesicle	Mitochondria	
	Q54678	--	Mm.214016	601771	Cyp1b1	CYP1B1	1		164														1
	P16381	A3602	Mm.106054	--	D1Pas1	D1PAS1	1	268															
	Q54734	A3619	Mm.7236	602202	Ddx6	DDX6T	2	267			134												2
	Q91VR5	A3656	Mm.251255	601257	Ddx1	DDX1	3	266	185		136												
	Q62167	A3653	Mm.289962	300160	Ddx3x	DDX3	4	259	186	149	136												
	Q61656	A0814	Mm.220038	180630	Ddx5	DDX5	2	260			137												
	Q08550	A1300	Mm.314923	601441	Dgkz	DGKZ	1		187														
	Q99J47	A3651	Mm.21475	--	9C0309	DKFZP566C084	1	261															1
	Q99QF9	A3641	Mm.29676	--	DKFZp761F069	DKFZP761F069	1	262															5
	Q9WU83	A3638	Mm.201322	603503	DPM1	DPM1	1	263															
	O35098	A3606	Mm.250414	608407	Dpysl4	DPYSL4	2	264			138									63			
	Q9JMG8	A3610	Mm.27732	608383	Dpysl5	DPYSL5	1	265												64			
	P17182	A3615	Mm.70666	172430	Eno1	ENO1	2	266	188											66	21		
	P21550	A3617	Mm.251322	131370	Eno3	ENO2	1		189											65	22		
	P17183	A3616	Mm.3913	131360	Eno2	ENO3	1		170											67	23		
	P70426	--	Mm.4336	133701	Ext2	EXT2	1		171														1
	Q63835	A3659	Mm.267478	604443	Acsl6	FACL6	2	267		150													1
	Q6QR0	A3584	Mm.236443	600212	Fasn	FASN	2	268		151													
	Q6ESW4	A3620	Mm.32840	--	2610037M15Rik	FLJ10642	1	269															
	P28037	A3594	Mm.30035	600249	Aldh1l1	FTHFD	1	270															
	Q6Z148	A0444	Mm.35345	604599	Bat5	G9A	1	271									106						
	P15105	A0424	Mm.210745	138290	Glul	GLNS	5	272	172	152	139	58					107			68	24		
	Q62VW5	A3592	Mm.28188	601336	Gcs1	GCS1	2	273	173														1
	Q62339	A1969	Mm.29457	605482	Gsto1	GSTO1	1					59											
	O54865	A1980	Mm.9445	139397	Gucy1b3	GUCY1B1	1			153													
	O70351		Mm.6994	300258	Hadh2	HADH2	1				140									69			
	Q8CGI8	--	Mm.28109	--	Hars2	HARS2	1		174														
	P97852		Mm.277857	601860	Hsd17b4	HSD17B4	1				141												
	O70503	A3633	Mm.22505	--	Hsd17b12	HSD17B12	1	274															1
	P97518	--	Mm.183042	602064	Impa1	IMPA1	1		175														
	Q62784	A3587	Mm.150420	600916	Inpp4a	INPP4A	1	275															
	O60275	A2429	Mm.33027	--	IQSEC2	IQSEC2	3	276		154	142												
	O75043	A3640	Mm.140138	607813	KIAA0455	LPPR4	1	277															4
	P06151	A0370	Mm.29324	150000	Ldha	LDH1	1									9371636	106			70			
	P16125	A0369	Mm.9745	150100	Ldhb	LDHB	1									9371636	109			71			
	P21396	A3672	Mm.21108	309850	Maoa	MAOA	1	278												72			
	Q99MR8	A3604	Mm.249016	210200	Mccc1	MCCC1	1	279															
	P48318	A0635	Mm.272120	605363	Gad1	MGAD	1	280															
	Q6CPL4	A1948	Mm.218286	604564	Mgat3	MGST3	3		176	155	143												4
	P34884	--	Mm.2326	153620	Mif	MIF	1		177											73			
	P46935	A1245	Mm.279923	602278	Nedd4	NEDD4	1	281															
	Q7Z5F1	--	Mm.98658	606384	NEDD4L	NEDD4L	1		178														
	Q01768	--	Mm.1260	156491	Nme2	NME2	1						23							74			
	Q91VJ8	--	Mm.29824	606793	Npepps	NPEPPS	1			156													
	Q6CGU4	A3635	Mm.27427	--	Dhrs4	DHRS4	1	282															
	O15294	--	Mm.259191	300255	OGT	OGT	2		179	157													
	Q08642	--	Mm.2296	607935	Pad2	PADI2	1		180														
	Q91ZA3	A3621	Mm.23876	232000	Pcca	PCCA	1	283															
	Q9WUA3	A1926	Mm.273874	171840	Pfkp	PFK-C	4	284	181	158	144									75			
	P12382	A1925	Mm.269649	171860	Pfkl	PFK-L	2	285		159							110			76			
	P47858	A0261	Mm.272582	232800	Pfkm	PFK-M	2	286		160										77			
	Q9DBJ1	--	Mm.315962	172250	Pgam1	PGAM1	3		182	161		60								78	25		
	P15259	--	Mm.219627	261670	PGAM2	PGAM2	1		183														
	P42337	A0153	Mm.260521	171834	Pik3ca	PI3-K	3	287		162		61					111						
	O09062	A3658	Hs.529436	600286	Pik4ca	PI4-K	3	288		163	145												
	Q9CXY9	A3622	Mm.331447	605087	Pigk	PIGK	1	289															1
	O70172	--	Mm.313977	603140	Pip5k2a	PIP5K2A	1			164													
	O70161	A3591	Mm.29836	606102	Pip5k1c	PIP5K1C	1	290															
	P53657	A0427	Mm.8359	266200	Pklr	PKLR	1									1629737	112			79			
	Q9CR09	--	Mm.22284	606114	Pkp4	PKP4	1		184														
	Q99JY8	A3632	Mm.348326	607125	Ppap2b	PPAP2B	1	291															6
	P10111	--	Mm.5246	123840	Ppia	PPIA	3		185	166		62								80	26		
	O88541	--	Mm.325816	123841	Ppib	PPIB	3		186	168	146												1
	Q60700	A3631	Mm.287178	311850	Prps1	PRPS1	1	292															
	Q62ZW5	A3657	Mm.127731	179035	Pycr1	PYCR1	1	293															
	P53534	A3588	Mm.222584	138050	Pygb	PYGB	1	294												81			
	Q925F3	--	Mm.214932	--	Rnf144	RNF144	1		187														1
	Q9EPK2	A3625	Mm.268141	312600	Rp2h	RP2	1	295															
	P07153	A3639	Mm.188544	180470	Rpn1	RPN1	1	296															1
	P25235	A3611	Mm.22130	180490	Rpn2	RPN2	1	297															3
	Q91YM8	--	Mm.197486	180410	Rrm1	RRM1	1		188														
	Q9Z2X7	--	Mm.131443	606477	Sir	SIR	1					63											
	P35700	A0371	Mm.30929	176763	Prdx1	TDPX2	2			167		64					113						
	P50137	--	Mm.290692	606781	Tkt	TKT	2		189	169										82			
	O76062	A2012	Mm.301370	603414	TM7SF2	TM7SF2	1		190														7
	P17751	A0433	Mm.4222	190450	Tpi1	TPI1	3	298		169		66					114			83	27		
	Q61171	A0372	Mm.270130	600538	Prdx2	TPX	1			170							115			84			
	Q9CZY3	--	Mm.278763	602995	Ube2v1	UBE2V1	1					66											
	Q9R172	A1052	Mm.258530		Uble1a	UBLE1A	1					67											
	Q00981	--	Mm.29807	191342	Uchl1	UCHL1	3		191	171		68								85			
	P56389	--	Mm.3571	601447	Usp5	USP5	1			172													
Mitochondrial enzymes	P17764	A3694	Mm.293233	203750	Acat1	ACAT1	2	299	192											86		24	
	Q99KI0	A0429	Mm.154561	100850	Aco2	ACO2	3	300	193			69					116			87		25	
	P39089	--	Mm.29189	103000	Ak1	AK1	3		194	173		70								88		26	
	Q6QWW0	A0432	Mm.186185	602316	Prdx6	AOP2	1			174							117			89		27	
	O55028	--	Mm.8903	Bckdk	BCKDK		1		195													28	
	P29147	A3702	Mm.293470	603063	Bdh	BDH	2	301		175												29	
	P30275	A3690	Mm.252145	123290	Ckmt1	CKMT1	6	302	196	176	147	71		13						90		30	
	Q9CZU6	A3688	Mm.58836	118950	Cs	CS	1	303														31	
	Q9CQM3	A3698	Mm.29196	123960	Cyc1	CYC1	1	304														32	1
	P53395	A3675	Mm.3636	248610	Dbt	DBT	1	305														33	
	Q64591	A3699	Mm.24395	222745	Decr1	DECR1	1	306														34	
	P08461	A0605	Mm.285076	109720	Dlat	DLAT	1	307									118			91		35	
	O08749	A3674	Mm.3131	246900	Dld	DLD	1	308														36	
	Q01205	A3682	Mm.296221	126063	Dlst	DLST	1	309												92		37	

Appendix 7

Protein Class	UniProt Acc	PPID Acc	Unigene cluster ID	OMIM	Gene name	Abbreviated name	Confidence (# of times identified)	Postsynaptic density datasets								Complexes			Other structures			# TM helices
								G2C #	Yosh #	Jordan #	Peng #	Li #	Satoh #	Walik #	PSD from Literature	MASC #	AMPA #	mGluR5 #	Synaptosome	Clathrin Vesicle	Mitochondria	
	Q08553	A0442	Mm.235263	602463	Dpysl2	DPYSL2	6	310	197	177	148		34		11799185	119			93	28	38	
	Q08600	A3783	Mm.4449	600440	Endog	ENDOG	1	311													39	
	Q9ZCT9	A3779	Mm.28336	231675	Etfdh	ETFDH	1	312													40	
	P26492	A3670	Mm.32102	606365	Gls2	GA	1	313													41	
	P16858	A0331	Mm.333399	138400	Gapd	GAPDH	7	314	196	178	149	72		14	1629737	120			94		42	
	P26443	A3680	Mm.10600	138130	Glud1	GLUD	5	315	199	179	150	73									43	
	P05202	A0423	Mm.230169	138150	Got2	GOT2	3		200	180		74				121			96		44	
	Q64521	A3667	Mm.3711	138400	Gpd2	GPD2	1	316													45	1
	Q70325	--	Mm.332910	138322	Gpx4	GPX4	1			181												
	Q64428	A3683	Mm.200497	600890	Hadha	HADHA	3	317	201			75									46	
	Q60587	A3684	Mm.291463	143450	Hadhb	HADHB	1	318													47	
	P17710	A3663	Mm.196605	142600	Hk1	HK1	3	319	202		151								86		48	
	Q9EQK1	A3690	Mm.2966	147650	Idh2	IDH2	1	320													49	
	Q9D6R2	A3629	Mm.279195	601149	Idh3a	IDH3A	1	321											97		50	
	Q91VA7	A3668	Mm.29590	604526	Idh3b	IDH3B	2	322	203												51	
	P70404	A3699	Mm.14825	300089	Idh3g	IDH3G	1	323											98		52	
	Q8R2M5	A3780	Mm.331142	--	2010002H18Rik	ILES	1	324													53	
	P08249	A3696	Mm.297096	154100	Mdh2	MDH2	3	326	204			76							89		54	
	P00405	A3778	Mitochondrial	516040	Mtco2	MTCO2	1	328											100		56	2
	Q60597	A3664	Mm.276348	203740	Ogdh	OGDH	1	327													55	
	Q9JLD0	A3665	Hs.17660	--	KIAA1290	OGDHL1	1	328													57	
	P04166	--	Mm.20242	--	Cyb5m	CMB5	1		206												59	1
	Q05920	A3662	Mm.1845	266150	Pc	PC	1	329													58	
	Q9Z0X1	A3781	Mm.240434	300169	Pdcd8	PDCD8	1	330													60	
	P35486	A3685	Mm.34775	312170	Pdha1	PDHA1	1	331											101		61	
	P49432	A3687	Mm.301527	179060	Pdhb	PDHB	1	332											102		62	
	P97700	A3731	Mm.296982	604165	Slc25a11	SLC25A11	2	333							11799185						63	1
	P07895	--	Mm.290876	147460	Sod2	SOD2	2		205			77							103		64	
	Q9Z2I9	A3621	Mm.38951	603921	Sucla2	SUCLA2	2	334				78									65	
	Q9WUM5	A3697	Mm.29845	--	Suclg1	SUCLG1	3	335	207			79							104		66	
	Q9CZ13	A3676	Mm.194811	191328	Uqcrc1	UQCRC1	1	336													67	
	P52551	A3677	Mm.334206	191329	Uqcrc2	UQCRC2	2	337		182									106		68	
Transcription and Translation																						
Transcription Elements	Q9JK31	A3789	Mm.173271	--	Atf7ip	ATF7IP	1	338														
	P46100	A2062	Mm.10141	300032	Atrx	ATRX	1		208													
	Q8UIG0	A2168	Mm.40331	605681	Baz1b	BAZ1B	1		209													
	Q61336	A0439	Mm.36241	--	Bcap37	BCAP37	5	339	210	183	152		25			122						
	Q62403	A3590	Mm.15456	606323	Cyfip2	CYFIP2	4	340	211	184	153											
	Q9JI19	A3798	Mm.329656	608295	Fibp	FIBP	1	341														
	P58501	A0298	Mm.347	--	Gcfc	GCFC	0									123						
	Q8C0E1	A3792	Mm.311439	160994	Hnrpm	HNRPM	1	342												29		
	O54910	A1330	Mm.5704	604546	Nfkbie	NFKBIE	1		212													
	Q9DC29	A3794	Mm.192706	608205	Nrbf1	NRBF1	1	343														
	P59759	--	Mm.270643	--	Mrtfb	MRTFB	1			185												
	Q9Z1N2	A3790	Mm.294154	601902	Orc1l	ORC1L	1	344														
	P24142	A3793	Mm.263862	176705	Phb	PHB	3	345		186		80							108		69	
	P42669	A3795	Mm.231802	600473	Pura	PURA	4	346	213	187		81										
	O35295	A3796	Mm.296150	608887	Purb	PURB	1	347														
	Q9R0S1	A3788	Mm.29827	602181	Snd1	SND1	2	348	214													
	P40646	A2191	Mm.42162	--	Sox7	SOX7	1		215													
	Q8C3V4	--	Mm.277406	600555	Stat1	STAT1	1		216													
	P40630	A2186	Mm.229292	600438	Tfam	TFAM	1			188												
	Q8K055	A3799	Mm.22786	602290	Trim32	TRIM32	1	349														
	Q9QZP7	A2922	Mm.254997	--	Zfp386	ZFP386	1		217													
Other nuclear/DNA binding	Q99MR0	--	Mm.31646	--	Actl6b	ACTL6	1					82										
	Q9Z1N7	A2027	Mm.350789	--	Arid3b	ARID3B	1		218													
	O54962	--	Mm.299167	603811	Banf1	BANF1	2		219	189												
	Q8NE35	--	Mm.371677	--	CPEB3	CPEB3	1		220													
	Q6CW03	A3213	Mm.14910	606062	Cspp6	CSPP6	1		221													
	P47875	A1021	Mm.196484	123676	Csrp1	CSRP1	2			190		83										
	Q9Z1X8	A2251	Mm.262287	607112	Fbxo2	FBXO2	1		222													
	Q8TF61	--	Hs.23158	--	FBXO41	FBXO41	1			191												
	P84249	A1373	Hs.546259	601128	H3-IX	H3F3A	1		223													
	Q8TDG4	A2152	Hs.480101	606769	HEL308	HEL308	1		224													
	P43274	--	Mm.1401	142220	Hist1h1e	HIST1H1E	1			192												
	P20001	--	Mm.261965	602786	HIST1H2AE	HIST1H2AE	1		225													
	Q811M0	--	Mm.227295	142750	Hist1h4h	HIST1H4H	1			193												
	Q9D2U9	--	Mm.28022	601831	Hist3h2ba	HIST3H2BA	1		226													
	O70592	A2547	Mm.316306	--	Alfa	HNRPA3	1			194										30		
	O35737	--	Mm.21740	601035	Hnrph1	HNRPH1	1					84										
	P70333	--	Mm.315909	601036	Hnrph2	HNRPH2	1					85										
	Q9JYF4	--	Mm.277906	--	JMJD1B	JMJD1B	1		227													
	Q9QWV4	--	Mm.10414	601402	Mlf1	MLF1	1		228													
	Q999Z1	--	Mm.218940	--	MMS19L	MMS19L	1		229													
	P28656	--	Mm.337558	--	Nasp1l1	NAP1L1	1		230													
	Q9EPT6	--	Mm.31979	607538	Ndel1	NDEL1	1		231													
	Q60862	A3791	Mm.3411	601182	Orc2l	ORC2L	1		232													
	P60335	--	Mm.274146	601209	Pcbp1	PCBP1	1					86										
	O00541	--	Mm.28659	605819	PES1	PES1	1		233													
	Q86Y13	--	Hs.154658	602327	PSD	PSD	1		234													
	Q61493	--	Mm.288788	602776	Rev3l	REV3	1			195												
	Q8JJC5	--	Hs.302772	--	SH3BGRL2	SH3BGRL2	1		235													
	P28370	A2045	Mm.229151	300012	SMARCA1	SNF2L1	1					87										
	Q60596	--	Mm.333661	194360	Xrcc1	XRCC1	1		236													
	Q99K28	--	Mm.43636	606908	Zfp289	ZFP289	1			196												
	Q9Y4E5	--	Mm.289103	--	ZNF451	ZNF451	1		237													
Elongation / RNA binding	P62632	A3801	Mm.2645	602959	Eef1a2	EEF1A2	5	350	238	197	154	88								31		
	O70251	A3802	Mm.2718	600955	Eef1b	EEF1B	1	351														
	P58252	A3804	Mm.289431	130610	Eef2	EEF2	4	352	239	198					12711090							
	P10126	A1820	Mm.196614	130590	Eef1a1	EF1A	3	353			155				1629737							
	Q9D8N0	A1968	Mm.247762	130593	Eef1g	EF1G	1	354														
	O70194	A3805	Mm.3955	603915	Eif3s7	EIF357	1	355														
	P10630	--	Mm.260064	601102	Eif4a2	EIF4A2	1		240													
	Q922K6	A3806	Mm.290022	603928	Eif4b	EIF4B	1	356														

Appendix 7

Protein Class	UniProt Acc	PPID Acc	Unigene cluster ID	OMIM	Gene name	Abbreviated name	Confidence (# of times identified)	Postsynaptic density datasets								Complexes			Other structures			# TM helices
								G2C #	Yosh #	Jordan #	Peng #	LI #	Satoh #	Walikonis #	PSD from Literature	MASC #	AMPA #	mGluR5 #	Synaptosome	Clathrin Vesicle	Mitochondria	
	P63074	A0973	Mm.3941	133440	Eif4e	EIF4E	1					89										
	P29341	A2508	Mm.371570	604679	Pabpc1	PABPC1	1	357														
	O54725	--	Mm.257276	605199	Sfpq	SFPQ	1			186												
Ribosomal Proteins	Q8BHL0	--	Mm.4071	150370	Larp1	LARP1	2		241		154									92		
	Q63742	A3286	Mm.172720	--	RBP34	RBP34	1	368														1
	P27659	--	Mm.290899	604163	Rpl3	RPL3	1		242													
	P50878	A3808	Mm.280083	180479	Rpl4	RPL4	3	368	243		157											
	P47962	A1222	Mm.4419	603634	Rpl5	RPL5	3	360	244		158											
	P47911	A3810	Mm.262021	603703	Rpl6	RPL6	4	361	245	200	159									35		
	P14148	A0366	Mm.290772	604164	Rpl7	RPL7	3	362	246		160					124				34		
	P12970	A3811	Mm.725	185640	Rpl7a	RPL7A	2	363			161									36		
	Q6ZWY7	A3812	Mm.30066	604177	Rpl8	RPL8	2	364			162											
	P51410		Mm.300271	603686	Rpl9	RPL9	1				163											
	Q6ZWY7		Mm.100113	312173	Rpl10	RPL10	1				164											
	P53026	A3813	Mm.306955	--	Rpl10a	RPL10A	3	366	247		165											
	Q9CXW4	--	Mm.276856	604175	Rpl11	RPL11	3		248	201	166											
	P23358		Mm.286100	180475	Rpl12	RPL12	2			202	167									38		
	P47963	A0367	Mm.319719	113703	Rpl13	RPL13	2	368			169					125						
	P19253	A0368	Mm.180456	--	Rpl13a	RPL13A	0									126						
	Q63507	A3814	Mm.289810	--	Rpl14	RPL14	2	367	249													
	Q6ZWZ7	--	Mm.276337	603661	Rpl17	RPL17	1			203												
	P35980	--	Mm.349277	604179	Rpl18	RPL18	2		250		169									37		
	P14118	--	Mm.309019	180466	RPL19	RPL19	2		251		170											
	P67985	--	Mm.307846	180474	Rpl22	RPL22	2		252		171											
	P23131	--	Mm.140380	603662	RPL23	RPL23	2		253		172											
	P62751		Mm.244545	602326	RPL23A	RPL23A	1				173											
	P83731		Mm.282814	604180	RPL24	RPL24	1				174											
	Q9CSM4	--	Mm.340658	607526	Rpl27	RPL27	2		254	204												
	P62889	--	Mm.358632	180467	RPL30	RPL30	1		255											38		
	P62900		Mm.298467	--	RPL31	RPL31	1				176											
	P47964	--	Mm.11376	--	Rpl36	RPL36	1		256													
	P63174		Mm.236817	604182	RPL38	RPL38	1				178											
	P14869	A3807	Mm.5286	180510	Rplp0	RPLP0	3	358			177	90										
	P47955	--	Mm.3158	180520	Rplp1	RPLP1	1		257													
	P02401		Mm.341719	180530	Rplp2	RPLP2	1				178											
	P25444	A3815	Mm.1129	603024	Rps2	RPS2	2	369			179											
	P62909	A3816	Mm.236868	600454	Rps3	RPS3	4	370	259	206	180									39		
	P97351	A3817	Mm.6957	180478	Rps3a	RPS3A	2	371			181											
	P47961	A3818	Mm.66	312760	RPS4	RPS4X	2	372	239													
	P46782	--	Mm.5291	603630	RPS5	RPS5	1		260											40		
	P10660	--	Mm.325584	180460	RPS6	RPS6	2		261		182											
	P23921	--	Mm.371579	603658	RPS7	RPS7	1		262													
	P09058	A3823	Mm.260904	600357	RPS8	RPS8	4	373	263	206	183									41		
	P46781	--	Mm.13944	603631	RPS9	RPS9	2		264		184											
	P63326	A1668	Mm.275810	603632	Rps10	RPS10	2		266		185											
	P63324	--	Mm.371846	603660	Rps12	RPS12	1		268													
	P62278		Mm.14796	180476	RPS13	RPS13	1				186											
	P62264	--	Mm.43778	130620	RPS14	RPS14	1		267													
	P62842	--	Mm.288212	180535	RPS15	RPS15	2		269		187											
	P62250		Mm.702	603675	RPS16	RPS16	2			207	188									42		
	P63274	--	Mm.42767	180472	RPS17	RPS17	2		269		188											
	P62270	--	Mm.371578	180473	RPS18	RPS18	2		270		190											
	Q9CZX8	--	Mm.300281	603474	Rps19	RPS19	3		271	208	191								107			
	Q921M2		Mm.21938	603982	Rps20	RPS20	1				192											
	P63221	--	Mm.286669	180477	RPS21	RPS21	1		272													
	P25111		Mm.292027	180465	RPS25	RPS25	1				193											
	P24051		Mm.270283	603702	Rps27l	RPS27	1				194											
	P49664	--	Mm.180003	191343	Rps27a	RPS27L	1		273													
	Q6ZWY3	--	Mm.30120	--		RPS27L	1		274													
	P62858	--	Mm.200920	603685	RPS28	RPS28	1		275													
Cytoskeletal and Cell Adhesion																						
Actin / ARP	Q35247	A0091	Mm.196173	102560	Actin	ACTIN	7	374	278	209	195	91	28		8260822	177				43		
	P53488	A0235	Mm.259045	604221	ACTR2	ACTR2	2	376			196											
	Q99JY9	A0234	Mm.183102	604222	Actr3	ACTR3	2	375			197									44		
	Q9R0Q6	A3287	Mm.34695	604220	Arpc1a	ARPC1A	2	377	277													
	O15144	A0233	Mm.337038	604224	ARPC2	ARPC2	2			210	198					128						
	O15145	A0232	Mm.275942	604225	ARPC3	ARPC3	3		278	211	199					129						
	P59999	A0231	Mm.289306	604226	ARPC4	ARPC4	3			212	200	92				130						
	Q9P1U1	A3339	Mm.150319	--	ARP3beta	ARP3BETA	1				201											
Tubulin	P68362	A0017	Mm.209290	602529	Tuba1	TUBULIN	8	378	279	213	202	93	27	15	8260822	131	8		106			
actinin	Q9JI91	A0007	Mm.37638	102573	Actn2	ACTN	4	379	280		203				9427241	132						
	Q9Z1P2	A2341	Mm.253564	102575	Actn1	ACTN1	7	380	281	214	204	94		18	9427241							
	O88990	A2343	Mm.5316	102574	Actn3	ACTN3	0									133						
	P57780	A2344	Mm.276042	604638	Actn4	ACTN4	3	381	282		205					134						
Spectrin	Q62261	A0010	Mm.123110	182790	Sptbn1	SPECTRIN	8	382	283	215	206	95	26	17	2307671	135						
MAPs	Q9QXZ0	A2342	Mm.3350	608271	Macf1	MACF1	4	383	284	216	207											
	P34926	A0085	Mm.36501	600178	Map1a	MAP1A	5	384	285	217	208	96						36				
	P14873	A0625	Mm.4173	157129	Map1b	MAP1B	4	385	286	218	209							39	109			
	P27816	A0778	Mm.217318	157132	MAP4	MAP4	1		287													
	Q15555	A0654	Mm.132237	605789	MAPRE2	MAPRE2	1		288													
	P10638	A0135	Mm.1287	157140	Mapt	MAPT	3	386	289		210								110			
	P20357	A0134	Mm.256996	157130	Map2	MTAP2	5	387	290	219	211				9427241	136		37				
	O55129	A3376	Mm.154087	601783	Stop	MTAP6	3	388		220	212											
Catenins	Q61301	A1914	Mm.34637	114025	Ctnna2	CATNA	5	389	291	221	213				9427241							
	Q02248	A0112	Mm.291929	116806	Ctnnb1	CATNB	5	390	292	222	214				9427241	137						
	P26231	A0089	Mm.18962	116805	Ctnna1	CTNNA1	1			223												
	P30999	A0219	Mm.35738	601045	Ctnnd1	CTNND1	2	391		224												
	O35927	A0469	Mm.321648	604275	Ctnnd2	CTNND2	4	392	293	225	215											
Myelin	P35802	A3425	Mm.241700	601275	Gpm6a	GPM6A	1	393												45		4
	P04370	A0350	Mm.252063	159430	Mbp	MBP	3		294	226	216					138			111	46		
	Q61885	A0351	Mm.210857	159465	Mog	MOG	4	394	295		217	97				139						3
	Q63912	A3590	Mm.93335	164345	Omg	OMG	2	395			218											
	Q62079	A0352	Mm.1269	300401	Plp1	PLP1	4	396	296		219	98				140	7					3
Cell recognition molecules	P97685	A0663	Mm.326702	--	Nfasc	NFASC	5	397	297	227	220	99										2

Appendix 7

Protein Class	UniProt Acc	PPID Acc	Unigene cluster ID	OMIM	Gene name	Abbreviated name	Confidence (# of times identified)	Postsynaptic density datasets								Complexes			Other structures			# TM helices
								G2C #	Yosh #	Jordan #	Peng #	LI #	Satoh #	Walikonis #	PSD from Literature	MASC #	AMPA #	mGluR5 #	Synaptosome	Clathrin Vesicle	Mitochondria	
	P97547	A3607	Mm.15125	--	Sdfr1	SDFR1	1	398												47		1
	Q62701	A1525	Mm.980	187380	Tnc	TENASCIN-C	1				221											
	Q05546	A1074	Mm.337795	601995	Tnr	TENASCIN-R	3	399	299		222											
Keratins	Q3CWC9	A3741	Mm.158190	--	2410039E07Rik	2410039E07RIK	1	400														
	Q3CV17	A3825	Mm.46376	607742	2310056N18Rik	FLJ22261	1				223											
	Q920F2	A1999	Mm.22657	148040	Krt2-5	KRT5	1	401														
	Q9Z332	A3846	Mm.302399	148041	Krt6a	KRT6A	1	402														
Other Cytoskeletal Proteins	Q3EPW6	A3976	Mm.217161	602330	Ablim1	ABLIM1	4	403	299	236	224					141						
	Q80WK6	--	Mm.254446	--	Ablim2	ABLIM2	1			229												
	P61164	A0660	Mm.3118	605143	ACTR1A	ACTR1A	2	404		230										48		
	Q9R1V4	A3860	Mm.89854	152120	Adam11	ADAM11	1	405														2
	Q9R1V6	A0373	Mm.125477	603709	Adam22	ADAM22	1	406								142						1
	Q9R1V7	A3881	Mm.124892	603710	Adam23	ADAM23	2	407			226											1
	Q9QYC0	A3676	Mm.289106	102680	Add1	ADD1	5	408	300	231	226				9427241			36				
	Q9QYB8	A0587	Mm.104155	102681	Add2	ADD2	4	409	301	232	227							39				
	Q9QY95	A3877	Mm.44106	601568	Add3	ADD3	3	410	302	233												
	Q02357	A0662	Mm.334444	182900	Ank1	ANK1	2			234	228											
	Q01484	A3873	Mm.220242	106410	ANK2	ANK2	4	411	303	235	229											
	Q61307	A3874	Mm.235960	600465	Ank3	ANK3	4	412	304	236	230											
	Q91ZU0	--	Mm.34765	--	Asb7	ASB7	1		305													
	O88737	A0285	Mm.20425	604020	Bsn	BASSOON	6	413	306	237	231	100		18		143		40				
	P47753	--	Mm.19142	601580	Capza1	CAPZ ALPHA-1	2			238	232											
	P47754	A0417	Mm.262039	601571	Capza2	CAPZ ALPHA-2	5	414	307	239	233	101				144						
	P47757	A0418	Mm.2945	601572	Capzb	CAPZ BETA	5	415	308	240	234	102				145						
	O15083	A3879	Mm.318004	--	ERC2	CAST	4	416		241			29		14723704							
	Q68VG1	--	Mm.224189	--		4733477G22Rik	CTTNBP2	2			243	235										
	P46592	A0999	Mm.4024	601442	Cfl1	CFL1	4		309	243	236	103								112		
	Q8R3F2	A3211	Mm.334999	--	Ckap4	CKAP4	1	417														
	Q8C5W0	A3870	Mm.244078	--	Clmn	CLMN	2	419		244												1
	O89053	A2143	Mm.290462	605006	Coro1a	CORO1A	1			245												
	O89046	A2141	Mm.276859	--	Coro1b	CORO1B	1				237											
	Q9VCQ0	A2142	Mm.260158	605269	Coro1c	CORO1C	2		310		238											
	Q9JJQ03	A2139	Mm.335229	605002	CORO2B	CORO2B	1								11799186							
	Q60598	A0083	Mm.205601	164765	Cttn	CORTACTIN	1				239					146						
	Q9QX74	A0081	Mm.323725	603290	Shank2	CORTBP-1	5	418	311	246	240				10414979	147						
	Q8R555	A3975	Mm.329851	606276	Crtac1	CRTAC1	2	420			241											
	Q9JP98	A0466	Mm.22417	601342	CSE1L	CSE1L	0									148						
	Q9QX56	A0436	Mm.19016	126660	Dbn1	DBN1	5	421	312	247	242	104				149						
	P46675	A0535	Mm.6712	125660	Des	DES	1		313													
	Q61495	A0291	Mm.37953	125670	Dsg1a	DESMOGLEIN	1				243					150						
	P11531	A0462	Mm.275608	310200	Dmd	DMD	1								1454867							
	Q61167	A0654	Mm.57874	605789	Magi3	EB2	1	422														
	O15075	A2330	Mm.210035	605070	EEA1	EEA1	1		314													
	Q99MI2	A3878	Mm.288860	607127	Rab6ip2	ELKS	4	423	315	248				19								
	Q9JMB3	A3965	Mm.131135	605331	Epb4.1l3	EPB4.1L3	4	424	316	249	244											
	Q9Z2H5	A3967	Mm.20852	602879	Epb41l1	EPB41L1	2	425		250												
	Q8JZV5	--	Mm.210863	125305	Epb4.9	EPB4.9	1			251												
	P26040	A0326	Mm.277812	123900	Vil2	EZRIN	1						30									
	Q9JJ38	A0204	Mm.39046	300017	Flnc	FILAMIN	0									151						
	Q96ID5	A3979	Hs.212511	--		MGC15730	FLJ90835	2	426				245										
	Q61553	A0224	Mm.289707	602689	Fscn1	FSCN1	1						31									
	P13020	A0419	Mm.21109	137350	Gsn	GELSOLIN	3	427			246		32			152						
	P46660	A0292	Mm.276251	605338	Ina	INTERNEXIN	8	428	317	252	247	106	33	20	9310398	153						
	Q9QYE7	A1987	Rn.34728	--	Itgad	ITGAX	1		318													1
	P70565	A0324	Mm.299774	173325	Jup	JUP	2			253		108										
	Q14008	A4009	Mm.168478	--	KIAA0097	KIAA0097	2	429			248											
	Q9ERD4	A3875	Mm.250641	--	Kidins220	KIDINS220	1	430														4
	Q61792	A3882	Mm.271967	602920	Lasp1	LASP1	3	431		254		107										
	P48679	A1374	Mm.243014	150330	Lmna	LMNA	1			255												
	P14733	A0430	Mm.4105	150340	Lmnb1	LMNB1	0									154						
	Q80052	A3991	Mm.130883	--	Lyric	LYRIC	1	432														1
	P80853	--	Mm.334578	606551	Lzts1	LZTS1	1			256												
	Q8ES.J4	A1027	Mm.192416	606671	Nckipsd	NCKIPSD	2	433		257												
	O35462	A0374	Mm.295283	162230	NF-H	NF-H	3			258	249				1829737							
	P08551	A0375	Mm.1956	162280	Nefl	NF-L	7	434	319	259	250	108		21	1829737	155						
	P08553	A0376	Mm.242832	162250	Nef3	NF-M	8	435	320	260	251	109	34	22	1829737	156						
	O35449	A3989	Mm.373627	--	Ngb	NGB	2	436		261												2
	P63005	A2074	Mm.56337	601545	Pafah1b1	PAFAH1B1	1			262												
	Q02DP4	A3983	Mm.34650	608134	Palm	PALM	3	437	321	263												
	Q96JQ0	--	Mm.334108	603067	OCHS1	PCDH16	1		322													2
	Q91Y06	--	Mm.196700	606332	Pcdhb13	PCDHB13	1		323													2
	Q99569	--	Mm.260938	604276	PKP4	PKP4	2		324		252											
	P30427	A1941	Mm.234912	601282	Plec1	PLECTIN	6	438	325	264	253	110		23								
	O75334		Mm.262890		PPFIA2	PPFIA2	1			265												
	O75145	A3969	Mm.277235	603144	PPFIA3	PPFIA3	2	439		266												
	Q91Z80	--	Mm.295105	603145	Ppfia4	PPFIA4	1			267												
	Q8JUV2	--	Mm.271744	176590	Pfn2	PFN2	1			268												
	Q8VC86	A0216	Mm.36769	602701	Slmap	SLMAP	0									157						1
	O70049	--	Mm.205647	603050	Abi1	SSH3BP1	2	440	326													
	Q62327	A3671	Mm.252321	--	Synpo	SYNPO	4	441	327	269	254							41				
	P37805	--	Mm.24183	607953	Tagln3	TAGLN3	2		328			111								113		
	P70567		Mm.249594	190930	Tmod1	TMOD1	1				256											
	P01831	A1218	Mm.3951	188230	Thy1	THY1	4	442	329		256		35							114	49	
	Q6ERQ9	A0148	Mm.41353	605035	Wasf1	WASF1	4	443	330	270	257											
	Q6NBL0	A3872	Mm.29043	--	DKFZp762B245	WBP3	1	444														
Other Cell Adhesion Molecules	O35889	A0101	Mm.181959	159559	Mlh4	AF6	3		331	271	258											
	Q61282	--	Mm.358571	155760	Agc1	AGC1	1		332									42				
	Q06890	A1276	Mm.200608	185430	Clu	APOJ	1	445														
	P96203	--	Mm.293599	602269	Arvcf	ARVCF	1			272												
	Q61137	A3896	Mm.329566	600904	Astn	ASTN	2	446			259											1
	Q6JHW6	A3897	Mm.22786	--	Astn2	ASTN2	1	447														
	O54691	A0758	Mm.8021	602346	Cntnap1	CASPR	3	448		273	260											1
	Q9Y6N8		Mm.234715	604555	CDH10	CDH10	1				261											1
	Q9WTR5		Mm.334841	601364	Cdh13	CDH13	1				262											
	P23242	A0493	Mm.338720	121014	Gja1	CONNEXIN-43	4	449	333	274	263											4

Appendix 7

Protein Class	UniProt Acc	PPID Acc	Unigene cluster ID	OMIM	Gene name	Abbreviated name	Confidence (# of times identified)	Postsynaptic density datasets							Complexes			Other structures			# TM helices
								G2C #	Yosh #	Jordan #	Peng #	Li #	Satoh #	Walik #	Literature PMID	MASC #	AMPA #	mGluR5 #	Synaptosome	Clathrin Vesicle	Mitochondria	
	P12960	A0757	Mm.4911	600016	Cntn1	CONTACTIN-1	3	460		276	264											
	Q61330	A1073	Mm.260861	190197	Cntn2	CONTACTIN-2	3	461	334					24								
	Q9QUP5	A3893	Mm.266790	115435	Hepbn1	CRTL1	3	462	336		266											
	P55066	A3892	Mm.268079	600826	Cspg3	CSPG3	3	463	338		268											
	Q9QY32	A3895	Mm.38496	606775	Cspg5	CSPG5	1	464														1
	P70587	A0150	Mm.132162	--	Lrrc7	DENSIN-180	7	456	337	279	267	117		25	9524323							
	P15924	A0454	Mm.355327	125647	DSP	DESMOPLAKIN	1				266						8					
	Q912U6	A1575	Mm.336625	--	Dkt	BPAG1	1		339													
	Q9NR18	A0507	Mm.277354	606944	ERBB2IP	ERBB2IP	2		339						11279090							
	P70589	--	Mm.25652	604418	Gjb6	GJB6	1		340													4
	Q9ESM3	--	Mm.294467	--	Hapln2	HAPLN2	1		341													
	Q62718	A3887	Mm.283138	607938	Nit	HNT	2	456			269								115			
	Q920J8	A3888	Mm.317293	--	Negr1	NEGR1	3	457	342		270								116			
	P11627	A0290	Mm.260568	308840	L1cam	L1CAM	3	458	343		271					188	43					1
	Q9JIA1	A0440	Mm.298251	604619	Lgi1	LGI1	3	459		277	272					189						
	Q9Z251	A0141	Mm.268025	603380	Lin7a	LIN7A	3	460					36		10341223	190						
	Q9Z252	A3899	Mm.20472	--	Lin7b	LIN7B	3	461		275					10341223							
	P20917	A3898	Mm.241355	159460	Mag	MAG	1	462														1
	P15116	A0219	Mm.257437	114020	Cdh2	N-CADHERIN	5	455	344	279	273				7723514	191						1
	P13595	A3884	Mm.4974	116930	Ncam1	NCAM1	2	464			274								117			2
	O35136	A3885	Mm.258759	602040	Ncam2	NCAM2	1	465														1
	P97796	A3894	Mm.42249	601907	Neo1	NEO1	1	466														1
	Q62765	A0129	Mm.316080	600568	Nlgn1	NLGN	3	467	345						9927700							1
	Q62888	A3889	Mm.151293	606479	Nlgn2	NLGN2	3	468	346	280												1
	Q62889	A3890	Mm.121508	300336	Nlgn3	NLGN3	3	469	347	281												1
	Q9Z0E0	A1163	Mm.276466	608458	Nrdn	NORBIN	2	470		282												
	Q9ULB1	A0130	Mm.329258	600565	NRXN1	NRXN1	2		348		275											1
	Q07314	A3891	Mm.291005	600567	Nrxn3	NRXN3	3	471	349		276											1
	P32736	A3886	Mm.39534	600632	Opcml	OPCML	2	472			277								118			
	Q9EP53	--	Mm.224054	605284	Tsc1	TSC1	1		350													
Synaptic Vesicles / Protein Transport																						
Clathrin	P08081	A0122	Mm.298875	118960	Clta	CLTA	1	473												60		
	P08082	A3901	Mm.290026	118970	Cltb	CLTB	1	474											119	61		
	P11442	A0119	Mm.254588	118955	Cltc	CLTC	4	476	351	283	278					192		44	120	62		
	Q08838	A0121	Mm.101650	600418	Amph	AMPH	2	478		284									121			
	O08839	A1109	Mm.4383	601248	Bin1	AMPHIPHYSIN2	2	477		285									122			
	P10107	A0647	Mm.248360	151690	Anxa1	ANXA1	1		352													
	Q61548	A0684	Mm.281651	607923	Snap91	AP180	1	479											123			
	P52303	A0686	Mm.274816	600157	Ap1b1	AP1B1	3	479	353		279								124			
	P17426	A0251	Mm.6877	601026	Ap2a1	AP2A1	4	480	354	286	280									63		
	P17427	A0252	Mm.253090	607242	Ap2a2	AP2A2	3	481		287	281								126	64		
	Q9DBG3	A0253	Mm.39053	601025	Ap2b1	AP2B1	2	482		288								45		65		
	P20172	A0254	Mm.18946	601024	AP2M1	AP2M1	3	483		289	282									66		
	Q00380	A0255	Mm.333597	602242	Ap2s1	AP2S1	2		355	290										67		
	Q9JME5	--	Mm.322894	602166	Ap3b2	AP3B2	1			291												
	O54774	A1324	Mm.209294	607246	Ap3d1	AP3D1	3	484		292	283											
	P35678	A3907	Mm.10647	--	Ap3m2	AP3M2	1	485														
	P35564	A0452	Mm.248827	114217	Canx	CALNEXIN	1	486														1
	P53621	A2102	Mm.30041	601924	COPA	COPA	1	487														
	O55029	A0237	Mm.261735	606990	Copb2	COPB2	1	488														
	Q9ERI6	A2349	Mm.255138	603432	Cyln2	CYLN2	2	489		293												
	Q01469	A0441	Mm.741	605168	FABP5	E-FABP	1	490								193						
	Q9QXY6	A1331	Mm.18526	605891	Ehd3	EHD3	1	491														
	Q9Z1Z3	A3926	Mm.139695	607263	Epn2	EPN2	1	492														
	O54922	A3921	Mm.22530	608163	Exoc7	EXO70	1	493														
	O54924	A3923	Mm.347360	--	Exo84	EXO84	1	494														
	O55126	A3913	Mm.12468	603004	Gbas	GBAS	3	495		294		113									70	
	O88630	A3936	Mm.20931	604026	Gosr1	GOSR1	1	496														1
	Q8JKY5	A3906	Mm.149954	605613	Hip1r	HIP1R	1	497												68		
	P97411	A2211	Mm.275683	147625	Ica1	ICA1	1			295												
	Q61595	A3531	Mm.296457	600381	Ktn1	KTN1	2	498			284											1
	Q69994	A4008	Mm.177724	--	Mgi87	MGI87	1	499													71	1
	Q9Y3A3	A3935	Mm.291037	--	PRE3	MOB3	1	500														
	P54921	A0596	Mm.104540	603215	Napa	NAPA	1	501											128	69		
	P28663	A3910	Mm.274308	--	Napb	NAPB	2	502					37							60		
	Q9QWZ7	A3911	Mm.292687	603216	Napg	NAPG	1	503														
	Q55126	A3914	Mm.293716	603249	Ncaprap1	NIPSNAP1	4	504	356	296			38								72	
	Q62443	A1697	Mm.5142	602367	Nptx1	NPTX1	1						39						127			
	P46460	A0117	Mm.260117	601633	Nsf	NSF	6	505	357	297	285	114			1829737	194		46		61		
	Q9ZQW5	A0501	Mm.4926	606512	Pacsin1	PACSIN1	2		358			115							129			
	O55011	A3927	Mm.235175	603025	Picalm	PICALM	2	506	359													
	Q9JM89	A3932	Mm.151332	--	Pex5	PXR2	1	507														
	Q80U40	--	Mm.233996	--	Rimbp2	RIMBP2	1			298												
	Q99NE5	A0035	Mm.328661	606629	Rims1	RIM1	4	508	360	298	286								129			
	Q9EQZ7	A1861	Mm.309296	606630	Rims2	RIMS2	2		361	300									130			
	Q8R361	--	Mm.220334	605536	Rsp11	RIP11	1			301												
	P47708	A0212	Mm.181166	--	Rph3a	RPH3A	6	509	362	302	287	116			7922236			47				
	Q62991	A3928	Mm.216511	--	Scfd1	SCFD1	1	510														
	P97878	A3925	Mm.31607	604469	Sec10l1	SEC10L1	1	511														
	Q9D4R7	A3924	Mm.159621	607880	Sec15l2	SEC15L2	1	512														
	Q9D4H1	A3920	Mm.293510	--	Sec5l1	SEC5L1	1	513														
	Q62825	A3922	Mm.261859	608186	Sec8l1	SEC8L1	1	514														
	Q8VIL3	--	Rn.7624	--	Zwint	SIP30	1					117										
	P60881	A0333	Mm.45953	600322	SNAP25	SNAP25	4	516	363	303	288					195			131	62		
	Q9QXY3	A3903	Mm.342665	--	P140	SNIP	4	518	364	304	289							48				
	O60493	--	Mm.277870	605930	SNX3	SNX3	1					118										
	Q64548	A0445	Mm.221275	600865	Rtn1	S-REX	2	517		305						196			132			2
	Q9QXG3	A0040	Mm.6225	186590	Stx1a	STX	3	519	365		290					197			133			1
	P61264	--	Mm.40343	--	Stx1b2	STX1B2	1			306									134			1
	O70439	A3940	Mm.248042	603217	Stx7	STX7	1	519											135			1
	Q99JB2	A3942	Mm.26751	--	Stom2	SYTL2	2	520					40									
	Q9ER00	A3939	Mm.28237	606892	Stx12	STX12	1	521											136	63		1
	Q08599	A0238	Mm.278865	602926	Stxbp1	STX0BP1	4	522		307	291		41			198	9			64	73	
	Q02563	A3902	Mm.200365	185860	Sv2	SV2	2	523							1519084				137			12
	Q15431	A3863	Mm.243649	602162	SYCP1	SYCP1	1	524														

Appendix 7

Protein Class	UniProt Acc	PPID Acc	Unigene cluster ID	OMIM	Gene name	Abbreviated name	Confidence (# of times identified)	Postsynaptic density datasets								Complexes			Other structures			# TM helices
								G2C #	Yosh #	Jordan #	Peng #	Li #	Satoh #	Walik #	PSD from Literature	MASC #	AMPA #	mGluR5 #	Synaptosome	Clathrin Vesicle	Mitochondria	
	Q88935	A0181	Mm.370211	313440	Syn1	SYN1	6	626	366	309	292	119	42	28	1429737				138			
	Q9QWV7	A3904	Mm.275845	600755	Syn2	SYN2	8	626	367	309	293	120	43	27	8034989				138			
	Q9QZA3	--	Mm.374805	602705	Syn3	SYN3	1			310												
	O55100	A0288	Mm.230301	603925	Syngr1	SYNAPTOGYRIN	1	627								169		49				4
	Q9WVG8	A3933	Mm.26032	603927	Syngr3	SYNGR3	2	628							9780194					96		4
	Q62910	A0120	Mm.187079	604297	Syn1	SYNJ1	2	629			294							50	140			
	Q62277	A3941	Mm.223674	313475	Syp	SYP	2	630							8210620				141	98		3
	P49096	A0039	Mm.289702	185605	Syt1	SYT1	5	631	368	311	295		44			170			142	97		1
	Q8R0N7	A3905	Mm.182654	604146	Syt7	SYT7	2	632			296								143			1
	Q9DCD6	A4027	Mm.284594	--	Tjp4	TJP4	1	633														
	Q9Z152	A0562	Mm.286868	604566	Stxbp6	TOMOSYN	1	634														
	Q91ZY8	--	Mm.184012	--	Trim9	TRIM9	2		369	312												
	Q8IYF7	--	Mm.17436	605073	TRIM37	TRIM37	1		370													
	Q62769	A0560	Mm.334606	--	Unc13a	UNC13A	4	635	371	313	297											
	Q8K0T7	--	Mm.41035	--	Unc13c	UNC13C	1			314												
	P63044	--	Mm.28643	185881	Vamp2	VAMP2	1			315								51		88		1
	Q9Z270	A3937	Mm.266767	605703	Vapa	VAPA	3	636		316	298									89		1
	Q9QY76	A3936	Mm.260456	605704	Vapb	VAPB	1	637														1
	Q01853	A3918	Mm.262053	601023	Vcp	VCP	1	638											144			
	Q9ZQQ4	A3908	Mm.276092	608550	Vps16	VPS16	2	639		317												
	Q8BMZ7	--	Mm.30028	--	Vps28	VPS28	1			318												
	Q9QZ88	--	Mm.216528	606932	Vps29	VPS29	1			319												
	Q63615	A3930	Mm.41372	--	Vps33a	VPS33A	1	640														
	Q61123	A3909	Mm.296520	606931	Vps35	VPS35	2	641		320										70		
	Q9H3A6	A1326	Mm.27389	605465	hVPS41	VPS41	1	642														
	P97390	A3929	Mm.263185	--	Vps45a	VPS45A	1	643														
	O88384	A3934	Mm.265929	603207	Vti1b	VTI1L1	1	644														1
Transporters	P06795	--	Mm.146649	--	Abcb1	ABCB1B	1		372										145			10
	Q9CXJ4	A3947	Mm.283914	605464	4833412N02Rik	ABCB8	1	645													74	4
	P55096	A3946	Mm.194462	170995	Abcd3	ABCD3	2	646		321												3
	Q9QZP4	A3948	Mm.285322	--	Mtch1	MTCH1	1	647													75	2
	Q99JZ8	A3949	Mm.28023	--	Mtch2	MTCH2	3	648	373	322											76	1
	P47802	A3956	Mm.280943	600605	Mtx1	MTX1	1	649													77	1
	O88441	A3957	Mm.292613	608555	Mtx2	MTX2	1	650													78	
	O08547	--	Mm.2551	604229	Sec22l1	SEC22L1	1			323												1
	Q01405	A3945	Mm.209207	--	Sec23a	SEC23A	2	651	374													
	Q63965	A3951	Mm.134191	--	Sfxn1	SFXN1	1	652											146		79	2
	Q9JHY2	A3952	Mm.36169	--	Sfxn3	SFXN3	3	653	375	324											80	2
	Q925N0	A3954	Mm.121485	--	Sfxn5	SFXN5	2	654		325											81	5
	P32089	A3750	Mm.229291	190315	Slc25a1	SLC25A1	1	655													82	
	P16035	A3776	Mm.298	600370	Slc25a3	SLC25A3	2	656	376										147		83	2
	P48962	A0395	Mm.16228	103220	Slc25a4	SLC25A4	5	657	377		299	121		28		171			148		84	3
	Q09073	A0396	Mm.658	300150	Slc25a5	SLC25A5	2	658			300					172			149		85	2
	O75746	A0443	Mm.30928	603667	SLC25A12	SLC25A12	3	659		326	301					173					86	
	Q9QXX4	A3747	Mm.24513	603859	Slc25a13	SLC25A13	1	660													87	
	Q9H936	A3773	Mm.33729	--	SLC25A22	SLC25A22	3	661	378		302					174					88	1
	P97441	A3943	Mm.1396	602878	Slc30a3	SLC30A3	1	662														6
	Q9CZW5	A3950	Mm.213292	606081	Tomm70a	TOMM70A	1	663													89	
	P40336	A3944	Mm.260703	605506	Vps26	VPS26	1	664														
Motor Proteins	Q9R1R2	A1178	Mm.29290	605493	Trim3	BERP	2	665		327												
	Q99JI3	A3970	Mm.222272	605853	Clasp2	CLASP2	4	666	379	328	303											
	Q99PJ8	A0656	Mm.167537	607376	Dctn3	DCTN2	1	667														
	Q9QUR2	A3969	Mm.272801	--	Dctn4	DCTN4	1	668														
	Q9D0M5	--	Mm.246436	608942	Dlc2	DLC2	1			329												
	P38650	A0133	Mm.181430	600112	Dnchc1	DNCH1	3	669	380		304											
	P63168	A0514	Mm.256858	601562	DNCL1	DNCLC1	3		381	330	305											
	Q8R1Q8	--	Mm.128627	--	Dnclic1	DNCLIC1	1			331												
	P70611	A0561	Mm.266301	604567	Doc2a	DOC2A	2		382		306											
	O88485	A0661	Mm.20893	603772	Dnci1	DNCI1	2	670		332												
	O88487	A2127	Mm.249479	603331	Dnci2	DNCI2	1	671														
	P60521	--	Mm.371666	607452	Gabarapl2	GABARAPL2	2		383	333												
	Q60760	--	Mm.40338	603127	Gas7	GAS7	1		384													
	Q8Q9F6	A3973	Mm.235123	600378	Immt	IMMT	4	672	385	334	307										90	
	Q670R4	A0571	Mm.40546	602442	Itsn1	INTERSECTIN-1	1	673														
	P70188	--	Mm.4651	601836	Kifap3	KIFAP3	1			335												
	P28740	A3966	Mm.355686	602591	Kif2	KIF2	3	674	386	336												
	Q61768	A3967	Mm.22744	602809	Kifb	KIF5B	2	675		337												
	P28736	--	Mm.256342	604593	Kif5c	KIF5C	1			338												
	O88448	A3968	Mm.271648	--	Klc2	KLC2	1	676								175						
	Q63781	A4048	Mm.306770	--	Mrlcb	MRLC2	2		387	339						176						
	P09541	--	Mm.247636	160770	Myl4	MYL4	1			340												
	Q60925	A4049	Mm.337074	--	Myl6	MYL6	2		388		308					177						
	P16475	A4049	Mm.337074	--	MYL6	MYL6	0															
	O61355	A3202	Mm.290003	160760	Mhc	MYH7	0									178						
	Q5JMH9	A3964	Mm.341248	--	Myo18a	MYO18A	4	677	389	341	309											
	QRVDD5	A3186	Mm.29677	160775	Myh9	MYH9	4	678	390	342	310					179						
	Q9JLT0	A3961	Mm.218233	160776	Myh10	MYH10	3	679		343	311					180			150			
	O08638	A2331	Mm.250705	160745	Myh11	MYH11	3	680	391	344						181						
	Q9H8B2	A3962	Mm.358743	608568	Myh14	MYH14	2	681	392													
	P46735	A3958	Mm.3390	606537	Myo1b	MYO1B	2	682		345						182						
	Q63357	A3959	Mm.151948	606539	Myo1d	MYO1D	1	683														
	Q9NQX4	--	Mm.41339	--	MYO5C	MYO5C	1				312											
	O54331	A3960	Mm.4040	600970	Myo6	MYO6	2	684	393													
	Q99MZ6	--	Mm.312342	606541	Myo7b	MYO7B	1				313											
	Q9QYF3	A0149	Mm.3645	160777	Myo5a	MYOSIN (V)	7	685	394	346	314	122		29	12711090	183						
	O60313	A3971	Mm.274285	605290	OPA1	OPA1	1	686													91	
	O08788	A0658	Mm.6919	601143	Dctn1	P150GLUED	3	687		347	315											
	P21807	A0459	Mm.2477	170210	Psph	PRKH1	2		395	348												
	P30622	A2345	Mm.241109	179838	RSN	RESTIN	1	688														
	P46097	--	Mm.5102	600104	Syt2	SYT2	1		396										151	71		1
	O35681	--	Mm.4624	600327	Syt3	SYT3	1		397													1
	Q8R0N5	--	Mm.354663	--	Syt5	SYT5	1		398										152			1
	P70596	A3972	Mm.291880	602926	Tmod2	TMOD2	5	689	399	349		123	45						153			
	Q63608	A3965	Mm.121878	191010	Tpm1	TPM1	4	690	400		316	124										
	P58774	--	Mm.646	190990	Tpm2	TPM2	1		401										154			

Appendix 7

Protein Class	UniProt Acc	PPID Acc	Unigene cluster ID	OMIM	Gene name	Abbreviated name	Confidence (# of times identified)	Postsynaptic density datasets								Complexes			Other structures			# TM helices
								G2C #	Yosh #	Jordan #	Peng #	LI #	Satoh #	Walik #	PSD from Literature	MASC #	AMPA #	mGluR5 #	Synaptosome	Clathrin Vesicle	Mitochondria	
	P21107	A3205	Mm.240839	191030	Tpm3	TPM3	3	691	402		317											
	P31000	A0850	Mm.268000	193060	Vim	VIM	2		403			126										
Uncharacterised / novel																						
	Q9DCZ4	--	Mm.371632	--	0610008C08Rik	0610008C08Rik	1			360												1
	Q8VDQ5	--	Mm.294159	--	0610011D08Rik	0610011D08RIK	1					126										
	Q9D595	A3994	Mm.20313	--	0710001P09Rik	0710001P09RIK	2	692		361											92	
	Q8R570	--	Mm.26680	--	1110031B06Rik	1110031B06Rik	1			362												
	Q78IS1	--	Mm.27606	--	1200002G13Rik	1200002G13RIK	1		404													
	Q8BJY8	--	Mm.120298	--	Nebl	1200007O21RIK	2		405			127										
	Q8BMB1	--	Mm.41903	--	1200014J11Rik	1200014J11RIK	1		406													
	Q9JKC6	--	Mm.87027	--	Bin88	1500001H12RIK	1		407													1
	Q9DB72	--	Mm.38347	--	1500005I02Rik	1500005I02RIK	1				318											
	Q9CW19	--	Mm.200385	--	1500041B16Rik	1500041B16RIK	1					128										
	Q8K2T2	A3996	Mm.157746	--	1700054N08Rik	1700054N08RIK	1	693														
	Q8WVI5	--	Mm.318691	--	RBAF600	1810009A16RIK	1					129										
	Q9D935	--	Mm.71682	--	1810009J06Rik	1810009J06RIK	1			363												
	Q80TP4	--	Mm.285968	--	Wdr47	1810073M12Rik	1			364												
	Q6P099	--	Mm.96211	--	Sri	2210417O06RIK	1		408													
	Q8C9Z0	--	Mm.337339	--	2310035C23Rik	2310035C23Rik	1			365												
	Q8K1A7	--	Mm.24829	--	Fbxo31	2310046N15RIK	1		409													
	Q922F4	--	Mm.181860	--	Tubb6	2310057H16RIK	1		410													
	Q69NN1	--	Mm.41459	--	2610039E05Rik	2610039E05Rik	1			366												
	Q9D023	--	Mm.195625	--	2610205H19Rik	2610205H19Rik	1			367												
	Q9CZU2	A3995	Mm.61582	--	2610528A17Rik	2610528A17RIK	1	694														
	Q9Z2D2	A3510	Mm.102970	--	Mtmr13	4833411B01RIK	1	695														
	Q9CVR0	--	Mm.227260	602660	4930542G03Rik	4930542G03Rik	1					130										
	Q8C0R4	--	Mm.72926	--	4932414J0	4932414J0	1					131										2
	Q8BUY8	--	Mm.271980	--	5330440H13Rik	5330440H13Rik	1			368												
	Q8BP77	--	Mm.350327	--	5730453H04Rik	5730453H04Rik	1			369												
	Q9CYH2	--	Mm.27227	--	5730469M10Rik	5730469M10Rik	1			360												1
	Q80T83	--	Mm.274493	--	5730538E15Rik	5730538E15Rik	1			361												
	Q8BX25	--	Mm.44490	--	DCAMKL2	6330415M09Rik	1			362												
	Q8BQV1	--	Mm.44355	--	6430514L14Rik	6430514L14Rik	1			363												
	Q80Y24	--	Mm.248820	608501	Prickle2	6720451F09Rik	1			364												
	Q9D217	A4003	Mm.257354	--	A330102H22Rik	A330102H22RIK	1	696														
	Q8C409	--	Mm.204836	--	A930041I02Rik	A930041I02Rik	1			365												
	Q8C0J2	A2084	Mm.272972	--	Apg16l	APG16L	0									184						
	Q8VDA6	--	Mm.30060	--	Avpi1	AVPI1	1					132										
	Q9CUI5	--	Mm.92705	--	Azi2	AZI2	1			366												
	Q9D5J9	A4015	Mm.199698	--	4930431B11Rik	BA342L8.1										186						
	Q8K4U9	--	Mm.240850	602968	Bcas1	BCAS1	1					133										
	Q9HAA5	--	Hs.486708	--	C8ORF103	C8ORF103	1		411													
	Q80W16	--	Mm.4880	--	Cbln1	CBLN1	1			367												
	Q8VCR7	--	Mm.335427	--	Cib	CIB	1			368												
	Q9D172	--	Mm.268691	601659	D10Jhu81e	D10ERTD214E	2		412	369												
	Q8C3Q5	--	Mm.347599	--	D430041B17	D430041B17	1			370												
	Q8BXF0	--	Mm.291609	--	D430041D05Rik	D430041D05Rik	1			371												1
	Q8C3K4	--	Mm.35077	--	D630045J12Rik	D630045J12Rik	1			372												
	Q8C8Q0	--	Mm.196943	--	D6Ertd349e	D6ERTD349E	1			373												
	Q9CU45	A3800	Mm.248687	--	6330407J23Rik	DJ403A15.3	1	697														1
	Q9H0C2	A3744	Mm.78691	--	DKFZp434N123	DKFZP434N1235	1		413												93	2
	Q8K2R1	A0281	Mm.23044	--	9630050M13Rik	DKFZP564A026	0									188						
	Q9NXY9	--	Hs.410889	--	DKFZP564D166	DKFZP564D166	1				319											
	Q8VD37	A3987	Mm.238094	--	3110007P09Rik	DKFZP761D221	1	698														
	Q8R1A4	A3966	Mm.260623	--	Dock7	DOCK7	1	699														
	Q9CR89	A3568	Mm.21501	--	Chchd3	FLJ20420	1	700														
	Q12914	A4002	Hs.502266	--	G2	G2	1	701														1
	O35127	--	Mm.3172	--	C10	GRCC10	2		414	374												
	Q9Y3I0	--	Mm.9257	--	HSPC117	HSPC117	1		415													
	Q96TC7	A3511	Mm.330492	--	hucep-10	HUCEP-10	1	702														1
	Q9HCG3	A2325	Mm.37803	606793	HOMOGT-1	HOMOGT1	1	703														
	Q96KE9	--	--	--	Hypothetical prot	hypothetical protein	1					134										
	Q8R0L7	--	Mm.271717	606644	Igsf8	IGSF8	1			375												1
	Q14700	A4000	Mm.271956	--	KIAA0090	KIAA0090	1	704														
	Q9CRD2	A3977	Mm.244512	607722	Kiaa0103	KIAA0103	2	705			320											
	Q922I2	A4006	Mm.260647	--	C920006C10Rik	KIAA0143	1	706														
	Q92508	--	Mm.37324	--	FAM38A	KIAA0233	1		416													23
	Q80U49	--	Mm.23689	--	AW555464	KIAA0284	1			376												
	O60302	A3912	Mm.134602	--	KIAA0555	KIAA0555	1	707														
	Q8C090	--	Mm.134338	--	AU040829	KIAA0672	1			377												
	Q9Y2G0	A4035	Mm.210266	--	KIAA0953	KIAA0953	1	708														
	Q80TJ8	--	Mm.299442	--	AK122446	KIAA1110	1			378												
	Q9UL73	A3992	Mm.186647	--	GPR158	KIAA1136	1	709														1
	Q9ULQ0	A3983	Mm.56097	--	FAM40B	KIAA1170	1	710														
	Q9ULK3	--	Hs.445585	--	KIAA1217	KIAA1217	2			379	321											
	Q96IB4	--	Hs.505516	--	KIAA1463	KIAA1463	1				322											
	Q9HCM3	A3554	Hs.490294	--	KIAA1549	KIAA1549	1	711														2
	Q9C0D0	--	Mm.180124	--	PHACTR1	KIAA1733	2			380	323											
	Q8TDC3	--	Hs.443752	--	KIAA1811	KIAA1811	1			381												
	Q96PY1	--	Mm.296626	--	KIAA1906	KIAA1906	1			382												1
	Q9JJK2	A3988	Mm.274904	--	Lancl2	LANCL2	1	712														
	Q92210	A4004	Mm.260538	604407	Letm1	LETM1	1	713													94	1
	Q8NE86	A4023	Mm.59732	--	C10orf42	LOC90550	1	714														3
	Q9K4V0	A3999	Mm.217027	--	Lrpprc	LRPPRC	1	715														
	Q9HAY2	--	Hs.306123	--	MAGEF1	MAGEF1	1		417													
	Q6TB98	--	Hs.130316	--	PRR7	MGC10772	1		418													1
	Q9D898	A3290	Mm.38155	--	Aspc9l	MGC3038	1				324											
	O35906	A4045	Mm.196596	--	C18orf8	MIC1	1	716														
	Q8IUF7	--	Mm.289150	--	mirin53	MINA	1		419													
	Q99KX1	--	Mm.29737	601401	Mif2	MLF2	1					136										
	P56379	--	Mm.318	--	Myl6	MP68	1			383												1
	P61956	--	Mm.29923	603042	SMT3H2	Smt3h2	1					136										
	Q8N8M3	--	Hs.5025	605491	NEBL	NEBL	1					137										
	O95897	--	Hs.169743	--	OLFM2	OLFM2	1		420													

Appendix 7

Protein Class	UniProt Acc	PPID Acc	Unigene cluster ID	OMIM	Gene name	Abbreviated name	Confidence (# of times identified)	Postsynaptic density datasets								Complexes			Other structures			# TM helices
								G2C #	Yosh #	Jordan #			Satoh #	Walik #	PSD from Literature	MASC #	AMPA #	mGluR5 #	Synaptosome	Clathrin Vesicle	Mitochondria	
	Q9CRA1	--	Mm.275648	605903	Pdlim7	PDLIM7	1		421													
	Q7TQG1	--	Mm.253559	607771	Plekha6	PEPP3	1			384												
	P14200	--	Mm.179378	176795	Pmch	PMCH	1		422													
	Q9NTK5	--	Mm.22681	--	PTD004	PTD004	1		423													
	Q8JNJ4	--	Mm.261333	--	SIPA1L1	SIPA1L1	3		424	385	326											
	P52823	--	Mm.20911	601185	STC1	STC1	1		425													
	Q9CUN3	A3993	Mm.260782	--	2810037C14Rik	SVH	1	617														1
	Q63128	--	Mm.157716	--	CPG2	SYNE1	2		426	386												
	O54784	--	Mm.14485	--	Tex9	TEX9	1		427													
	Q98Q82	--	Hs.12082	--	tiga1	TIGA1	1		428													1
	Q9CR67	--	Mm.23217	--	Tmem33	TMEM33	1			387												
	Q16890	--	Mm.7821	604069	TPD52L1	TPD52L1	1					138										
	P14701	--	Mm.296922	600763	Tpt1	TPT1	1		429										155			
	Q86TV6	--	Mm.295441	--	TTC7B	TTC7B	2	618		388												
	Q9EQN7	A3997	Mm.29729	605784	Tthh1	TTYH1	1	619														4
	Q9ER48	--	Mm.29722	--	Usmg5	USMG5	1			389												1
	Q80ZI9	--	Mm.2654	604734	Wdr1	WDR1	1			390												
	Q8CBE3	--	Mm.284654	--	Wdr37	WDR37	1			391												
	Q92049	A4010	Mm.30850	--	Wdr7	WDR7	3	620		392	326											
Removed from databases																						
	--	--	--	--	gi:892272	gi:892272	2		430				46									
	--	--	--	--	gi:27733035	gi:27733035	1			393												
	--	--	--	--	gi:27712644	gi:27712644	1			394												
	--	--	--	--	gi:27666466	gi:27666466	1					139										
	--	--	--	--	gi:18601785	gi:18601785	2		431			140										
	--	--	--	--	gi:27683621	gi:27683621	1					141										
	--	--	--	--	gi:20909043	gi:20909043	1					142										
	--	--	--	--	gi:27684483	gi:27684483	1					143										
	--	--	--	--	gi:27718001	gi:27718001	1			395												
	--	--	--	--	gi:20864488	gi:20864488	1					144										
	--	--	--	--	gi:27682685	gi:27682685	1			396												1
	--	--	--	--	gi:20881893	gi:20881893	1					145										
	--	--	--	--	gi:20843007	gi:20843007	1			397												1
	--	--	--	--	gi:27689767	gi:27689767	1			398												
	--	--	--	--	gi:27677960	gi:27677960	1			399												1
	--	--	--	--	gi:20876984	gi:20876984	1					146										
	--	--	--	--	gi:20884391	gi:20884391	1					147										
	--	--	--	--	gi:27731499	gi:27731499	1					148										
	--	--	--	--	gi:27666085	gi:27666085	1			400												
Other cell origin																						
	P08758	--	Mm.1620	131230	ANXA5	ANXA5	2		432			149							156	72		
	Q60771	A1892	Mm.4425	--	Cldn11	CLDN11	1		433													4
	P23943	--	Mm.284855	--	Edn2	EDN2	1		434													
	P58107	--	Mm.259929	--	EPPK1	EPPK1	1		435													
	P14480	A1493	Mm.30063	--	Fgb	FGB	1					150										
	P02680	A1494	Mm.16422	--	Fgg	FGG	1		436													
	P47819	A0534	Mm.1239	137780	Gfap	GFAP	4		437	401	327	151										
	Q9ERL8	--	Mm.87312	--	Gmfb	GMFB	1		438										157			
	P01946	A1741	Mm.196110	--	Hba1	HBA1	1		439										158			
	P02091	A1742	Mm.286567	--	Hbb	HBB	1		440											73		
	P19637	A1496	Mm.154660	173370	Plat	PLAT	1		441													
mGLUR5 complex proteins not yet found in the PSD																						
	GI	Name																				
	130228	Phospholipase C-delta					0											62				
	19705535	Phosphofurin acidic cluster sorting protein 1a (PACS1)					0											63				
	3347955	Phosphofurin acidic cluster sorting protein 1b (PACS1b)					0											64				
	7441446	Destrin					0											65				
	8393564	Histamine receptor H1					0											66				
	266495	Myristoylated alanine-rich C-kinase substrate (MARCKS)					0											67				
	9507115	Solute carrier family 1, member 3					0											68				
	111839	Inositol 1,4,5-triphosphate receptor 2 (IP3R2)					0											69				
	9437326	NBC (Na/HCO3)-like protein					0											60				
	16758542	Nuclear factor regulated by interleukin 3					0											61				
	18677712	Periaxin					0											62				
	16758594	Phospholipase C-epsilon					0											63				
	16758716	Voltage-dependent calcium channel, L-type beta					0											64				
	21687094	Voltage-activated calcium channel, T-type alpha 1					0											65				

Appendix 7 - Integration of PSP datasets. Our data from MS and immunoblotting analyses of the PSD, MASC and the AMPA receptor complex were integrated with data from 6 other MS analyses of PSD fractions ("Yosh", (Yoshimura et al., 2004); "Jordan", (Jordan et al., 2004); "Peng", (Peng et al., 2004); "Li", (Li et al., 2004); "Satoh", (Satoh et al., 2002); "Walik", (Walikonis et al., 2000), PSD proteins from literature searching (PMID numbers for the papers are included); MS analysis of the mGluR5 receptor complex, (Farr et al., 2004); MS analysis of rat cortical synaptosomes, (Witzmann et al., 2005); rat brain clathrin-coated vesicles (Blondeau et al., 2004) and mouse brain mitochondria (Mootha et al., 2003). Each dataset has a running number, UniProt, PPID, Unigene cluster and OMIM accession numbers for each entry if available. The "Confidence" score is the number of times a given protein was found in the PSD, proteins found more than once are considered to be of higher confidence and constitute the Consensus PSD. Transmembrane helices were predicted using TMHMM Server v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM/).

Appendix 8

						Times detected in PSDs								Complexes		
						8	7	6	5	4	3	2	1			
Protein Class	**UniProt Acc**	**PPID Acc**	**Unigene cluster**	**OMIM**	**Gene name**	Consensus PSD								**MA**	**mG**	**AM**
Channels and Receptors																
Glutamate Receptors	Q62683	A0001	Mm.278672	138249	Grin1		X							X		
	O08948	A0002	Mm.2953	138253	Grin2a		X							X	X	
	Q62684	A0003	Mm.322010	138252	Grin2b		X							X		
	P23819	A0297	Mm.220224	138247	Gria2			X								X
	P23818	A0296	Mm.4920	138248	Gria1				X							X
	Q9Z2W9	A0299	Mm.327681	305915	Gria3				X							X
	P19493	A0300	Mm.209263	138246	Gria4						X					X
	Q9EPV6	A0103	Mm.370186	604473	Grm1						X			X		
	Q9QYS2	A1353	Mm.318966	601115	Grm3						X					
	P39087	A0022	Mm.332838	138244	Grik2							X		X		
	P31424	A0327	Mm.235018	604102	Grm5							X		X		
	P35400	A0453	Mm.240881	604101	Grm7							X				
	Q61625	A0328	Mm.321227	602368	Grid2								X			
	Q03391	A0005	Mm.152584	602717	Grin2d								X			
Other ligand receptors	Q9WV18	A0603	Mm.32191	603540	Gabbr1				X							
	O88917	A0038	Mm.260733	--	Lphn1					X						
	P18504	A0503	Mm.6834	602376	Gabra1						X					
	O88871	A2000	Mm.101909	607340	Gpr51						X					
	Q9Z173	A3463	Mm.273631	--	Lphn3							X				
	O55022	A3464	Mm.9052	300435	Pgrmc1								X			
Ca2+-ATPases	O55143	A3467	Mm.227583	108740	Atp2a2								X			
	P11505	A3465	Mm.166944	108731	Atp2b1								X		X	
	Q9R0K7	A0508	Mm.321755	108733	Atp2b2								X			
	Q64568	A3466	Mm.210095	300014	Atp2b3								X			
	Q64542	A0114	Mm.188617	108732	Atp2b4								X	X		
NA+/K+-ATPases	Q9Z1G6	A0342	Mm.44101	182350	Atp1a3						X			X	X	
	P06685	A0340	Mm.280103	182310	Atp1a1							X		X	X	X
	P06686	A0341	Mm.207432	182340	Atp1a2								X		X	
	Q9WV28	A0343	Mm.337950	607321	Atp1a4								X			
	P14094	A0344	Mm.4550	182330	Atp1b1								X		X	
Voltage- and Ligand-gated Ca2+-channel	Q61290	A3471	Mm.298091	601013	Cacna1e					X						
	P54285	A3476	Mm.3544	601958	Cacnb3						X					
	P54282	A0593	Mm.334658	601011	Cacna1a							X			X	
	Q01815	A0770	Mm.236039	114205	Cacna1c							X				
	O08532	A3474	Mm.173392	114204	Cacna2d1							X				
	Q7TPF2	--	Mm.41252	114207	Cacnb1							X				
	O55017	A0592	Mm.4424	601012	Cacna1b								X			
	Q9EQG2	A3472	Mm.273084	607082	Cacna2d2								X			
	Q9Z1L5	A3473	Mm.370172	606399	Cacna2d3								X			
	Q9JJV5	--	Mm.241121	606403	Cacng3								X			
Voltage-gated K+-channel	Q8VHF1	--	Mm.343607	600150	Kcnma1							X				
	P63142	A0764	Mm.56930	176262	Kcna2							X				
	P52190	A1867	Mm.140760	600504	Kcnj4								X			
	Q9Z343	A3477	Mm.40615	602235	Kcnq2								X			
	Q64284	A3478	Mm.302496	601142	Kcnab2								X			
Voltage-dependent anion channels	Q60932	A0425	Mm.3555	604492	Vdac1				X					X		
	Q60930	A0426	Mm.262327	193245	Vdac2					X				X		
	Q60931	A3479	Mm.227704	--	Vdac3					X						
Na+-channel	P15390	--	Mm.218743	603967	Scn4a							X				
	Q99250	A0598	Hs.93485	601219	SCN2A								X			
Ca2+-release channels	P11881	A0154	Mm.227912	147265	Itpr1							X			X	
	Q9ERN6	A3480	Mm.239871	180902	Ryr2								X			
Other Channels and Receptors	Q03555	A0637	Mm.341742	603930	Gphn					X						
	P12023	A1169	Mm.277585	104760	App							X				
	Q9Z1G4	A0414	Mm.340818	192130	Atp6v0a1							X				
	Q9UI12	A0409	Mm.27082	608861	ATP6V1H							X				
	O14514	A1017	Mm.43133	602682	BAI1							X				
	O60242	--	Mm.336569	602684	BAI3							X				
	P43006	A0422	Mm.267547	600300	Slc1a2							X		X		
	Q08406	A3490	Mm.272210	118946	Cntfr								X			
	O08530	A3491	Mm.982	601974	Edg1								X			
	Q80VZ2	--	Mm.3249	602188	Epha4								X			
	P54762	A0542	Mm.22897	600600	EPHB1								X			
	Q63645	--	Mm.1614	600933	F2rl1								X			
	Q920E0	--	Rn.81079	--	F2rl3								X			
	O35658	A1608	Mm.30049	601269	C1qbp								X			
	P60894	--	Mm.167625	605188	GPR85								X			
	O88703	A3486	Mm.12956	602781	Hcn2								X			
	Q9NZN1	--	Mm.267669	300206	IL1RAPL1								X			
	Q61291	A1175	Mm.271854	107770	Lrp1								X			
	P15209	A0506	Mm.130054	600456	Ntrk2								X			
	P70206	A3482	Mm.3789	601055	Plxna1								X			
	P70207	A3483	Mm.2251	601054	Plxna2								X			
	Q63495	A1094	Mm.3383	600214	Ager								X			
	Q91V14	A3489	Mm.252987	606726	Slc12a5								X		X	
	Q62634	A4034	Mm.255631	605208	Slc17a7								X			
	Q9EST0	A3481	Mm.314497	605556	Slc4a10								X			
	P16283	A3488	Mm.5053	106195	Slc4a3								X			
	O88343	A3485	Mm.41044	603345	Slc4a4								X		X	
	P48768	A3484	Mm.241147	601901	Slc8a2								X			
MAGUKs / Adaptors / Scaffolders																
PDZ-domain containing scaffolders	Q9Z216	A0104	Mm.37533	604798	Homer1	X								X	X	
	Q62108	A0013	Mm.27256	602887	Dlg4	X								X		
	Q63622	A0014	Mm.257035	603583	Dlg2			X						X		
	P70175	A0016	Mm.4615	300189	Dlg3			X						X		
	Q9WU13	A0075	Mm.358922	604999	Shank1			X						X	X	
	Q9R093	A0137	Mm.197074	--	Cnksr2				X							
	Q62402	A0015	Mm.382	601014	Dlg1				X					X		
	Q9JLU4	A0074	Mm.146855	606230	Shank3					X						
	Q9QYX6	A0568	Mm.332219	604918	Pclo						X					
	P39447	A0086	Mm.4342	601009	Tjp1						X			X		

Appendix 8

						Times detected in PSDs								Complexes		
						8	7	6	5	4	3	2	1			
Protein Class	**UniProt Acc**	**PPID Acc**	**Unigene cluster**	**OMIM**	**Gene name**	**Consensus PSD**								**MA**	**mG**	**AM**
	O70589	A0125	Mm.327591	300172	Cask							X				
	Q9WV34	A0436	Mm.36242	600723	Mpp2							X		X		
	O88910	A0435	Mm.20449	601114	Mpp3							X		X		
	Q925T6	A0301	Mm.196692	604597	Grip1							X			X	
	Q9C0E4	A0302	Mm.333264	--	GRIP2								X			
	Q99JP6	--	Mm.10022	604800	Homer3								X			
	Q62083	A0449	Mm.259464	605926	Prkcabp								X			
	Q9H0F6	A0602	Mm.41463	--	DKFZp434N1923								X			
	Q9WV37	A0793	Mm.41288	606959	Mpp6								X			
	P70625	A1881	Mm.104744	607709	Tjp2								X			
non-PDZ-domain containg scaffolders	Q60437	A1015	Mm.197534	605475	Baiap2	X										
	P97841	A0107	Mm.311840	605445	Dlgap1				X					X		
	Q9Y2A7	A0545	Mm.25203	604891	NCKAP1					X						
	P24587	A0143	Mm.311452	604688	Akap5						X			X		
	O35413	A0109	Mm.211096	--	Argbp2						X					
	O88881	A0228	Mm.342085	--	BEGAIN						X					
	Q8VHK2	A3990	Mm.70989	--	Caskin1						X					
	O08917	A3496	Mm.2931	606998	Flot1						X					
	O94938	A4033	Hs.511386	--	KIAA0856						X					
	Q60803	A0862	Mm.27431	601896	Traf3						X					
	P97837	A0447	Mm.323861	605438	Dlgap2							X				
	Q8R550	--	Mm.286495	300374	Sh3kbp1							X				
	Q62417	A0100	Mm.210815	605264	Sorbs1							X				
	P25304	A3494	Mm.273098	103320	Agrn								X			
	O54931	A0796	Mm.348266	604582	Akap2								X			
	O70333	A0275	Mm.109329	604594	1200020A08Rik								X			
	Q9Y2H0	--	Mm.22094	--	DLGAP4								X			
	Q60902	A3493	Mm.288894	--	Eps15l1								X			
	Q60634	A3497	Mm.130227	131560	Flot2								X			
	Q60760	A0270	Mm.273117	601523	Grb10								X			
	Q60631	A0126	Mm.6900	108355	Grb2								X	X		
	Q9EQR5	A3495	Mm.272712	--	2810038K19Rik								X			
	Q9EPM9	A3492	Mm.256536	604889	Nbea								X			
	P98083	A0171	Mm.86595	600560	Shc1								X			
	Q62245	A0159	Mm.60975	182530	Sos1								X			
	Q99ME2	--	Mm.193395	--	Wdr6								X			
	Q99P24	A0009	Mm.46044	604001	Akap9								X	X		
	Q8K3H0	A0188	Mm.202322	604299	Appl1									X		
	O35145	A0162	Mm.2720	604641	Mapk8ip1									X		
	P68040	A0124	Mm.5305	176981	GNB2L1									X		
14-3-3	P35213	A0467	Mm.34319	601289	Ywhab					X					X	
	P35215	A0361	Mm.332735	601288	Ywhaz					X				X	X	
	P62260	A0280	Mm.234700	605066	YWHAE						X			X		
	P68510	A0633	Mm.332314	113508	YWHAH						X			X	X	
	P61982	A0362	Mm.233813	605356	YWHAG						X			X	X	
	P35216	A0907	Mm.289630	--	Ywhaq						X				X	
	Q9JJ20	A0901	Mm.44482	601290	Sfn								X			
Ser/Thr Kinases					**Gene name**											
	P11798	A0011	Mm.131530	114078	Camk2a	X								X		
	P28652	A1851	Mm.4857	607707	Camk2b		X							X		
	P63319	A0259	Mm.7980	176980	Prkcg			X						X		
	P15791	A1853	Mm.255822	607708	Camk2d				X							
	P11730	A1852	Mm.235182	602123	Camk2g				X							
	Q9JLM8	A3506	Mm.295263	604742	Dcamkl1					X						
	P12369	A0306	Mm.25594	176912	Prkar2b					X				X		
	P04411	A0168	Mm.207496	176970	Prkcb1					X				X		
	Q05512	A3505	Mm.258986	600526	Mark2						X					
	P26817	A0813	Mm.254144	109635	Adrbk1						X					
	Q9JM02	A3503	Mm.42067	--	Map4k6						X					
	P20444	A0307	Mm.222178	176960	Prkca						X					
	P16054	A0144	Mm.287660	176975	Prkce						X			X		
	Q60737	A0753	Mm.23692	115440	Csnk2a1							X			X	
	Q99PS2	A4036	Mm.30199	600863	Csnk1e							X				
	Q63092	A3499	Mm.274540	--	1G5							X				
	P68181	A4037	Mm.16766	176892	PRKACB							X		X		
	P70336	A3501	Mm.276024	604002	Rock2							X				
	Q9WUN2	A3507	Mm.34580	604834	Tbk1							X				
	Q9UKE5	--	Mm.126193	--	TNIK							X				
	Q9UPV4	A3502	Mm.116833	--	AAK1								X			
	Q60823	A0466	Mm.177194	164731	Akt2								X	X		
	O54875	A3504	Mm.27397	--	Cdc42bpb								X			
	P49615	A0667	Mm.298798	123831	Cdk5								X			
	Q92772	--	Mm.44963	603442	CDKL2								X			
	O76039	--	Mm.336310	300203	CDKL5								X			
	P97633	A0896	Mm.26908	600505	Csnk1a1								X			
	P67871	--	Mm.7405	115441	CSNK2B								X			
	P63086	A0181	Mm.196581	176948	Mapk1								X	X		
	Q9WV60	A0222	Mm.200770	605004	Gsk3b								X	X		
	Q60700	A1730	Mm.172097	600447	Map3k12								X			
	P97820	A0177	Mm.19073	604666	Map4k4								X			
	P31938	A0266	Mm.248907	176872	Map2k1								X	X		
	P12367	A0305	Mm.253102	176910	Prkar2a								X			
	P31750	A0189	Mm.6645	164730	Akt1								X			
	Q9DBC7	A0304	Mm.30039	188830	Prkar1a								X			
	Q63433	A0871	Mm.213000	601032	Pkn1								X			
	Q61410	A3500	Mm.263002	601591	Prkg2								X			
	Q9CST0	A0337	Mm.374825	608938	Rps6kb1								X			
	P18654	A0268	Mm.328476	300075	Rps6ka3								X	X		
	Q9Z1W9	A0166	Mm.198414	607648	Stk39								X	X		
	Q8TBQ5	A3529	Mm.260934	608128	DUET								X			
	Q91YW5	A0265	Mm.8385	601795	Mapk3									X		
	O09110	A0491	Mm.18494	602315	Map2k3									X		

Appendix 8

						Times detected in PSDs								Complexes		
						8	7	6	5	4	3	2	1			
Protein Class	UniProt Acc	PPID Acc	Unigene cluster	OMIM	Gene name	Consensus PSD								MA	mG	AM
	Q61831	A0489	Mm.39253	602897	Mapk10								X	X		
	Q63932	A0267	Mm.275436	601263	Map2k2								X	X		
	O35406	A0164	Mm.3906	603014	Map2k7								X	X		
	Q62767	A0287	Mm.170276	602747	Dusp4								X	X		
	Q9Z2A0	A0260	Mm.10504	605213	Pdpk1								X	X		
	Q99N57	A0138	Mm.184163	164760	Raf1								X	X		
Tyr Kinase																
	P39688	A0019	Mm.4848	137025	Fyn							X				
	P05480	A0018	Mm.22845	190090	Src							X		X		
	P25911	A0522	Mm.317331	165120	Lyn								X			
	P06240	A0463	Mm.293753	153390	Lck								X			
	P16879	A3508	Mm.48757	190030	Fes								X			
	Q9QVP9	A0030	Mm.21613	601212	Ptk2b								X	X		
	Q04736	A0553	Mm.4558	164880	Yes								X			
Protein Phosphatases																
	Q9Z1G2	A0147	Mm.277629	176875	Ppp1ca		X									
	O35274	A0273	Mm.229087	603325	Ppp1r9b			X								
	P63088	A4042	Mm.280784	176914	Ppp1cc				X					X		
	P62143	A4043	Mm.241931	600590	PPP1CB					X						
	P63329	A0123	Mm.331389	114105	Ppp3ca						X			X		
	Q96DH3	A0295	Mm.294138	605983	PPP2R1A						X			X		
	Q9QW00	A2624	Mm.258771	601576	Ptpsigma						X					
	O35867	A0510	Mm.332901	602468	Ppp1r9a							X				
	P13353	A0132	Mm.260288	176915	Ppp2ca							X		X		
	Q64487	A1198	Mm.184021	601598	Ptprd							X				
	Q9WUT8	A1202	Mm.41639	176891	Ptprz1							X				
	Q9D7X3	--	Mm.196295	600183	Dusp3								X			
	Q8R3L3	--	Mm.266191	608867	Dusp10								X			
	Q8VEL2	A4016	Mm.197816	--	1110061O04Rik								X			
	Q8CB81	--	Mm.341988	--	Ppm1e								X			
	P35235	A0020	Mm.8681	176876	Ptpn11								X	X		
	Q9EQ17	A0223	Mm.29855	179590	Ptprf								X			
	O60228	A3509	Mm.35483	603560	SBF1								X			
	O35299	A0289	Mm.3294	600658	Ppp5c									X		
	P54830	A0118	Mm.4654	176879	Ptpn5									X		
G-proteins and Modulators																
G-proteins	Q9Z2Q6	A3518	Mm.20365	602724	SEPT5			X							X	
	P18872	A0043	Mm.251445	139311	Gnao1				X						X	
	O55131	A3520	Mm.270259	603151	SEPT7				X					X	X	
	Q9QWG8	A0069	Mm.2344	139380	Gnb1					X				X		
	Q9R1T4	A3519	Mm.260036	--	SEPT6					X						
	P01118	--	Hs.505033	190070	KRAS						X					
	P42208	A3515	Mm.242324	601506	SEPT2						X					
	Q9Z1S5	A3516	Mm.309707	608314	SEPT3						X					
	Q9ESF7	A3525	Mm.274399	608418	SEPT8						X					
	P08752	A0055	Mm.196464	139360	Gnai2							X				
	O70443	A0052	Mm.32595	139160	Gnaz							X				
	P62880	A0070	Mm.30141	139390	GNB2							X		X		
	P54314	A0073	Mm.17604	604447	Gnb5							X				
	P63012	A0037	Mm.5083	179490	Rab3a							X		X	X	
	P05810	A0203	Mm.27348	179550	Rala							X		X		
	P84096	--	Mm.259795	179505	RHOG							X				
	Q9NVA2	A3524	Mm.236587	--	SEPT11							X				
	Q9QZR6	A3521	Mm.38450	604061	SEPT9							X				
	Q91Z61	--	Mm.44995	607862	Diras1								X			
	Q96HU8	A4017	Mm.25648	--	DIRAS2								X			
	P21279	A0042	Mm.195898	600998	Gnaq								X		X	
	P63095	A0051	Mm.125770	139320	GNAS								X	X		
	Q61411	A0200	Mm.334313	190020	Hras1								X	X		
	Q9NYS0	--	Mm.25648	604496	NKIRAS1								X			
	Q9H067	A3974	Mm.261491	--	DKFZp586D0222								X			
	O08989	A3512	Mm.2045	608435	Mras								X			
	P08556	A0199	Mm.256975	164790	Nras								X			
	O88565	A0359	Mm.329123	179512	Rab5a								X			
	P55258	A0620	Mm.162811	165040	Rab8a								X			
	P61027	A0621	Mm.74596	--	Rab10								X		X	
	P35284	A3513	Mm.248313	--	Rab12								X	X		
	P15154	A0194	Mm.292510	602048	RAC1								X	X		
	Q9TU25	--	Mm.1972	602049	RAC2								X			
	P62835	A0323	Mm.333868	179520	RAP1A								X			
	P28661	A3517	Mm.2214	603696	SEPT4								X			
	P29387	A0072	Mm.139192	--	Gnb4									X		
	P53994	A0363	Mm.240224	179509	Rab2									X		
	Q9JJD4	A0364	Mm.28650	179513	Rab6a									X		
	Q9JKM7	A0365	Mm.143789	--	Rab37									X		
	P62826	A0152	Mm.297440	601179	RAN									X		
	Q80ZJ1	A0033	Mm.261448	179540	5830461H18Rik									X		
Modulators	O88938	A0276	Mm.8321	605629	Cit		X							X	X	
	Q9ESK6	A0024	Mm.291291	603384	Syngap1		X							X		
	Q9JIF2	A0214	Mm.260934	604605	Kalirin				X					X		
	Q9JHW8	A3514	Mm.336761	605476	Centg1						X					
	Q9Z1P0	A3528	Mm.196153	606058	Rapgef4						X					
	P06837	A0215	Mm.1222	162060	Gap43						X			X		
	Q9Z272	A3532	Mm.290182	608434	Git1						X					
	Q96D85	A2428	Mm.196943	--	IQSEC1						X					
	Q9C0H5	A3530	Mm.291713	--	KIAA1688						X					
	Q7L804	--	Mm.24167	608599	RAB11FIP2						X					
	Q9ESG7	A0843	Mm.239329	607560	Arhgef2							X				
	O08757	A0646	Mm.244068	605477	Arhgef7							X				
	Q96P47	--	Mm.250703	--	CENTG3							X				
	Q96JJ3		Mm.35064	606421	ELMO2							X				

Appendix 8

						Times detected in PSDs							Complexes		
						8	7	6	5	4	3	2	1			
Protein Class	**UniProt Acc**	**PPID Acc**	**Unigene cluster**	**OMIM**	**Gene name**	Consensus PSD								**MA**	**mG**	**AM**
	Q9NYI0	A2433	Mm.32525	--	PSD3							X				
	Q04690	A0196	Mm.255596	162200	Nf1							X		X		
	Q9D8S3	--	Mm.258910	--	Arfgap3								X			
	Q9QX73	A0707	Mm.44841	300429	Arhgef9								X			
	O75689	--	Hs.135183	608114	CENTA1								X			
	Q9UPQ3	--	Mm.291135	608651	CENTG2								X			
	Q8TDL2	--	Mm.29629	--	Dab2ip								X			
	Q8IZD9		Mm.150259	603123	DOCK3								X			
	Q9ESQ7	A2431	Hs.154658	602327	Psd								X			
	Q9JL26	A4046	Mm.138913	604656	Fmnl1								X			
	P97379		Mm.290530	--	G3bp2								X			
	P50398	A0567	Mm.205830	300104	Gdi1								X		X	
	Q9DAS9	A0067	Mm.234342	--	Gng12								X			
	Q921M4	A3209	Mm.106376	602580	Golga2								X			
	O35550	A0589	Mm.7087	603616	Rabep1								X			
	Q923H2	A1069	Mm.143717	605991	Ngef								X			
	Q8CHG7	A0034	Mm.31220	--	Rapgef2								X			
	Q94764	A4047	Mm.148781	603854	RANBP9								X			
	Q8K2L6	--	Mm.180763	600278	Rap1ga1								X			
	Q8BU92	--	Hs.167371	606136	A330066M24Rik								X			
	P70392	--	Mm.248630	606614	Rasgrf2								X			
	P49801	--	Mm.153013	603894	Rgs6								X			
	O54829	A0932	Mm.7956	602517	Rgs7								X			
	P49804	--	Hs.458417	607189	Rgs8								X			
	Q9QZB0	--	Mm.44606	607191	Rgs17								X			
	Q9Z0G0	A1159	Mm.20945	605072	Rgs19ip1								X			
	Q60992	--	Mm.179011	600428	Vav2								X			
	Q91ZT5	A0470	Mm.256131	--	Fgd4									X		
Signalling molecules and Enzymes																
Heat shock / Chaperones / Chaperonins	P08109	A0146	Mm.290774	600816	Hspa8			X								
	P17879	A2007	Mm.6388	140550	Hspa1b			X						X		
	P20029	A0451	Mm.330160	138120	Hspa5					X						
	Q9QYJ0	A3535	Mm.279692	--	Dnaja2						X					
	P38647	A2010	Mm.209419	600548	Hspa9a						X					
	P11499	A2003	Mm.2180	140572	Hspcb						X					
	O54946	A1578	Mm.290110	--	Dnajb6						X					
	P11983	A3541	Mm.32019	186980	Cct1							X				
	P80314	A3540	Mm.247788	605139	Cct2							X				
	P80315	A3538	Mm.296985	605142	Cct4							X				
	P80316	A3543	Mm.282158	--	Cct5							X				
	P54102	A3536	Mm.27897	602837	Dnaja1							X				
	P19226	A3550	Mm.1777	118190	Hspd1							X				
	O43301	A3534	Mm.39739	--	HSPA12A							X				
	Q91YN9	--	Hs.55220	603882	Bag2								X			
	P80318	A3544	Mm.256034	600114	Cct3								X			
	P80317	A3545	Mm.153159	104613	Cct6a								X			
	P80313	A3539	Mm.289900	605140	Cct7								X			
	P42932	A3542	Mm.328673	--	Cct8								X			
	P23927	--	Mm.178	123590	Cryab								X			
	Q99M87	--	Mm.248337	--	Dnaja3								X			
	Q9QYI5	--	Mm.248776	604139	Dnajb10								X			
	O75061	A0685	Mm.76494	608375	DNAJC6								X			
	P08113	A2004	Mm.87773	191175	Tra1								X			
	Q61699	A3533	Mm.270681	--	Hsph1								X			
	P04792	--	Mm.13849	602195	HSPB1								X			
	Q9NVH1	A3537	Mm.21353	--	DNAJC11								X			
	Q922R8	A3548	Mm.222825	--	Txndc7								X			
	O55042	A0905	Mm.17484	160090	Snca								X			
	Q9CY50	A3549	Mm.138725	600868	Ssr1								X			
Phosphodiesterases	P16330	A3553	Mm.15711	123830	Cnp				X							
	Q01062	A3551	Mm.247564	602658	Pde2a						X					
	Q8VBU5	--	Mm.20181	600127	Pde4b								X			
	Q01063	--	Mm.343333	600129	Pde4d								X			
	Q9QYJ5	A3552	Mm.87161	--	Pde10a								X			
	Q8VID8	--	Mm.246613	604961	Pde11a_v1								X			
Adenylate/Guanylate cyclase	Q9WVI4	A1979	Mm.371695	601244	Gucy1a2								X			
Secreted / secretory / signalling	Q61361	A3556	Mm.4598	--	Bcan					X						
	Q62059	A3555	Mm.158700	118661	Cspg2						X					
	O88998	A3557	Mm.43278	605366	Olfm1							X				
	O35382	A1304	Mm.265512	608185	Sec8l1								X			
ATP synthases	P56480	A0378	Mm.238973	102910	Atp5b					X						
	P50516	A0397	Mm.217787	607027	Atp6v1a					X				X		
	P62815	A0400	Mm.249096	606939	Atp6v1b2					X						
	Q03265	A0377	Mm.276137	164360	Atp5a1						X			X		
	P51863	A0403	Mm.17708	607028	Atp6v0d1						X			X		
	P50518	A0404	Mm.29045	108746	Atp6v1e1						X					
	Q925I1	A3559	Mm.241152	--	Atad3a						X					
	P35435	A0379	Mm.12677	108729	Atp5c1								X	X		
		A0383	Mm.251152	603270									X			
	Q9Z1G3	A0401	Mm.276618	603097	Atp6v1c1								X			
	P57746	A0402	Mm.311549	--	Atp6v1d								X			
NADH-Ubiquinone Oxidoreductase	Q16795	A3563	Mm.29939	603834	NDUFA9						X					
	P28331	A3561	Mm.290791	157655	NDUFS1						X					
	Q9DCT2	A3564	Mm.30113	603846	Ndufs3						X					
	Q9D6J6	A0431	Mm.2206	600532	Ndufv2							X		X		
	Q62425	--	Mm.41926	603833	Ndufa4								X			
	Q9DCJ5	--	Mm.19834	603359	Ndufa8								X			
	Q99LC3	A3560	Mm.248778	603835	Ndufa10								X			
	Q91WD5	A3562	Mm.21669	602985	Ndufs2								X			
Polyubiquitin	Q62317	A1257	Mm.331	191340	Rps27a						X					
	P62986	--	Mm.43005	191321	UBA52								X			
	Q63446	A1256	Mm.282093	191339	Rps27a								X			

Appendix 8

						Times detected in PSDs							Complexes		
						8	7	6	5	4	3	2	1			
Protein Class	**UniProt Acc**	**PPID Acc**	**Unigene cluster**	**OMIM**	**Gene name**	Consensus PSD								MA	mG	AM
Proteases / protease inhibitors	O88543	A3570	Mm.40	604665	Cops3								X			
	P46101	A3565	Mm.42078	126141	Dpp6								X			
	Q9EQF2	A3571	Mm.19958	110900	Kel								X			
	O35593	A3566	Mm.218198	607173	Psmd14								X			
	Q9R1P0	A3567	Mm.30270	--	Psma4								X			
	Q63569	--	Mm.289832	186852	Psmc3								X			
	O88761	A3568	Mm.280013	--	Psmd1								X			
	Q13200	--	Mm.243234	606223	PSMD2								X			
	Q99JB5	A3569	Mm.273152	--	Psmd8								X			
	Q9Z2X3	--	Mm.17640	603480	Psmd10								X			
	P48594	--	Mm.283677	600518	SERPINB4								X			
	Q9JJI7	--	Mm.243926	606797	St14								X			
	O88399	A1731	Mm.236068	--	Synj2								X			
Development	O88558	A3579	Mm.37249	606322	Cyfip1						X					
	O88741	A3576	Mm.18218	606598	Gdap1						X					
	Q9D2N4	A3577	Mm.94371	601239	Dtna							X				
	Q8VE33	A3980	Mm.102080	--	Gdap1l1							X				
	Q05175	A3574	Mm.29586	605940	Basp1								X			
	P18572	A3573	Mm.726	109480	Bsg								X			
	Q9CYN5	A3575	Mm.86523	--	5730405I09Rik								X			
	Q9JIQ6	A3998	Mm.39388	607802	Cnnm1								X			
	Q9CQQ2	A4001	Mm.87417	606626	Daam1								X			
	Q9QXL0	A3578	Mm.248788	602865	Dbccr1								X			
	Q03173	--	Mm.87759	--	Enah								X			
	Q9QZY2	A3581	Mm.41812	--	Gprin1								X			
	Q62813	A3572	Mm.310524	603241	Lsamp								X			
	Q8K377	--	Mm.292568	--	Lrrtm1								X			
	O43300	--	Mm.39900	--	LRRN2								X			
	Q86VH4	--	Mm.94135	--	LRRTM4								X			
	O08919	A2635	Mm.255487	604018	Numbl								X			
	Q924Y6	--	Mm.32831	605186	Wif1								X			
Other signalling molecules	Q61358	A0095	Mm.44736	602377	Dnm1				X					X	X	
	P62157	A0008	Mm.285993	114180	CaM					X				X	X	
	P39052	A1927	Mm.39292	602378	Dnm2						X				X	
	Q9NRX7	--	Mm.335310	607815	Eb-1						X					
	P04925	A4012	Mm.648	176640	Prnp						X					
	Q9WV31	A0277	Mm.25405	--	Arc							X		X		
	Q9NZT1	--	Mm.21075	605183	CALML5							X				
	P49621	--	Mm.242576	604070	Dgkb							X				
	Q8R3T9	A1928	Mm.29567	--	Dnm3							X				
	O70465	A4013	Mm.254610	--	Odz4							X				
	O08957	--	Mm.232930	607409	Nrn1							X				
	O88377		Mm.39700	603261	Pip5k2b							X				
	O94811	--	Mm.39752	608773	Tppp							X				
	Q9WUD1	A3978	Mm.277599	607207	Stub1							X				
	O35639	--	Mm.7214	106490	Anxa3								X			
	Q91VR9	--	Mm.5061	--	Arf2								X			
	O35143	--	Mm.2171	--	Atpif1								X			
	O75498	A1114	Mm.9749	604515	BLNK								X			
	P63030	--	Mm.288510	--	Brp44l								X			
	O08918	--	Mm.3527	603203	Ccng2								X			
	P27274	A1630	Mm.247265	107271	Cd59								X			
	Q99KW9	A4005	Mm.334685	--	Tip								X			
	P61810	A0666	Mm.142275	603460	Cdk5r1								X			
	Q9JLH7	A4007	Mm.28297	608202	Cdk5rap3								X			
	Q9Z140		Mm.5249	605688	Cpne6								X			
	Q9Z0J4	A0262	Mm.44249	163731	Nos1								X	X		
	Q9Z1B3	A0330	Mm.330607	607120	Pkd1								X	X	X	
	Q9JHK5	A0919	Mm.98232	173570	Plek								X			
	Q9ES97	--	Mm.246990	604249	Rtn3								X			
	Q9JK11	--	Mm.192580	604475	Rtn4								X			
	P05942	--	Mm.3925	114210	S100a4								X			
	P06702	A1551	Mm.2128	123886	S100A9								X			
	P50114	A1342	Mm.235998	176990	S100b								X			
	P07092	A1547	Mm.3093	177010	Serpine2								X			
	O88602	A0473	Mm.277338	602911	Cacng2								X	X		
	O55106	A2120	Mm.311915	--	Strn								X			
	P58404	--	Mm.21612	--	Strn4								X			
	O75962	A0739	Hs.130031	601893	TRIO								X			
	P20156	A2164	Mm.297818	602186	Vgf								X			
	P61206	A0421	Mm.221298	103190	ARF3									X		
	Q61337	A0476	Mm.4387	603167	Bad									X		
	P47728	A0437	Mm.2755	114051	Calb2									X	X	
	P47713	A0106	Mm.4186	600522	Pla2g4a									X		
	Q00731	A0434	Mm.282184	192240	Vegf									X		
	P35637	A0217	Mm.277680	137070	FUS									X		
	P53762	A0488	Mm.250265	126110	Arnt									X		
	P35569	A0484	Mm.4952	147545	Irs1									X		
	Q12888	A0286	Mm.215389	605230	TP53BP1									X		
	Q62077	A0012	Mm.44463	172420	Plcg1									X		
	Q9NQ38	A4014	Mm.35369	605010	SPINK5									X		
Other enzymes	P05064	A1923	Mm.275831	103850	Aldoa				X						X	
	P15105	A0424	Mm.210745	138290	Glul				X					X		
	P97427	A3605	Mm.290995	602462	Crmp1					X						
	Q62167	A3653	Mm.289662	300160	Ddx3x					X						
	Q9WUA3	A1926	Mm.273874	171840	Pfkp					X						
	P05063	A0428	Mm.7729	103870	Aldoc						X			X	X	
	P07335	A3693	Mm.16831	123280	Ckb						X					
	Q9NQC7	--	Mm.24282	605018	CYLD						X					
	Q91VR5	A3656	Mm.251255	601257	Ddx1						X					
	O60275	A2429	Mm.33027	--	IQSEC2						X					

Appendix 8

Protein Class	UniProt Acc	PPID Acc	Unigene cluster	OMIM	Gene name	Times detected in PSDs: Consensus PSD 8	7	6	5	4	3	2	1	Complexes: MA	mG	AM
	Q9CPU4	A1948	Mm.218286	604564	Mgst3						X					
	Q9DBJ1	--	Mm.315962	172250	Pgam1						X					
	P42337	A0153	Mm.260521	171834	Pik3ca						X			X		
	O08662	A3658	Hs.529438	600286	Pik4ca						X					
	P10111	--	Mm.5246	123840	Ppia						X					
	O88541	--	Mm.325816	123841	Ppib						X					
	P17751	A0433	Mm.4222	190450	Tpi1						X			X		
	Q00981	--	Mm.29807	191342	Uchl1						X					
	Q91V92	A3586	Mm.282039	108728	Acly							X				
	O54734	A3619	Mm.7236	602202	Ddost							X				
	Q61656	A0814	Mm.220038	180630	Ddx5							X				
	O35098	A3606	Mm.250414	608407	Dpysl4							X				
	P17182	A3615	Mm.70666	172430	Eno1							X				
	Q63835	A3659	Mm.267478	604443	Acsl6							X				
	Q9EQR0	A3584	Mm.236443	600212	Fasn							X				
	Q9Z2W5	A3592	Mm.28188	601336	Gcs1							X				
	O15294	--	Mm.259191	300255	OGT							X				
	P12382	A1925	Mm.269649	171860	Pfkl							X		X		
	P47858	A0261	Mm.272582	232800	Pfkm							X				
	P35700	A0371	Mm.30929	176763	Prdx1							X		X		
	P50137	--	Mm.290692	606781	Tkt							X				
	Q9R1T2	A1052	Mm.258530	--	Uble1a								X			
	Q925C4	A3582	Mm.31374	200350	Acac								X			
	P55266	--	Mm.316628	601059	Adar								X			
	Q9D1E8	A3630	Mm.24117	--	Agpat5								X			
	P51635	--	Mm.30085	103830	Akr1a1								X			
	P24549	A3585	Mm.250866	100640	Aldh1a1								X			
	P11274	A0902	Mm.333722	151410	BCR								X			
	Q99PU5	A3593	Mm.20592	--	Lpd								X			
	O88738	--	Mm.290908	605638	Birc6								X			
	P24270	--	Mm.4215	115500	Cat								X			
	Q8VDP6	--	Mm.28219	605893	Cdipt								X			
	P52825	--	Mm.307620	600650	Cpt2								X			
	Q9D5V5	--	Mm.218910	601741	Cul5								X			
	Q64678	--	Mm.214016	601771	Cyp1b1								X			
	P16381	A3652	Mm.108054	--	D1Pas1								X			
	O08560	A1300	Mm.314923	601441	Dgkz								X			
	Q99J47	A3651	Mm.21475	--	BC003479								X			
	Q9BQF9	A3641	Mm.29678	--	DKFZp761F069								X			
	Q9WU83	A3638	Mm.201322	603503	DPM1								X			
	Q9JMG8	A3610	Mm.27732	608383	Dpysl5								X			
	P21550	A3617	Mm.251322	131370	Eno3								X			
	P17183	A3616	Mm.3913	131360	Eno2								X			
	P70428	--	Mm.4336	133701	Ext2								X			
	Q9ESW4	A3620	Mm.32840	--	2610037M15Rik								X			
	Q8TE79	A3634	Mm.11827	--	Comtd1								X			
	P28037	A3594	Mm.30035	600249	Aldh1l1								X			
	Q9Z148	A0444	Mm.35345	604599	Bat8								X	X		
	Q9Z339	A1969	Mm.29457	605482	Gsto1								X			
	O54865	A1980	Mm.9445	139397	Gucy1b3								X			
	O70351	--	Mm.6994	300256	Hadh2								X			
	Q9DD18	--	Mm.28109	--	Hars2								X			
	P97852	--	Mm.277857	601860	Hsd17b4								X			
	O70503	A3633	Mm.22505	--	Hsd17b12								X			
	P97618	--	Mm.183042	602064	Impa1								X			
	Q62784	A3587	Mm.150420	600916	Inpp4a								X			
	O75043	A3640	Mm.140138	607813	KIAA0455								X			
	P06151	A0370	Mm.29324	150000	Ldha								X	X		
	P16125	A0369	Mm.9745	150100	Ldhb								X	X		
	P21396	A3672	Mm.21108	309850	Maoa								X			
	Q99MR8	A3604	Mm.249016	210200	Mccc1								X			
	P48318	A0635	Mm.272120	605363	Gad1								X			
	P34884	--	Mm.2326	153620	Mif								X			
	P46935	A1245	Mm.279923	602278	Nedd4								X			
	Q7Z5F1	--	Mm.98668	606384	NEDD4L								X			
	Q01768	--	Mm.1260	156491	Nme2								X			
	Q91VJ8	--	Mm.29824	606793	Npepps								X			
	Q9EQU4	A3635	Mm.27427	--	Dhrs4								X			
	Q08642	--	Mm.2296	607935	Padi2								X			
	Q91ZA3	A3601	Mm.23876	232000	Pcca								X			
	P15259	--	Mm.219627	261670	PGAM2								X			
	Q9CXY9	A3622	Mm.331447	605087	Pigk								X			
	O70172	--	Mm.313977	603140	Pip5k2a								X			
	O70161	A3591	Mm.29836	606102	Pip5k1c								X			
	P53657	A0427	Mm.8359	266200	Pklr								X	X		
	Q9CR08	--	Mm.22284	606114	Pop4								X			
	Q99JY8	A3632	Mm.348326	607125	Ppap2b								X			
	Q9D7G0	A3631	Mm.287178	311850	Prps1								X			
	Q922W5	A3657	Mm.127731	179035	Pycr1								X			
	P53534	A3588	Mm.222584	138550	Pygb								X			
	Q925F3	--	Mm.214932	--	Rnf144								X			
	Q9EPK2	A3625	Mm.288141	312600	Rp2h								X			
	P07153	A3639	Mm.188544	180470	Rpn1								X			
	P25235	A3611	Mm.22130	180490	Rpn2								X			
	Q91YM8	--	Mm.197486	180410	Rrm1								X			
	Q9QZX7	--	Mm.131443	606477	Srr								X			
	O76062	A2012	Mm.301370	603414	TM7SF2								X			
	Q61171	A0372	Mm.270130	600538	Prdx2								X	X		
	Q9CZY3	--	Mm.278783	602995	Ube2v1								X			
	P56399	--	Mm.3571	601447	Usp5								X			
Mitochondrial enzymes	P16858	A0331	Mm.333399	138400	Gapd		X							X		
	P30275	A3690	Mm.252145	123290	Ckmt1			X								

Appendix 8

Protein Class	UniProt Acc	PPID Acc	Unigene cluster	OMIM	Gene name	Times detected in PSDs: Consensus PSD 8	7	6	5	4	3	2	1	Complexes: MA	mG	AM
	O08553	A0442	Mm.235283	602463	Dpysl2			X						X		
	P26443	A3680	Mm.10600	138130	Glud1				X							
	Q99KI0	A0429	Mm.154581	100850	Aco2						X			X		
	P39069	--	Mm.29189	103000	Ak1						X					
	P05202	A0423	Mm.230169	138150	Got2						X			X		
	Q64428	A3683	Mm.200497	600890	Hadha						X					
	P17710	A3663	Mm.196605	142600	Hk1						X					
	P08249	A3696	Mm.297096	154100	Mdh2						X					
	Q9WUM5	A3697	Mm.29845	--	Suclg1						X					
	P17764	A3694	Mm.293233	203750	Acat1							X				
	P29147	A3702	Mm.293470	603063	Bdh							X				
	Q91VA7	A3668	Mm.29590	604526	Idh3b							X				
	P97700	A3731	Mm.296082	604165	Slc25a11							X				
	P07895	--	Mm.290876	147460	Sod2							X				
	Q9Z2I9	A3621	Mm.38951	603921	Sucla2							X				
	P32551	A3677	Mm.334206	191329	Uqcrc2							X				
	Q9QWW0	A0432	Mm.186185	602316	Prdx6								X	X		
	O55028	--	Mm.8903	--	Bckdk								X			
	Q9CZU6	A3688	Mm.58836	118950	Cs								X			
	Q9D0M3	A3698	Mm.29196	123980	Cyc1								X			
	P53395	A3675	Mm.3636	248610	Dbt								X			
	Q64591	A3699	Mm.24395	222745	Decr1								X			
	P08461	A0605	Mm.285076	109720	Dlat								X	X		
	O08749	A3674	Mm.3131	246900	Dld								X			
	Q01205	A3682	Mm.296221	126063	Dlst								X			
	O08600	A3783	Mm.4449	600440	Endog								X			
	Q9DCT9	A3779	Mm.28336	231675	Etfdh								X			
	P28492	A3670	Mm.30102	606365	Gls2								X			
	Q64521	A3667	Mm.3711	138430	Gpd2								X			
	O70325	--	Mm.332810	138322	Gpx4								X			
	Q60587	A3684	Mm.291463	143450	Hadhb								X			
	Q9EQK1	A3660	Mm.2966	147650	Idh2								X			
	Q9D6R2	A3629	Mm.279195	601149	Idh3a								X			
	P70404	A3669	Mm.14825	300089	Idh3g								X			
	Q8R2M5	A3780	Mm.331142	--	2010002H18Rik								X			
	P00405	A3778	Mitochondrial	516040	Mtco2								X			
	Q60597	A3664	Mm.276348	203740	Ogdh								X			
	Q9ULD0	A3665	Hs.17860	--	KIAA1290								X			
	P04166	--	Mm.20242	--	Cyb5m								X			
	Q05920	A3662	Mm.1845	266150	Pc								X			
	Q9Z0X1	A3781	Mm.240434	300169	Pdcd8								X			
	P35486	A3685	Mm.34775	312170	Pdha1								X			
	P49432	A3687	Mm.301527	179060	Pdhb								X			
	Q9CZ13	A3676	Mm.194811	191328	Uqcrc1								X			
Transcription and Translation																
Transcription Elements	Q61336	A0439	Mm.36241	--	Bcap37				X					X		
	Q924D3	A3580	Mm.154358	606323	Cyfip2					X						
	P42669	A3795	Mm.231802	600473	Pura					X						
	P24142	A3793	Mm.263862	176705	Phb						X					
	Q9R0S1	A3788	Mm.29827	602181	Snd1							X				
	Q9JK31	A3789	Mm.173271	--	Atf7ip								X			
	P46100	A2062	Mm.10141	300032	ATRX								X			
	Q9UIG0	A2168	Mm.40331	605681	BAZ1B								X			
	Q9JI19	A3798	Mm.329656	608296	Fibp								X			
	Q9D0E1	A3792	Mm.311439	160994	Hnrpm								X			
	O54910	A1330	Mm.57043	604548	Nfkbie								X			
	Q9DCS3	A3794	Mm.192706	608205	Nrbf1								X			
	P59759	--	Mm.270643	--	Mrtfb								X			
	Q9Z1N2	A3790	Mm.294154	601902	Orc1l								X			
	O35295	A3796	Mm.296150	608887	Purb								X			
	P40646	A2191	Mm.42162	--	Sox7								X			
	Q8C3V4	--	Mm.277406	600555	Stat1								X			
	P40630	A2186	Mm.229292	600438	Tfam								X			
	Q8K055	A3799	Mm.22786	602290	Trim32								X			
	Q9QZP7	A2922	Mm.254997	--	Zfp386								X			
	P58501	A0298	Mm.347	--	Gcfc									X		
Other nuclear/DNA binding	O54962	--	Mm.299167	603811	Banf1							X				
	P47875	A1021	Mm.196484	123876	Csrp1							X				
	Q99MR0	--	Mm.31646	--	Actl6b								X			
	Q9Z1N7	A2027	Mm.350789	--	Arid3b								X			
	Q8NE35	--	Mm.371677	--	CPEB3								X			
	Q9CW03	A3213	Mm.14910	606062	Cspg6								X			
	Q9Z1X8	A2251	Mm.262287	607112	Fbxo2								X			
	Q8TF61	--	Hs.23158	--	FBXO41								X			
	P84249	A1373	Hs.546259	601128	H3-IX								X			
	Q8TDG4	A2152	Hs.480101	606769	HEL308								X			
	P43274	--	Mm.14101	142220	Hist1h1e								X			
	P28001	--	Mm.261665	602786	HIST1H2AE								X			
	Q811M0	--	Mm.227295	142750	Hist1h4h								X			
	Q9D2U9	--	Mm.28022	601831	Hist3h2ba								X			
	O70592	A2547	Mm.316306	--	APa								X			
	O35737	--	Mm.21740	601035	Hnrph1								X			
	P70333	--	Mm.315909	601036	Hnrph2								X			
	Q9NYF4	--	Mm.277906	--	JMJD1B								X			
	Q9QWV4	--	Mm.10414	601402	Mlf1								X			
	Q969Z1	--	Mm.218940	--	MMS19L								X			
	P28656	--	Mm.337558	--	Nap1l1								X			
	Q9EPT6	--	Mm.31979	607538	Ndel1								X			
	Q60862	A3791	Mm.3411	601182	Orc2l								X			
	P60335	--	Mm.274146	601209	Pcbp1								X			
	O00541	--	Mm.28659	605819	PES1								X			
	Q86YI3	--	Hs.154658	602327	PSD								X			

Appendix 8

Protein Class	UniProt Acc	PPID Acc	Unigene cluster	OMIM	Gene name	Times detected in PSDs: 8	7	6	5	4	3	2	1	Complexes: MA	mG	AM
						Consensus PSD										
	Q61493	--	Mm.288788	602776	Rev3l								X			
	Q9UJC5	--	Hs.302772	--	SH3BGRL2								X			
	P28370	A2045	Mm.229151	300012	SMARCA1								X			
	Q60596	--	Mm.333661	194360	Xrcc1								X			
	Q99K28	--	Mm.43636	606908	Zfp289								X			
	Q9Y4E5	--	Mm.289103	--	ZNF451								X			
Elongation / RNA binding	P62632	A3801	Mm.2645	602959	Eef1a2				X							
	P58252	A3804	Mm.289431	130610	Eef2					X						
	P10126	A1820	Mm.196614	130590	Eef1a1						X					
	O70251	A3802	Mm.2718	600655	Eef1b								X			
	Q9D8N0	A1968	Mm.247762	130593	Eef1g								X			
	O70194	A3805	Mm.3955	603915	Eif3s7								X			
	P10630	--	Mm.260084	601102	Eif4a2								X			
	Q922K6	A3806	Mm.290022	603928	Eif4b								X			
	P63074	A0973	Mm.3941	133440	Eif4e								X			
	P29341	A2508	Mm.371570	604679	Pabpc1								X			
	O54725	--	Mm.257276	605199	Sfpq								X			
Ribosomal Proteins	P47911	A3810	Mm.262021	603703	Rpl6					X						
	P62909	A3816	Mm.236868	600454	Rps3					X						
	P09058	A3823	Mm.260904	600357	RPS8					X						
	P50878	A3808	Mm.280083	180479	Rpl4						X					
	P47962	A1222	Mm.4419	603634	Rpl5						X					
	P14148	A0366	Mm.290772	604166	Rpl7						X			X		
	P53026	A3813	Mm.336955	--	Rpl10a						X					
	Q9CXW4	--	Mm.276856	604175	Rpl11						X					
	P14869	A3807	Mm.5286	180510	Rplp0						X					
	Q9CZX8	--	Mm.300281	603474	Rps19						X					
	Q8BHL0	--	Mm.4071	150370	Lamr1							X				
	P12970	A3811	Mm.725	185640	Rpl7a							X				
	Q9Z237	A3812	Mm.30066	604177	Rpl8							X				
	P23358		Mm.286100	180475	Rpl12							X				
	P47963	A0367	Mm.319719	113703	Rpl13							X		X		
	Q63507	A3814	Mm.289810	--	Rpl14							X				
	P35980	--	Mm.349277	604179	Rpl18							X				
	P14118	--	Mm.309019	180466	RPL19							X				
	P67985	--	Mm.307846	180474	Rpl22							X				
	P23131	--	Mm.140380	603662	RPL23							X				
	Q9CSM4	--	Mm.340658	607526	Rpl27							X				
	P25444	A3815	Mm.1129	603624	Rps2							X				
	P97351	A3817	Mm.6957	180478	Rps3a							X				
	P47961	A3818	Mm.66	312760	RPS4							X				
	P10660	--	Mm.325584	180460	RPS6							X				
	P46781	--	Mm.13944	603631	RPS9							X				
	P63326	A1668	Mm.275810	603632	Rps10							X				
	P62842	--	Mm.288212	180535	RPS15							X				
	P62250	--	Mm.702	603675	RPS16							X				
	P63274	--	Mm.42767	180472	RPS17							X				
	P62270	--	Mm.371578	180473	RPS18							X				
	Q63742	A3286	Mm.172720	--	RBP34								X			
	P27659	--	Mm.290899	604163	Rpl3								X			
	P51410	--	Mm.300271	603686	Rpl9								X			
	Q6PDV7	--	Mm.100113	312173	Rpl10								X			
	Q6ZWZ7	--	Mm.276337	603661	Rpl17								X			
	P62751	--	Mm.244545	602326	RPL23A								X			
	P83731	--	Mm.282814	604180	RPL24								X			
	P62889	--	Mm.358632	180467	RPL30								X			
	P62900	--	Mm.298467	--	RPL31								X			
	P47964	--	Mm.11376	--	Rpl36								X			
	P63174	--	Mm.238817	604182	RPL38								X			
	P47955	--	Mm.3158	180520	Rplp1								X			
	P02401	--	Mm.341719	180530	Rplp2								X			
	P46782	--	Mm.5291	603630	RPS5								X			
	P23821	--	Mm.371579	603658	RPS7								X			
	P63324	--	Mm.371846	603660	Rps12								X			
	P62278	--	Mm.14798	180476	RPS13								X			
	P62264	--	Mm.43778	130620	RPS14								X			
	Q921M2	--	Mm.21938	603682	Rps20								X			
	P63221	--	Mm.289669	180477	RPS21								X			
	P25111	--	Mm.292027	180465	RPS25								X			
	P24051	--	Mm.270283	603702	Rps27l								X			
	P49664	--	Mm.180003	191343	Rps27a								X			
	Q6ZWY3	--	Mm.30120	--									X			
	P62858	--	Mm.200920	603685	RPS28								X			
	P19253	A0368	Mm.180458	--	Rpl13a									X		
Cytoskeletal and Cell Adhesion Molecules																
Actin / ARP	O35247	A0091	Mm.196173	102560	Actin		X							X		
	O15145	A0232	Mm.275942	604225	ARPC3						X			X		
	P59999	A0231	Mm.289306	604226	ARPC4						X			X		
	P53488	A0235	Mm.259045	604221	ACTR2							X				
	Q99JY9	A0234	Mm.183102	604222	Actr3							X				
	Q9R0Q6	A3287	Mm.34695	604220	Arpc1a							X				
	O15144	A0233	Mm.337038	604224	ARPC2							X		X		
	Q9P1U1	A3339	Mm.150319	--	ARP3beta								X			
Tubulin	P68362	A0017	Mm.209290	602529	Tuba1	X								X		X
actinin	Q9Z1P2	A2341	Mm.253564	102575	Actn1		X									
	Q9JI91	A0007	Mm.37638	102573	Actn2					X				X		
	P57780	A2344	Mm.276042	604638	Actn4						X			X		
	O88990	A2343	Mm.5316	102574	Actn3									X		
Spectrin	Q62261	A0010	Mm.123110	182790	Sptbn1	X								X		
MAPs	P34926	A0085	Mm.36501	600178	Map1a				X						X	
	P20357	A0134	Mm.256966	157130	Map2				X					X	X	
	Q9QXZ0	A2342	Mm.3350	608271	Macf1					X						

Appendix 8

						Times detected in PSDs								
						8	7	6	5	4	3	2	1	Complexes
Protein Class	UniProt Acc	PPID Acc	Unigene cluster	OMIM	Gene name	Consensus PSD								MA mG AM
	P14873	A0625	Mm.4173	157129	Map1b					X				X
	P10638	A0135	Mm.1287	157140	Mapt						X			
	O55129	A3376	Mm.154087	601783	Stop						X			
	P27816	A0778	Mm.217318	157132	MAP4								X	
	Q15555	A0654	Mm.132237	605789	MAPRE2								X	
Catenins	Q61301	A1914	Mm.34637	114025	Ctnna2				X					
	Q02248	A0112	Mm.291928	116806	Ctnnb1				X					X
	O35927	A0469	Mm.321648	604275	Ctnnd2					X				
	P30999	A0218	Mm.35738	601045	Ctnnd1							X		
	P26231	A0089	Mm.18962	116805	Ctnna1								X	
Myelin	Q61885	A0351	Mm.210857	159465	Mog					X				X
	Q62079	A0352	Mm.1268	300401	Plp1					X				X X
	P04370	A0350	Mm.252063	159430	Mbp						X			X
	Q63912	A3590	Mm.93335	164345	Omg							X		
	P35802	A3425	Mm.241700	601275	Gpm6a								X	
Cell recognition molecules	P97685	A0663	Mm.326702	--	Nfasc				X					
	Q05546	A1074	Mm.337795	601995	Tnr						X			
	P97547	A3607	Mm.15125	--	Sdfr1								X	
	Q62701	A1525	Mm.980	187380	Tnc								X	
Keratins	Q9CWC9	A3741	Mm.158190	--	2410039E07Rik								X	
	Q9CV17	A3825	Mm.46378	607742	2310058N18Rik								X	
	Q920F2	A1669	Mm.22657	148040	Krt2-5								X	
	Q9Z332	A3846	Mm.302399	148041	Krt6a								X	
Other Cytoskeletal Proteins	P46660	A0292	Mm.276251	605338	Ina	X								X
	P08553	A0376	Mm.242832	162250	Nef3	X								X
	P08551	A0375	Mm.1956	162280	Nefl		X							X
	O88737	A0285	Mm.20425	604020	Bsn			X						X X
	P30427	A1941	Mm.234912	601282	Plec1			X						
	Q9QYC0	A3876	Mm.289106	102680	Add1				X					X
	P47754	A0417	Mm.262039	601571	Capza2				X					X
	P47757	A0418	Mm.2945	601572	Capzb				X					X
	Q9QX74	A0081	Mm.323725	603290	Shank2				X					X
	Q9QXS6	A0438	Mm.19016	126660	Dbn1				X					X
	Q9EPW6	A3976	Mm.217161	602330	Ablim1					X				X
	Q9QYB8	A0587	Mm.104155	102681	Add2					X				X
	Q01484	A3873	Mm.220242	106410	ANK2					X				
	Q61307	A3874	Mm.235960	600465	Ank3					X				
	O15083	A3879	Mm.318004	--	ERC2					X				
	P45592	A0999	Mm.4024	601442	Cfl1					X				
	Q99MI2	A3878	Mm.288860	607127	Rab6ip2					X				
	Q9JMB3	A3865	Mm.131135	605331	Epb4.1l3					X				
	Q9Z327	A3871	Mm.252321	--	Synpo					X				X
	P01831	A1218	Mm.3951	188230	Thy1					X				
	Q9ERQ9	A0148	Mm.41353	605035	Wasf1					X				
	Q9QYB5	A3877	Mm.44106	601568	Add3						X			
	P13020	A0419	Mm.21109	137350	Gsn						X			X
	Q61792	A3882	Mm.271967	602920	Lasp1						X			
	O35482	A0374	Mm.298283	162230	NF-H						X			
	Q9Z0P4	A3883	Mm.34650	608134	Palm						X			
	P61164	A0660	Mm.3118	605143	ACTR1A							X		
	Q9R1V7	A3881	Mm.124892	603710	Adam23							X		
	Q02357	A0662	Mm.334444	182900	Ank1							X		
	P47753	--	Mm.19142	601580	Capza1							X		
	Q8BVG1	--	Mm.224189	--	4732477G22Rik							X		
	Q8C5W0	A3870	Mm.244078	--	Clmn							X		
	Q8VCQ5	A2142	Mm.260158	605269	Coro1c							X		
	Q8R555	A3975	Mm.329851	606276	Crtac1							X		
	Q9Z2H5	A3867	Mm.20852	602879	Epb41l1							X		
	Q96ID5	A3979	Hs.212511	--	MGC15730							X		
	P70565	A0324	Mm.299774	173325	Jup							X		
	Q14008	A4009	Mm.168478	--	KIAA0097							X		
	O35449	A3989	Mm.373627	--	Ng5							X		
	Q99569	--	Mm.260938	604276	PKP4							X		
	O75145	A3869	Mm.277235	603144	PPFIA3							X		
	O76049	--	Mm.205647	603050	ABI1							X		
	P37805	--	Mm.24183	607953	Tagln3							X		
	Q9ESJ4	A1027	Mm.192416	606671	Nckipsd							X		
	Q80WK6	--	Mm.254446	--	Ablim2								X	
	Q9R1V4	A3880	Mm.89854	155120	Adam11								X	
	Q9R1V6	A0373	Mm.125477	603709	Adam22								X	X
	Q91ZU0	--	Mm.34785	--	Asb7								X	
	Q8R3F2	A3211	Mm.334999	--	Ckap4								X	
	O89053	A2143	Mm.290482	605000	Coro1a								X	
	O89046	A2141	Mm.276859	--	Coro1b								X	
	Q9UQ03	A2139	Mm.335229	605002	CORO2B								X	
	Q60598	A0083	Mm.205601	164765	Cttn								X	X
	P48675	A0535	Mm.6712	125660	Des								X	
	Q61495	A0291	Mm.37953	125670	Dsg1a								X	X
	P11531	A0462	Mm.275608	310200	Dmd								X	
	Q61167	A0654	Mm.57874	605789	Mapre3								X	
	Q15075	A2330	Mm.210035	605070	EEA1								X	
	Q8JZV5	--	Mm.210863	125305	Epb4.9								X	
	P26040	A0326	Mm.277812	123900	Vil2								X	
	Q61553	A0224	Mm.289707	602689	Fscn1								X	
	Q9QYE7	A1687	Rn.34728	--	Itgad								X	
	Q9ERD4	A3875	Mm.250641	--	Kidins220								X	
	P48679	A1374	Mm.243014	150330	Lmna								X	
	Q9D052	A3991	Mm.130883	--	Lyric								X	
	P60853	--	Mm.334576	606551	Lzts1								X	
	P63005	A2074	Mm.56337	601545	Pafah1b1								X	
	Q96JQ0	--	Mm.334108	603057	DCHS1								X	
	Q91Y06	--	Mm.196700	606332	Pcdhb13								X	

Appendix 8

						Times detected in PSDs								
						8	7	6	5	4	3	2	1	Complexes
Protein Class	UniProt Acc	PPID Acc	Unigene cluster	OMIM	Gene name	Consensus PSD								MA mG AM
	O75334	--	Mm.252890	603143	PPFIA2								X	
	Q91Z80	--	Mm.295105	603145	Ppfia4								X	
	Q9JJV2	--	Mm.271744	176590	Pfn2								X	
	P70567	--	Mm.249594	190930	Tmod1								X	
	Q9NSL0	A3872	Mm.29043	--	DKFZp762B245								X	
	Q9UP98	A0446	Mm.22417	601342	CSE1L									X
	Q9JJ38	A0204	Mm.39046	300017	Flnc									X
	P14733	A0430	Mm.4105	150340	Lmnb1									X
	Q8VC86	A0216	Mm.36769	602701	Slmap									X
Other Cell Adhesion Molecules	P70587	A0150	Mm.132162	--	Lrrc7		X							
	P15116	A0219	Mm.257437	114020	Cdh2				X					X
	P23242	A0493	Mm.338720	121014	Gja1					X				
	O35889	A0101	Mm.181959	159559	Mllt4						X			
	O54991	A0758	Mm.8021	602346	Cntnap1						X			
	P12960	A0757	Mm.4911	600016	Cntn1						X			
	Q61330	A1073	Mm.260861	190197	Cntn2						X			
	Q9QUP5	A3893	Mm.266790	115435	Hapln1						X			
	P55066	A3892	Mm.268079	600826	Cspg3						X			
	P11627	A0290	Mm.260568	308840	L1cam						X			X X
	Q9JIA1	A0440	Mm.298251	604619	Lgi1						X			X
	Q9Z0J8	A3888	Mm.317293	--	Negr1						X			
	Q62765	A0129	Mm.316080	600568	Nlgn1						X			
	Q62888	A3889	Mm.151293	606479	Nlgn2						X			
	Q62889	A3890	Mm.121508	300336	Nlgn3						X			
	Q07314	A3891	Mm.291005	600567	Nrxn3						X			
	Q9Z251	A0141	Mm.268025	603380	Lin7a						X			X
	Q9Z252	A3899	Mm.20472	--	Lin7b						X			
	Q61137	A3896	Mm.329586	600904	Astn							X		
	Q9NR18	A0507	Mm.277354	606944	ERBB2IP							X		
	Q62718	A3887	Mm.283138	607938	Nt							X		
	P13595	A3884	Mm.4974	116930	Ncam1							X		
	Q9Z0E0	A1163	Mm.276466	608458	Ncdn							X		
	Q9ULB1	A0130	Mm.329258	600565	NRXN1							X		
	P32736	A3886	Mm.39634	600632	Opcml							X		
	Q61282	--	Mm.358571	155760	Agc1								X	X
	Q06890	A1276	Mm.200608	185430	Clu								X	
	P98203	--	Mm.293599	602269	Arvcf								X	
	Q9UHW6	A3897	Mm.22786	--	ASTN2								X	
	Q91ZU6	A1575	Mm.336625	--	Dst								X	
	Q9Y6N8	--	Mm.234715	604555	CDH10								X	
	Q9WTR5	--	Mm.334841	601364	Cdh13								X	
	Q9QY32	A3895	Mm.38496	606775	Cspg5								X	
	P15924	A0454	Mm.355327	125647	DSP								X	X
	P70689	--	Mm.25652	604418	Gjb6								X	
	Q9ESM3	--	Mm.294467	--	Hapln2								X	
	P20917	A3898	Mm.241355	159460	Mag								X	
	O35136	A3885	Mm.258759	602040	Ncam2								X	
	P97798	A3894	Mm.42249	601907	Neo1								X	
	Q9EP53	--	Mm.224354	605284	Tsc1								X	
Synaptic Vesicles and Protein Transport														
Clathrin	P11442	A0119	Mm.254588	118955	Cltc					X				X X
	P08081	A0122	Mm.298875	118960	Clta									
	P08082	A3901	Mm.290026	118970	Cltb									
Synaptic vesicle	O88935	A0161	Mm.370211	313440	Syn1	X								
	Q9QWV7	A3904	Mm.275845	600755	Syn2	X								
	P46460	A0117	Mm.260117	601633	Nsf			X						X X
	P47708	A0212	Mm.181166	--	Rph3a			X						X
	P46096	A0039	Mm.289702	185605	Syt1				X					X
	P17426	A0251	Mm.6877	601026	Ap2a1					X				
	Q62768	A0560	Mm.334606	--	Unc13a					X				
	O55125	A3914	Mm.293716	603249	Nipsnap1					X				
	Q99NE5	A0035	Mm.328661	606629	Rims1					X				
	P60881	A0333	Mm.45953	600322	SNAP25					X				X
	Q9QXY3	A3903	Mm.342665	--	P140					X				X
	O08599	A0238	Mm.278865	602926	Stxbp1					X				X X
	P52303	A0686	Mm.274816	600157	Ap1b1						X			
	P17427	A0252	Mm.253090	607242	Ap2a2						X			
	P20172	A0254	Mm.18946	601024	AP2M1						X			
	O54774	A1324	Mm.209294	607246	Ap3d1						X			
	O55126	A3913	Mm.12468	603004	Gbas						X			
	Q9QXG3	A0040	Mm.6225	186590	Stx1a						X			X
	Q9Z270	A3937	Mm.266767	605703	Vapa						X			
	O08838	A0121	Mm.101650	600418	Amph							X		
	O08839	A1109	Mm.4383	601248	Bin1							X		
	Q9DBG3	A0253	Mm.39053	601025	Ap2b1							X		X
	Q00380	A0255	Mm.333597	602242	Ap2s1							X		
	Q9EP81	A2346	Mm.255138	603432	Cyln2							X		
	Q61595	A3531	Mm.296457	600381	Ktn1							X		
	P28663	A3910	Mm.274308	--	Napb							X		
	Q9Z0W5	A0501	Mm.4926	606512	Pacsin1							X		
	O55011	A3927	Mm.235175	603025	Picalm							X		
	Q9EQZ7	A1861	Mm.309296	606630	Rims2							X		
	Q64548	A0445	Mm.221275	600865	Rtn1							X		X
	Q02563	A3902	Mm.200365	185860	Sv2							X		
	Q9WVG8	A3933	Mm.26032	603927	Syngr3							X		
	Q62910	A0120	Mm.187079	604297	Synj1							X		X
	Q62277	A3941	Mm.223674	313475	Syp							X		
	Q9R0N7	A3905	Mm.182654	604146	Syt7							X		
	Q99JB2	A3942	Mm.26751	--	Stoml2							X		
	Q91ZY8	--	Mm.184012	--	Trim9							X		
	Q920Q4	A3908	Mm.276092	608550	Vps16							X		
	Q61123	A3909	Mm.296520	606931	Vps35							X		

Appendix 8

Protein Class	UniProt Acc	PPID Acc	Unigene cluster	OMIM	Gene name	Times detected in PSDs: Consensus PSD 8	7	6	5	4	3	2	1	Complexes: MA	mG	AM
	P10107	A0647	Mm.248360	151690	Anxa1								X			
	Q61548	A0684	Mm.281651	607923	Snap91								X			
	Q9JME5	--	Mm.322894	602166	Ap3b2								X			
	P53678	A3907	Mm.10647	--	Ap3m2								X			
	P35564	A0452	Mm.248827	114217	Canx								X			
	P53621	A2102	Mm.30041	601924	COPA								X			
	O55029	A0237	Mm.261735	606990	Copb2								X			
	Q01469	A0441	Mm.741	605168	FABP5								X	X		
	Q9QXY6	A1331	Mm.18526	605891	Ehd3								X			
	Q9Z1Z3	A3926	Mm.139695	607263	Epn2								X			
	O54922	A3921	Mm.22530	608163	Exoc7								X			
	O54924	A3923	Mm.347360	--	Exo84								X			
	O88630	A3936	Mm.20931	604026	Gosr1								X			
	Q9JKY5	A3906	Mm.149954	605613	Hip1r								X			
	P97411	A2211	Mm.275683	147625	Ica1								X			
	O88994	A4008	Mm.177724	--	Mg87								X			
	Q9Y3A3	A3935	Mm.291037	--	PREI3								X			
	Q8K0T7	--	Mm.41035	--	Unc13c								X			
	P54921	A0596	Mm.104540	603215	Napa								X			
	Q9CWZ7	A3911	Mm.292687	603216	Napg								X			
	Q62443	A1697	Mm.5142	602367	Nptx1								X			
	Q9JMB9	A3932	Mm.151332	--	Pex5l								X			
	Q80U40	--	Mm.233996	--	Rimbp2								X			
	Q8R361	--	Mm.220334	605536	Rip11								X			
	Q62991	A3928	Mm.216511	--	Scfd1								X			
	P97878	A3925	Mm.31607	604469	Sec10l1								X			
	Q9D4R7	A3924	Mm.159621	607880	Sec15l2								X			
	Q9D4H1	A3920	Mm.293510	--	Sec5l1								X			
	Q62825	A3922	Mm.261859	608186	Sec6l1								X			
	Q8VIL3	--	Rn.7624	--	Zwint								X			
	O60493	--	Mm.277870	605930	SNX3								X			
	P61264	--	Mm.40343	--	Stx1b2								X			
	O70439	A3940	Mm.248042	603217	Stx7								X			
	Q9ER00	A3939	Mm.28237	606892	Stx12								X			
	Q15431	A3863	Mm.243849	602162	SYCP1								X			
	Q9QZA3	--	Mm.374805	602705	Syn3								X			
	O55100	A0288	Mm.230301	603925	Syngr1								X	X	X	
	Q9DCD5	A4027	Mm.284594	--	Tjp4								X			
	Q9Z152	A0562	Mm.286868	604586	Stxbp5								X			
	Q8IYF7	--	Mm.17436	605073	TRIM37								X			
	P63044	--	Mm.28643	185881	Vamp2								X		X	
	Q9QY76	A3938	Mm.260456	605704	Vapb								X			
	Q01853	A3918	Mm.262053	601023	Vcp								X			
	Q8BMZ7	--	Mm.30028	--	Vps28								X			
	Q9QZ88	--	Mm.216528	606932	Vps29								X			
	Q63615	A3930	Mm.41372	--	Vps33a								X			
	Q9H348	A1326	Mm.27389	605485	HVPS41								X			
	P97390	A3929	Mm.263185	--	Vps45a								X			
	O88384	A3934	Mm.265929	603207	Vti1b								X			
Transporters	P48962	A0395	Mm.16228	103220	Slc25a4				X					X		
	Q99LZ6	A3949	Mm.28023	--	Mtch2						X					
	Q9JHY2	A3952	Mm.36169	--	Sfxn3						X					
	O75746	A0443	Mm.30928	603667	SLC25A12						X			X		
	Q9H936	A3773	Mm.33729	--	SLC25A22						X			X		
	P55096	A3946	Mm.194462	170995	Abcd3							X				
	Q01405	A3945	Mm.209207	--	Sec23a							X				
	Q925N0	A3954	Mm.121485	--	Sfxn5							X				
	P16036	A3776	Mm.298	600370	Slc25a3							X				
	Q09073	A0396	Mm.658	300150	Slc25a5							X		X		
	P06795	--	Mm.146649	--	Abcb1								X			
	Q9CXJ4	A3947	Mm.283914	605464	4833412N02Rik								X			
	Q9QZP4	A3948	Mm.285322	--	Mtch1								X			
	P47802	A3956	Mm.280943	600605	Mtx1								X			
	O88441	A3957	Mm.292613	608555	Mtx2								X			
	O08547	--	Mm.2551	604029	Sec22l1								X			
	Q63965	A3951	Mm.134191	--	Sfxn1								X			
	P32089	A3750	Mm.229291	190315	Slc25a1								X			
	Q9QXX4	A3747	Mm.24513	603859	Slc25a13								X			
	P97441	A3943	Mm.1396	602878	Slc30a3								X			
	Q9CZW5	A3950	Mm.213292	606081	Tomm70a								X			
	P40336	A3944	Mm.260703	605506	Vps26								X			
Motor Proteins	Q9QYF3	A0149	Mm.3645	160777	Myo5a		X							X		
	P70566	A3972	Mm.291880	602928	Tmod2				X							
	Q99JI3	A3970	Mm.222272	605853	Clasp2					X						
	Q9D9F6	A3973	Mm.235123	600378	Immt					X						
	Q8VDD5	A3166	Mm.29677	160775	Myh9					X				X		
	Q9JMH9	A3964	Mm.341248	--	Myo18a					X						
	Q63608	A3965	Mm.121878	191010	Tpm1					X						
	P38650	A0133	Mm.181430	600112	Dnchc1						X					
	P63168	A0514	Mm.256858	601562	DNCL1						X					
	P28740	A3966	Mm.355686	602591	Kif2						X					
	Q9JLT0	A3961	Mm.218233	160776	Myh10						X			X		
	O08638	A2331	Mm.250705	160745	Myh11						X			X		
	O08788	A0658	Mm.6919	601143	Dctn1						X					
	P21107	A3205	Mm.240839	191030	Tpm3						X					
	Q9R1R2	A1178	Mm.29290	605493	Trim3							X				
	P70611	A0561	Mm.266301	604567	Doc2a							X				
	O88485	A0661	Mm.20893	603772	Dnci1							X				
	P60521	--	Mm.371666	607452	Gabarapl2							X				
	Q61768	A3967	Mm.223744	602809	Kif5b							X				
	Q63781	A4048	Mm.306770	--	Mrlcb							X		X		
	Q9H882	A3962	Mm.358743	608568	Myh14							X				

Appendix 8

Protein Class	UniProt Acc	PPID Acc	Unigene cluster	OMIM	Gene name	Times detected in PSDs: Consensus PSD 8	7	6	5	4	3	2	1	Complexes: MA	mG	AM
	Q60605	A4049	Mm.337074	--	Myl6							X		X		
	P46735	A3958	Mm.3390	606537	Myo1b							X		X		
	Q64331	A3960	Mm.4040	600970	Myo6							X				
	P21807	A0459	Mm.2477	170710	Prph							X				
	P31000	A0850	Mm.268000	193060	Vim							X				
	Q99KJ8	A0656	Mm.167537	607376	Dctn2								X			
	Q9QUR2	A3969	Mm.272801	--	Dctn4								X			
	Q9D0M5	--	Mm.246436	608942	Dlc2								X			
	Q8R1Q8	--	Mm.128627	--	Dnclic1								X			
	O88487	A2127	Mm.249479	603331	Dnci2								X			
	Q60780	--	Mm.40338	603127	Gas7								X			
	Q9Z0R4	A0571	Mm.40546	602442	Itsn1								X			
	P70188	--	Mm.4651	601836	Kifap3								X			
	P28738	--	Mm.256342	604593	Kif5c								X			
	O88448	A3968	Mm.271648	--	Klc2								X	X		
	P09541	--	Mm.247636	160770	Myl4								X			
	Q63357	A3959	Mm.151948	606539	Myo1d								X			
	Q9NQX4	--	Mm.41339	--	MYO5C								X			
	Q99MZ6	--	Mm.312342	606541	Myo7b								X			
	O60313	A3971	Mm.274285	605290	OPA1								X			
	P30622	A2345	Mm.241109	179838	RSN								X			
	P46097	--	Mm.5102	600104	Syt2								X			
	O35681	--	Mm.4824	600327	Syt3								X			
	Q9R0N5	--	Mm.358663	--	Syt5								X			
	P58774	--	Mm.646	190990	Tpm2								X			
	P16475	A4049	Mm.337074	--	MYL6											
	O61355	A3202	Mm.290003	160760	Mhc									X		
Uncharacterised / novel	Q9UNU4	--	Mm.261333	--	SIPA1L1						X					
	Q920I9	A4010	Mm.30850	--	Wdr7						X					
	Q9D5I5	A3994	Mm.20313	--	0710001P09Rik							X				
	Q9D172	--	Mm.268691	601659	D10Jhu81e							X				
	O35127	--	Mm.3172	--	C10							X				
	Q9CRD2	A3977	Mm.244512	607722	Kiaa0103							X				
	Q9ULK3	--	Hs.445885	--	KIAA1217							X				
	Q9C0D0	--	Mm.160124	--	PHACTR1							X				
	Q8BJY6	--	Mm.120298	--	Nebl							X				
	Q63128	--	Mm.157716	--	CPG2							X				
	Q86TV6	--	Mm.295441	--	TTC7B							X				
	Q9DCZ4	--	Mm.371632	--	0610008C08Rik								X			
	Q8VD05	--	Mm.294159	--	0610011D08Rik								X			
	Q8R570	--	Mm.26680	--	1110031B06Rik								X			
	Q78IS1	--	Mm.27606	--	1200002G13Rik								X			
	Q8BMB1	--	Mm.41903	--	1200014J11Rik								X			
	Q9JKC6	--	Mm.87027	--	Bm88								X			
	Q9DB72		Mm.38347	--	1500005I02Rik								X			
	Q9CW19	--	Mm.200385	--	1500041B16Rik								X			
	Q8QZT2	A3996	Mm.157746	--	1700054N08Rik								X			
	Q8WV67	--	Mm.318691	--	RBAF600								X			
	Q9D935	--	Mm.71682	--	1810009B06Rik								X			
	Q80TP4	--	Mm.285968	--	Wdr47								X			
	Q6P069	--	Mm.96211	--	Sri								X			
	Q8C9Z0	--	Mm.337339	--	2310035C23Rik								X			
	Q8K1A7	--	Mm.24829	--	Fbxo31								X			
	Q922F4	--	Mm.181860	--	Tubb6								X			
	Q8BNN1	--	Mm.41459	--	2610039E05Rik								X			
	Q9D023	--	Mm.195625	--	2610205H19Rik								X			
	Q9CZU2	A3995	Mm.61682	--	2610528A17Rik								X			
	Q9Z2D2	A3510	Mm.102970	--	Mtmr13								X			
	Q9CVR0	--	Mm.227260	602660	4930542G03Rik								X			
	Q8C0R4	--	Mm.72926	--	4932414J0								X			
	Q8BUY8	--	Mm.271980	--	5330440H13Rik								X			
	Q8BP77	--	Mm.355327	--	5730453H04Rik								X			
	Q9CYH2	--	Mm.27227	--	5730469M10Rik								X			
	Q80T83	--	Mm.274493	--	5730538E15Rik								X			
	Q8BX25	--	Mm.44490	--	DCAMKL2								X			
	Q8BQV1	--	Mm.44355	--	6430514L14Rik								X			
	Q80Y24	--	Mm.248820	608501	Prickle2								X			
	Q9D217	A4003	Mm.257354	--	A330102H22Rik								X			
	Q8C409	--	Mm.204836	--	A930041I02Rik								X			
	Q8C0J2	A2084	Mm.272972	--	Apg16l									X		
	Q8VDA6	--	Mm.30060	--	Avpi1								X			
	Q9CUI5	--	Mm.92705	--	Azi2								X			
	Q9D5J9	A4015	Mm.199698	--	4930431B11Rik									X		
	Q8K4U9	--	Mm.240850	602968	Bcas1								X			
	Q9HAA5	--	Hs.486708	--	C6ORF103								X			
	Q80W16	--	Mm.4880	--	Cbln1								X			
	Q8VCR7	--	Mm.335427	--	Cib								X			
	Q8C3Q5	--	Mm.347599	--	D430041B17								X			
	Q8BXF0	--	Mm.291609	--	D430041D05Rik								X			
	Q8C3K4	--	Mm.35077	--	D630045J12Rik								X			
	Q8C8Q0	--	Mm.196943	--	D6Ertd349e								X			
	Q9CU45	A3800	Mm.248687	--	6330407J23Rik								X			
	Q9H0C2	A3744	Mm.78691	--	DKFZp434N1235								X			
	Q8K2R1	A0281	Mm.23044	--	9630050M13Rik									X		
	Q9NXY9	--	Hs.410889	--	DKFZP564D166								X			
	Q8VD37	A3987	Mm.238094	--	3110007P09Rik								X			
	Q8R1A4	A3986	Mm.260623	--	Dock7								X			
	Q9CRB9	A3558	Mm.21501	--	Chchd3								X			
	Q12914	A4002	Hs.502266	--	G2								X			
	Q9Y3I0	--	Mm.9257	--	HSPC117								X			
	Q96TC7	A3511	Mm.330492	--	hucep-10								X			
	Q9HCQ3	A2325	Mm.37803	608793	HCMOGT-1								X			

Appendix 8

						Times detected in PSDs								Complexes		
						8	7	6	5	4	3	2	1			
Protein Class	UniProt Acc	PPID Acc	Unigene cluster	OMIM	Gene name	Consensus PSD								MA	mG	AM
	Q95KE9	--	--	--	Hypothetical protein								X			
	Q8R0L7	--	Mm.271717	606644	Igsf8								X			
	Q14700	A4000	Mm.271956	--	KIAA0090								X			
	Q922I2	A4006	Mm.260647	--	C920006C10Rik								X			
	Q92508	--	Mm.37324	--	FAM38A								X			
	Q80U49	--	Mm.23689	--	AW555464								X			
	O60302	A3912	Mm.134602	--	KIAA0555								X			
	Q8C090	--	Mm.134338	--	AU040829								X			
	Q9Y2G0	A4035	Mm.210266	--	KIAA0953								X			
	Q80TJ8	--	Mm.299442	--	AK122446								X			
	Q9ULT3	A3992	Mm.166647	--	GPR158								X			
	Q9ULQ0	A3983	Mm.56097	--	FAM40B								X			
	Q96IB4	--	Hs.505516	--	KIAA1463								X			
	Q9HCM3	A3554	Hs.490294	--	KIAA1549								X			
	Q8TDC3	--	Hs.443752	--	KIAA1811								X			
	Q96PY1	--	Mm.296626	--	KIAA1906								X			
	Q9JJK2	A3988	Mm.274904	--	Lancl2								X			
	Q9Z2I0	A4004	Mm.260538	604407	Letm1								X			
	Q8NE86	A4023	Mm.69732	--	C10orf42								X			
	Q8K4V0	A3999	Mm.217027	--	Lrpprc								X			
	Q9HAY2	--	Hs.306123	--	MAGEF1								X			
	Q8TB68	--	Hs.130316	--	PRR7								X			
	Q9D898	A3290	Mm.38155	--	Arpc5l								X			
	O35606	A4045	Mm.196596	--	C18orf8								X			
	Q8IUF7	--	Mm.289150	--	mina53								X			
	Q99KX1	--	Mm.29737	601401	Mlf2								X			
	P56379	--	Mm.318	--	Mp68								X			
	P61956	--	Mm.29923	603042	SMT3H2								X			
	Q8N8M3	--	Hs.5025	605491	NEBL								X			
	O95897	--	Hs.169743	--	OLFM2								X			
	Q9CRA1	--	Mm.275648	605903	Pdlim7								X			
	Q7TQG1	--	Mm.253559	607771	Plekha6								X			
	P14200	--	Mm.179378	176795	Pmch								X			
	Q9NTK5	--	Mm.22661	--	PTD004								X			
	P52823	--	Mm.20911	601185	STC1								X			
	Q9CUN3	A3993	Mm.260782	--	2810037C14Rik								X			
	O54764	--	Mm.14485	--	Tex9								X			
	Q96Q82	--	Hs.12082	--	tiga1								X			
	Q9CR67	--	Mm.23217	--	Tmem33								X			
	Q16890	--	Mm.7821	604069	TPD52L1								X			
	P14701	--	Mm.296922	600763	Tpt1								X			
	Q9EQN7	A3997	Mm.29729	605784	Ttyh1								X			
	Q9ER48	--	Mm.29722	--	Usmg5								X			
	Q80ZI9	--	Mm.2654	604734	Wdr1								X			
	Q8CBE3	--	Mm.284654	--	Wdr37								X			
Removed from databases	--	--	--	--	gi:892272							X				
		--	--	--	gi:27733035							X				
		--	--	--	gi:27712644								X			
		--	--	--	gi:27666466								X			
		--	--	--	gi:18601785								X			
		--	--	--	gi:27683621								X			
		--	--	--	gi:20909043								X			
		--	--	--	gi:27684483								X			
		--	--	--	gi:27718001								X			
		--	--	--	gi:20864488								X			
		--	--	--	gi:27682685								X			
		--	--	--	gi:20881803								X			
		--	--	--	gi:20843007								X			
		--	--	--	gi:27689767								X			
		--	--	--	gi:27677960								X			
		--	--	--	gi:20876984								X			
		--	--	--	gi:20884391								X			
		--	--	--	gi:27731499								X			
		--	--	--	gi:27686085								X			
Other cell origin	P47819	A0534	Mm.1239	137780	Gfap					X						
	P08758	--	Mm.1620	131230	ANXA5							X				
	Q60771	A1892	Mm.4425	--	Cldn11								X			
	P23943	--	Mm.284855	--	Edn2								X			
	P58107	--	Mm.259929	--	EPPK1								X			
	P14480	A1493	Mm.30063	--	Fgb								X			
	P02680	A1494	Mm.16422	--	Fgg								X			
	Q9ERL8	--	Mm.87312	--	Gmfb								X			
	P01946	A1741	Mm.196110	--	Hba1								X			
	P02091	A1742	Mm.288567	--	Hbb								X			
	P19637	A1496	Mm.154660	173370	Plat								X			

Appendix 8 - Summary of integrated PSD and multiprotein complex datasets. The number of times a protein was found in the PSD (7 MS analyses of the PSD and individual publications) is shown along with the components of 3 multiprotein complexes (MA, MASC; mG, mGluR5 receptor complex and AM, AMPA receptor complex).

Appendix 9

Proteomics in the Neurosciences. In Proteomics: Biomedical and Pharmaceutical Applications. Edited by Hondermarck, H. (Kluwer Academic Press).

Chapter 3

Proteomics in the Neurosciences

MARK O COLLINS, HOLGER HUSI and SETH G N GRANT
Division of Neuroscence, University of Edinburgh, 1 George Square, Edinburgh, EH8 9JZ UK

1. INTRODUCTION

The field of proteomics aims to provide a comprehensive view of the characteristics and function of every protein present within a given biological system. The genome, which encodes the entire complement of genes in an organism, has been seen as the fundamentally complex component of cells. However it has emerged in recent years that its complexity is overshadowed by that of the transcriptome, which describes the full complement of mRNA transcripts transcribed from the genome of the cell. This in turn gives rise to yet another level of complexity see in that of the cellular proteome. This proteome is derived from the variety of alternative splicing events that occur at the level of transcription and the use of alternative start and stop sites and frame-shifting which occurs at the level of translation. This complexity is compounded by the 300 or more different protein modifications, which have been, reported (Aebersold and Goodlett 2001) of which phosphorylation and glycosylation are prototypic.

Comprehensive proteome analysis also seeks to elucidate protein sub-cellular distribution and protein-protein interactions, which are crucial in determining the integration of large numbers of proteins into interaction networks, and functional molecular complexes that ultimately perform a myriad of cellular functions. Finally, the proteome is highly dynamic and temporal and thus describes the complement of proteins expressed in a cell at a single time point.

Appendix 9

Proteomic analysis in any system encompasses many areas and integration of this data into biologically meaningful knowledge is the ultimate goal. The progress and future objectives of each area in relation to the field of neuroscience is discussed and the future of bioinformatic integration of this data is suggested (figure 1).

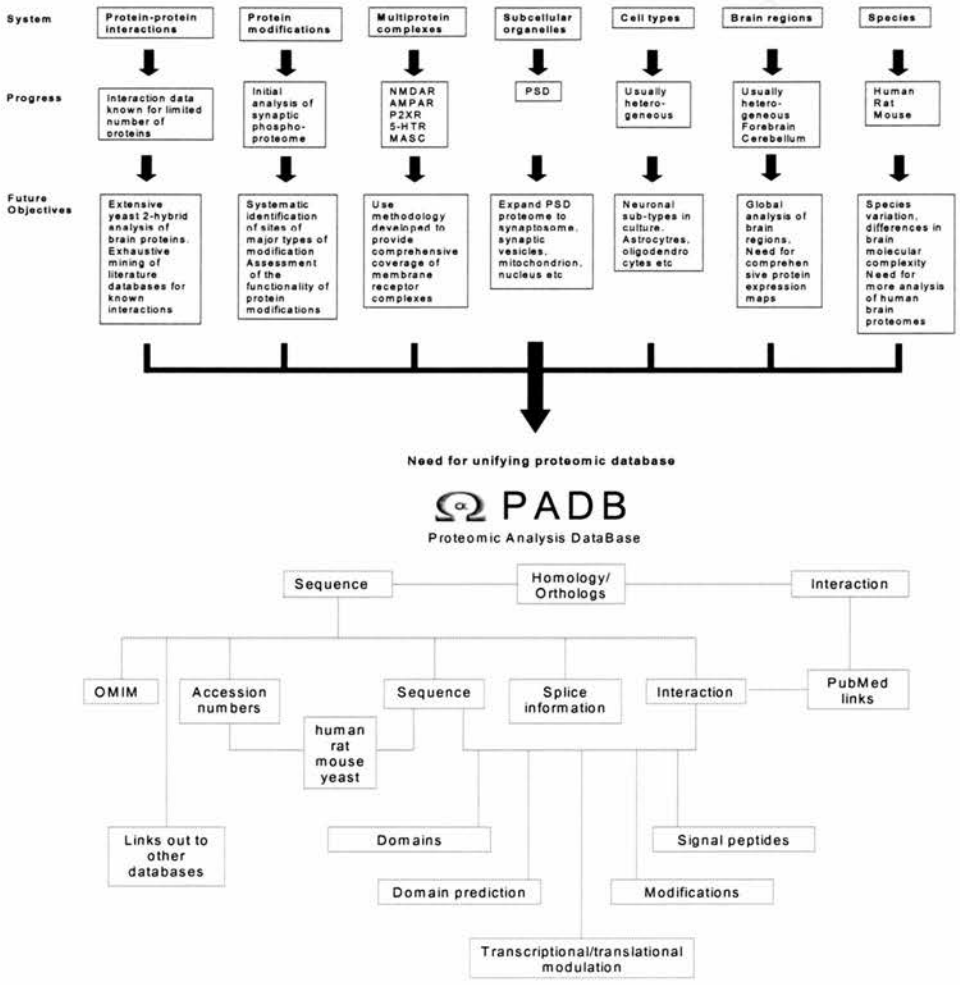


Figure 1. (a) Overview of the progress and objectives of proteomics and the structure of PADB, a proteomic analysis database.

Appendix 9

2. GLOBAL NEUROPROTEOMICS: APPROACHES & OBJECTIVES

2.1 Two-dimensional electrophoresis

The biochemical approach of two-dimensional electrophoresis which has become the classical proteomic approach to whole proteome analysis has the capacity to display a large number of proteins expressed the studied system under given physiological conditions. Construction of global expression maps for defined proteomes is the most widely used application of proteomics and when used in combination with mass spectrometry (MS) techniques can be a powerful approach. There have been a number of studies focused on global neuroproteomics from whole brain analysis to the analysis of synaptic components. Two-dimensional maps have been constructed for whole human (Langen, Berndt et al. 1999) mouse (Gauss, Kalkum et al. 1999) and rat brains (Fountoulakis, Schuller et al. 1999) to analyse the entire brain proteome. These studies yielded more than 200 discrete proteins with significant over lap between datasets. The mouse brain proteome study is the largest to date with 8767 resolved spots and over 300 proteins identified by a combination of proteomics and genetic methods, (figure 2).

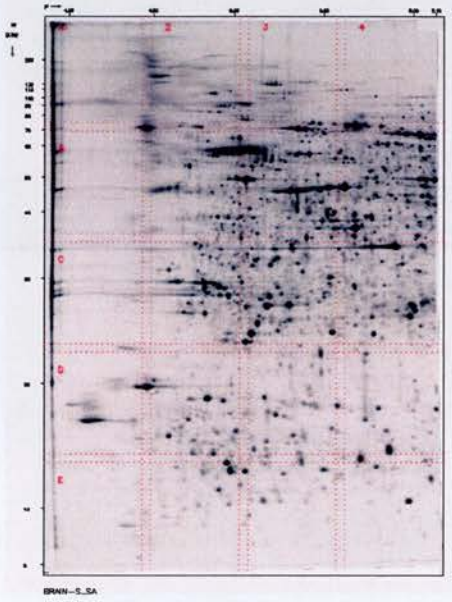


Figure 2. Protein standard map of the mouse brain supernatant fraction obtained by large-gel 2-D electrophoresis followed by densitometry and computer-assisted spot detection (image from Gauss, C., Kalkum, M., Lowe, M., Lehrach, H. & Klose, J. Analysis of the mouse proteome. (I) Brain proteins: separation by two-dimensional electrophoresis and identification by mass spectrometry and genetic variation. Electrophoresis 20, 575-600. (1999)

Appendix 9

A number of brain region proteomes have been studied to investigate differences in protein expression. Preliminary analysis of the mouse cerebellar proteome has identified 30 proteins (Beranova-Giorgianni, Giorgianni et al. 2002) and analysis of the porcine cerebellum led to identification of 56 spots (Friso and Wikstrom 1999). A developmental proteomic study of the rat cerebellum yielded resolution of over 3000 spots and identification of 67 of these (Taoka, Wakamiya et al. 2000). Most proteins showed an increase in abundance as the cerebellum matured, however, 42 spots appeared to be exclusively expressed in the immature cerebellum. Some of the latter were identified by MS and included proteins with defined roles in nervous system development.

Although, the approach of using 2D in conjunction with MS techniques is a classical proteomic approach, it is evident that it falls short as a sole platform for global proteomics. The range of biochemically different proteins, co-migration of proteins and difficulty of resolution of low abundance molecules has hampered the use of this technology to comprehensively analyse highly complex mixtures of proteins present in brain preparations. Problems associated with 2D resolution of membrane proteins have been overcome to some extent (Santoni, Molloy et al. 2000), however gel-free proteomic solutions such as multidimensional protein identification technology (MudPIT) (Wu, MacCoss et al. 2003) allow greatly increased coverage of the membrane proteome. However, 2D-electrophoresis will remain a useful tool in proteomics, due to sophisticated image analysis, software and fluorescent dyes, which now permit more sensitive and accurate comparative proteome analysis.

2.2 Subcellular Proteomics

Another approach which has proved to be very successful in the systematic identification of proteins, has been used to identify the molecular constituents of the post-synaptic proteome (PSP) (Husi 2003). The post-synaptic density is an amorphous structure usually 400 nm long and 40 nm wide and is located beneath the synaptic membrane, and is visible under the electron microscope as tight complexes of post-synaptic junctional proteins (figure 3). It is well established that the regulation of synaptic proteins that occur in the PSP plays an important role in synaptic plasticity, a mechanism which is believed to underlie the synaptic mechanism of learning and memory (Kennedy 1993).

Appendix 9

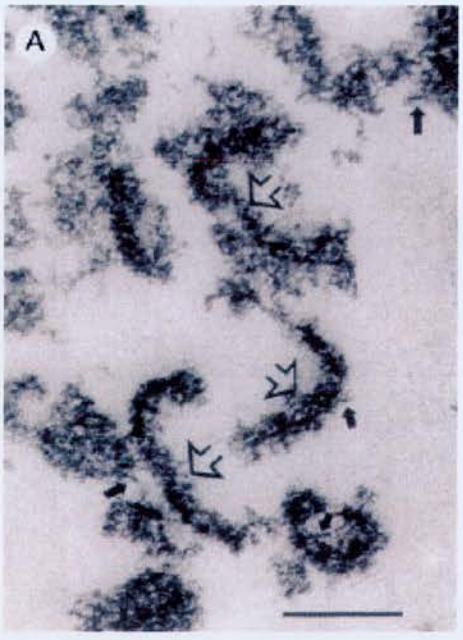


Figure 3. PSDs appear as slightly curved bars (open arrows indicate the cytoplasmic side). Adhering postsynaptic membranes are absent. Although little substructure is apparent, occaisional rod-like elements (4.5 x 28nm, small arrows) can be seen. (Image from Ziff, E.B. Enlightening the postsynaptic density. Neuron 19, 1163-1174 (1997))

The PSP was purified from mouse forebrain by density centrifugation as described previously (Carlin, Grab et al. 1980) with minor modifications. Protein samples were separated by SDS-PAGE, stained by Coomassie blue, and the entire gel lane was excised into 42 individual protein bands, reduced, alkylated and digested with trypsin . The resultant peptides were identified by online LC-MS/MS analysis and corresponded to 620 unique proteins. This approach has far exceeded past studies, one of which only identified 26 molecules in the post-synaptic density (Walikonis, Jensen et al. 2000). Another study, which employed a 2D electrophoresis approach, resolved 1700 spots. MS identified a total of 90 spots, containing 47 different protein species. Unsurprisingly, within the dataset reported by Husi et al, were many previously described synaptic proteins including ion channels, scaffold proteins, signaling enzymes and cytoskeletal proteins. The majority of the proteins were not known to be postsynaptic proteins. Although this proteomic strategy attempted to use a relatively unbiased strategy, it cannot be completely comprehensive as low abundance proteins may escape detection, and the selection of whole forebrain as starting material will overlook molecules found in small subsets of neurons. A further limitation imposed by the method of biochemical isolation is that some proteins may be contaminants not normally found at synapses. Despite this concern they

observed a significant number of known synaptic proteins and a lack of many nuclear and other proteins that would not be expected. A systematic study using electron microscopy would be necessary to confirm or exclude each protein's synaptic localization.

This postsynaptic protein profile is useful in several ways. First, it appears that many of these proteins appear to be excellent candidates for various aspects of cell biology at the postsynaptic terminal. For example, 24 components of the protein synthetic machinery (elongation factors, ribosomal proteins) were identified; proteins encoded by synaptically localized mRNAs; cytoskeletal components and modulators of the cytoskeleton; motor proteins involved with ion channel trafficking channels (e.g. KIF proteins, NSF, AMPA); ubiquitination mechanisms and so on. Since many of these proteins have well characterized functions in non-neuronal settings, their identification may facilitate tests of their function at synapses. A second insight relates to the overall organization of the postsynaptic terminal into sets of macromolecular complexes. In addition to components of NMDA receptor complexes and AMPA complexes, there are ribosomal and other complexes, consistent with the view that the postsynaptic terminal is organized into macromolecular machines. Third, a number of proteins are involved with neurological diseases and this list may prove useful in directing human genetic studies of synaptic pathology. In addition to its use in understanding the biology of synapses, this dataset may prove useful in guiding human genetic studies of the nervous system.

The approach of biochemical purification combined with high-resolution mass spectrometry used for the definition of the PSP could be easily applied to many other sub-cellular structures. Indeed, subcellular structures, such as axo-glial junctions and nodes of Ranvier would be prime candidates for such an approach. Also, refined or new biochemical approaches to isolate structures such as dendrites, axons and growth cones would facilitate definition of other neuronal proteomes.

2.3 Neuro-phosphoproteomics

Another global neuroproteomic study carried out by Collins et al, (unpublished results) focused on the neurophosphoproteome, i.e. the collection of phosphoproteins at the synapse. Protein phosphorylation is an essential regulator of protein function and signaling mechanisms in any biological system. It has been estimated that one third of all proteins are phosphorylated in the mammalian cell and if it is assumed that one third of all proteins are expressed in the brain, then a conservative estimate of the number of phosphoproteins expressed in the brain would be in the order of 2000-3000. This is underpinned by the fact that up to 5% of the expressed genome may code for kinases.

Appendix 9

Protein phosphorylation has been shown to be intimately involved in the assembly and function of the post-synaptic density (Yamauchi 2002), synaptic vesicle turnover (Turner, Burgoyne et al. 1999), LTP (Hedou and Mansuy 2003) and synaptic plasticity/memory (Schulman 1995). It is therefore obvious that study of protein phosphorylation on a proteomic platform will yield a wealth of data relevant to the molecular basis of brain function.

Identification of phosphoproteins and characterisation of phosphorylation sites has proved difficult for many reasons, most notably due to the low stochiometry and transient nature of protein phosphorylation. This has been especially true for global analysis of phosphorylation and there was a need for specific enrichment techniques to be developed to tackle these issues.

A large-scale phosphoprotein isolation strategy was developed using a pre-existing methodology for phosphopeptide purification. Phosphoproteins were purified by large-scale metal affinity chromatography (Andersson and Porath 1986)' (Muszynska, Andersson et al. 1986) from denatured synaptosomes which contain the pre- and post-synaptic terminals, including the postsynaptic density. Purified phosphoproteins were digested in solution and analysed by DRA-LC-MS-MS. This approach has yielded the largest mammalian proteomic dataset to date, identifying 236 potential phosphoproteins from many protein classes (channels, transporters/ transporting ATPases, scaffolding and adaptor proteins, kinases, phosphatases, small G-proteins and modulators, other signaling molecules, metabolic enzymes, cell adhesion and cytoskeletal proteins, nuclear/transcriptional, ribosomal and novel proteins. 28 phosphopeptides were detected and characterised and contained 37 phosphorylated residues. Of these 236 putative phosphoproteins it was shown by extensive literature mining that 81 of these had been shown previously to be phosphorylated, thus bringing confidence to the purification strategy used.

This dataset has considerable overlap with the post-synaptic density proteome, with one hundred common proteins (Husi 2003), fifty-four of which, are known to be phosphorylated. Also, thirty-one proteins contained in the MAGUK-associated signaling complex (MASC) (Husi 2003) are found in this phosphoproteomic dataset with twenty-four of these proteins known to be phosphorylated. This phosphoproteomic approach is being optimised in a number of ways, from using a double IMAC strategy, i.e. sequential purification of phosphoproteins and then phosphopeptides, to using more refined MS techniques such as precursor-ion scanning. These changes should yield much better coverage and depth to the synaptic phosphoproteome. Integration of this data on dynamic regulation of proteins into phosphorylation networks (figure 4), and in also into the PSP network will add an extra dimension and understanding to the organisation and dynamism of the post-synaptic density.

Appendix 9

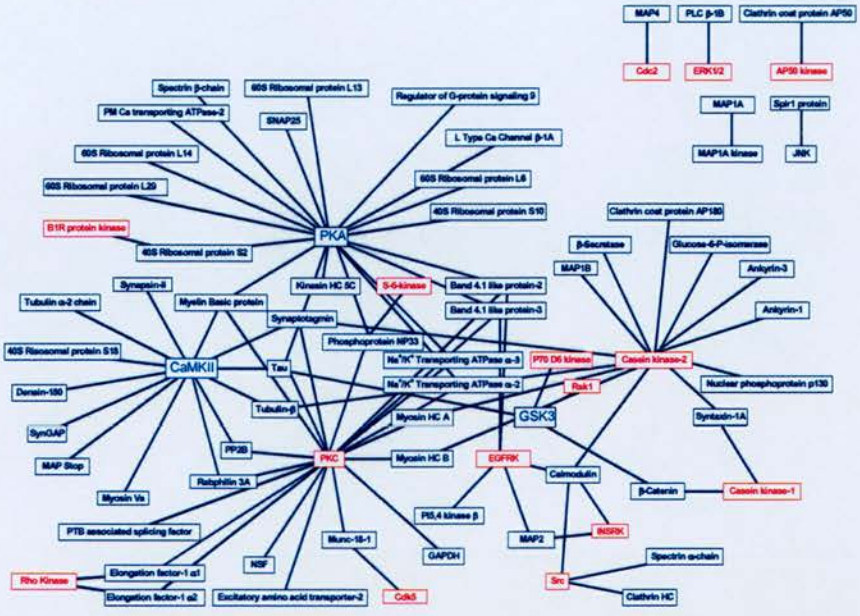


Figure 4. Literature based kinase substrate interaction map of proteins identified in the phosphoproteomic dataset by Collins et al. Proteins shown in red were not detected but included to show interactions with their substrates within the dataset.

The next major challenge for phosphoproteomics will be identification of cognate kinases and phosphatases for the potential wealth of phosphorylation sites which will arise from large-scale phosphorylation site studies. Also, other functional aspects of protein phosphorylation will need to be addressed in a high throughput fashion, namely the effects of phosphorylation on protein-protein interactions and protein stability. The development of high-density peptide and protein arrays (Lesaicherre, Uttamchandani et al. 2002) should facilitate the study of some of these aspects of the functional consequences of protein phosphorylation in a large-scale manner.

Appendix 9

3. MULTIPROTEIN COMPLEXES: A PROTEOMIC APPROACH

3.1 NMDA receptor-adhesion protein signaling complex

Protein-protein interactions are the basis on which the cellular and structure and function are built and interaction partners are an immediate lead into biological function which can be exploited for therapeutic purposes. Proteomics has been shown to make a crucial contribution to the study of protein-protein interactions (Neubauer, Gottschalk et al. 1997). Proteomic approaches to tackle multiprotein complexes usually involve purification of the entire complex by a variety of affinity methods, and protein identification by western blotting or mass spectrometry based approaches. The first proteomic analysis of multiprotein complexes relevant to the brain was the purification and identification of the molecular constituents of the NMDA receptor-adhesion protein signaling complexes (NRC) (Husi and Grant 2001) (Husi, Ward et al. 2000) and this will be described as a prototypic approach to multiprotein complex proteomics.

3.1.1 Methodology

Similar to most neurotransmitter receptors, the NMDAR has a low abundance in neurones compared with the overall protein content. Therefore, it is important to generate highly enriched samples containing the target molecule. This can be achieved by a variety of methods, generally involving affinity chromatography steps. These can range from ligand affinity (drugs, substrates or co-factors) to the more widely used approach of immunoprecipitation, involving specific antibodies to the molecule of interest. A major drawback in using antibody-based approaches for proteomic analysis is the introduction of the IgG itself, which interferes with the analysis by overshadowing proteins of interest that co-migrate during the gel-separation. This same problem arises with protein-affinity based methods, in which a recombinant protein carries one or several affinity-tags (e.g. poly-histidines), or is expressed as a GST-fusion protein. This problem can be eliminated using either the tandem affinity purification (TAP) method (Rigaut, Shevchenko et al. 1999) or with the use of non-protein based methods, which exploit substrate or ligand affinities, such as peptide-based isolation procedures (Husi and Grant 2001). All of these ligands, groups or proteins are then coupled to a resin either irreversibly (usually in the case of small molecules), or through a secondary affinity linkage specific for the bait-molecule. Neuronal extracts are then incubated with the resins bearing

Appendix 9

the bait and bound proteins separated from each other as well as from the matrix by conventional gel analysis.

Identification of isolated proteins can be readily achieved using western blotting or mass spectrometry. Western blotting (also known as immunoblotting) utilizes specific antibodies to detect a protein that has been transferred from a gel onto a membrane. A major advantage of western blotting is its sensitivity and ability to detect amounts of protein beyond the range of current sequencing-based or mass spectrometry methodologies. It is obvious that this approach is biased by the assumption that a given molecule might be present, and cannot be used to identify unknown or unsuspected proteins. However, there is a large amount of neurobiological evidence concerning the putative involvement of molecules that influence many target proteins and such proteins pose ideal targets for western blotting approaches. Additionally, changes in protein levels, as a result of modulation or modification in the receptor's environment, can easily be visualized using specific antibodies against the proteins under investigation. A disadvantage of this method is its labour intensive nature and thus difficulty in performing on a large scale.

The main approaches to protein identification are peptide-mass fingerprinting using matrix-assisted laser desorption/ionization (MALDI) (Refs [(Henzel, Billeci et al. 1993), (Shevchenko, Wilm et al. 1996)) and electrospray ionization using a tandem mass spectrometer (Shevchenko, Chernushevich et al. 1997) or a combination of both (Shevchenko, Loboda et al. 2000). Recent developments led to new methodologies using direct liquid chromatography/tandem mass spectrometry (LC/MS-MS) techniques, which allow fast and reliable mapping of very low amounts of peptides in an almost completely automated fashion (Shevchenko, Chernushevich et al. 2000). Furthermore, using this method, it is also possible to determine if a protein or peptide carries any post-translational modifications, such as glycosylation or phosphorylation (Pandey, Podtelejnikov et al. 2000).

The sample is usually separated on a gel and the bands of interest excised and treated with trypsin in order to generate fragments of proteins that can subsequently be extracted from the gel slice. There are several reasons for analyzing peptides rather than proteins. The peptide's mass can be measured at much greater precision than the mass of intact proteins. Peptides give rise to a library of molecular masses that are directly derived from the proteins, which can be identified using computational techniques (Berndt, Hobohm et al. 1999).

Using the approaches described above the NRC was shown to comprise 77 proteins organized into receptor, adaptor, signaling, cytoskeletal and novel proteins, of which 30 are implicated from binding studies and another 19 participate in NMDAR signaling (Figure 5) (Husi, Ward et al. 2000).

Appendix 9

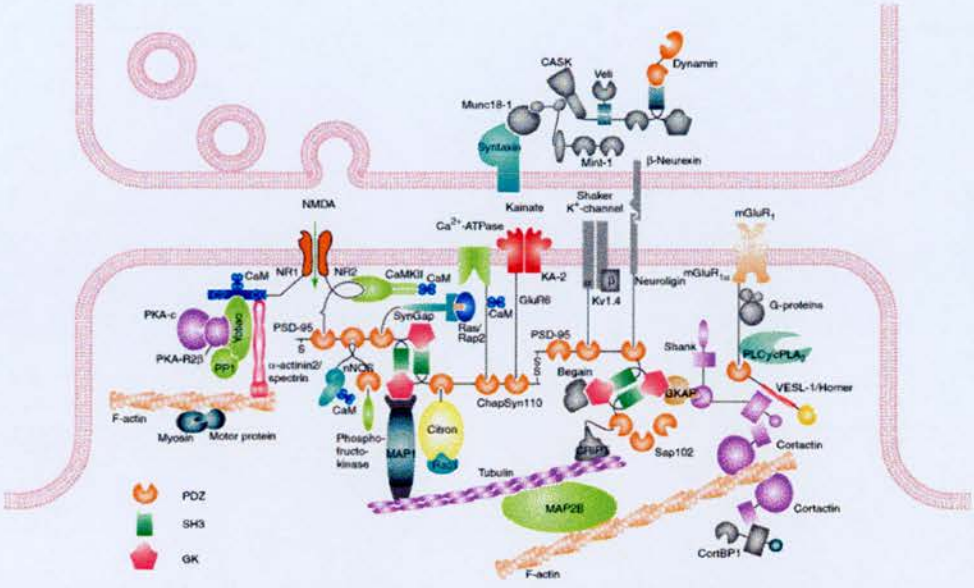


Figure 5. Schematic diagram of synaptic multiprotein complexes. Post-synaptic complexes of proteins associated with the NMDA receptor and PSD-95, found at excitatory mammalian synapses, are shown. Individual proteins are illustrated with arbitrary shapes and known interactions indicated. Proteins shown in colour are those found in a proteomic screen, whereas those shown in grey are inferred from bioinformatic studies. The specific protein–protein interactions are predicted, based on published reports from yeast two-hybrid studies. Membrane proteins (such as receptors, channels and adhesion molecules) are attached to a network of intracellular scaffold, signaling and cytoskeletal proteins, as indicated.

NMDAR and metabotropic glutamate receptor subtypes were linked to cadherins and L1 cell-adhesion molecules in complexes lacking AMPA receptors. These neurotransmitter-adhesion receptor complexes were bound to kinases, phosphatases, GTPase-activating proteins and Ras with effectors including MAPK pathway components. A striking feature of the composition is that there appears to be 'modules' or sets of signaling proteins that are known to comprise key components of signal transduction pathways that can be distinctly regulated. For example, all of the molecules necessary to induce the phosphorylation of mitogen-activated protein kinase (MAPK) following NMDAR stimulation are present in the NRC (including CaMKII, SynGAP, Ras, MEK and ERK) (Figure 5). There modules may allow the NMDAR and mGluR to integrate signals within the complex and then couple to down stream cellular effector mechanisms, such as trafficking

of AMPA receptors and cytoskeletal changes that mediate structural and physiological plasticity. Furthermore, at least 18 NRC constituents are regulated by synaptic activity indicating that the composition of the complex is dynamic. In the hippocampus, these activity-dependent genes are known to undergo specific temporal changes following the induction of plasticity.

Genetic or pharmacological interference with 15 NRC proteins impairs learning and with 22 proteins alters synaptic plasticity in rodents. Mutations in three human genes (NF1, Rsk-2, L1) are associated with learning impairments, indicating the NRC also participates in human cognition.

This NRC dataset was further expanded using a combination of peptide affinity chromatography and mass spectrometry identification of new components of the complex (Husi H et al 2003). Purification of MASCs (MAGUK-associated signaling complexes) was achieved by peptide affinity chromatography using a hexapeptide corresponding to the C-terminus of the NMDA-R2B subunit. This approach extended the NRC dataset to comprise 184 proteins which constitute the MASC(Husi 2003). Also, α-amino-3-hyroxy-5-methyklisoxazole-4-proprionic acid (AMPA) complexes were purified using immuno-affinity chromatography, resulting in the identification of ten unique proteins. These two receptor complex datasets, together with the PSD provides for the first time a general composition of the postsynaptic proteome, which totaled 698 distinct proteins. The main overlap was seen between MASC and the PSD, with 108 common proteins.

3.2 Other Membrane Receptor Complexes

Since the characterisation of the NRC, two other receptor complexes have been reported. P2X receptors are ATP-gated ion channels in the plasma membrane, and activation of the P2X7 receptor also leads to rapid cytoskeletal re-arrangements such as membrane blebbing. 11 proteins in human embryonic kidney cells that interact with the rat P2X7 receptor were identified by affinity purification followed by mass spectroscopy and immunoblotting (Kim, Jiang et al. 2001). Also using a proteomic approach based on peptide affinity chromatography followed by mass spectrometry and immunoblotting, another study identified 15 proteins that interact with the C- terminal tail of the 5-hydroxytryptamine 2C (5-HT (2C)) receptor, a GPCR (Becamel, Alonso et al. 2002). These proteins include several synaptic multidomain proteins containing one or several PDZ domains (PSD-95 and the proteins of the tripartite complex Veli3-CASK-Mint1), proteins of the actin/spectrin cytoskeleton and signaling proteins. Co-immunoprecipitation experiments showed that 5-HT (2C) receptors interact with PSD-95 and the Veli3-CASK- Mint1 complex in vivo. Electron microscopy also indicated a synaptic enrichment of Veli3 and 5-HT (2C) receptors and their co-localisation in microvilli of choroidal cells. These

results indicate that the 5-HT (2C) receptor is associated with protein networks that are important for its synaptic localization and its coupling to the signaling machinery (Becamel, Alonso et al. 2002).

It is now emerging that multiprotein signaling complexes or "signaling machines", of which the NRC is prototypic, are responsible for orchestrating the complex signaling events at ionotrophic glutamate receptors (Husi, Ward et al. 2000), ATP receptors (Kim, Jiang et al. 2001) and G-protein coupled receptors (Becamel, Alonso et al. 2002). They seem to share a common organisation of receptor, protein scaffold, (in which the signaling molecules are localised) and membrane to cytoskeletal interactions. Therefore, study of individual receptor multiprotein complexes such as the NRC will most likely provide a mechanistic basis that will be very useful and applicable for the systematic identification and study of membrane receptor complexes associated with the estimated 100 other receptor types believed to exist in the brain.

4. DISEASE PROTEOMICS OF THE BRAIN

4.1 Alzheimer's Disease as a case study

The main approach of proteomic analysis of brain pathology is to use comparative 2D-page separation of constituent proteins in diseased versus normal samples. This is usually followed by gel image analysis to identify alterations in the intensity of specific spots compared to the normal state and subsequent excision of the abnormal spots and identification by mass spectrometry or peptide sequencing. The main types of samples, which have been analysed by these means, are human biopsies or postmortem samples and mouse models of human disease.

There have been a number of studies focused on Alzheimer's disease (AD), an increasingly prevalent neurodegenerative disease. Comparative proteome analysis was performed on post-mortem brain tissue samples from patients with AD and compared with profiles from non-demented control brain tissue (Schonberger, Edgar et al. 2001). Proteins were resolved by 2D-PAGE and identified by NH_2-terminal sequencing. Thirty-seven proteins differed significantly between AD and normal tissue and could be grouped into several functional categories. A number of proteins identified in this study, such as α–crystallin, superoxide dismutase, glyceraldehyde-3-phosphate dehydrogenase, and dihydropyrimidinase-related protein had been implicated previously in the pathogenesis of AD.

Appendix 9

This work has been complemented by studies of mouse models of AD, namely mice transgenically overexpressing glycogen synthase kinase-3β (GSK-3β) and microtubule-associated protein tau, both of which have been implicated in the pathogenesis of AD. A proteomic approach was used to identify cellular abnormalities in mice over expressing human tau with the aim of understanding the role of tau in the pathogenesis of AD. Proteins were resolved by 2D-PAGE and 34 proteins whose expression levels in wild-type and tau transgenic mice differed at least 1.5 fold were identified by ESI-MS (Tilleman, Van den Haute et al. 2002). A similar study by the same group was performed on GSK-3β overexpressing mice (Tilleman, Stevens et al. 2002). GSK-3β is a serine-threonine protein kinase capable of phosphorylating tau and hyperphosphorylated tau is a principle component of neurofibrillary tangles, which are characteristic of AD brains. It is regulated by serine (Inhibitory) and tyrosine (stimulatory) phosphorylation and also by protein complex formation and its intracellular localisation. GSK-3β is a key regulator of regulating neuronal plasticity and gene expression controlling the function of many metabolic, signaling and structural proteins. It also regulates cell survival, as it facilitates a variety of proapoptotic mechanisms (Harwood 2001). It has been linked to all of the primary abnormalities associated with AD and was shown to interact with components of the plaque producing amyloid system (Grimes and Jope 2001).

A comparison between GSK-3β transgenic mice and wild type revealed 51 proteins whose expression differed significantly. Twenty-five of these proteins have been identified to be present in the postsynaptic proteome, 14 of which are components of NMDAR-PSD-95 complexes that constitute MASCs.

There was a significant decrease in the relative abundance of cytoskeletal and energy metabolism proteins and a significant increase in proteins involved in signal transduction and oxidative stress in the GSK-3β transgenic mice. A comparison of the profile of proteins with altered expression in the tau and GSK-3β transgenic mice showed that there were a number of proteins with similar changes consistent with the mutual involvement of these two proteins in the process of neurodegeneration.

Complementary large-scale approaches to the study of AD and other brain diseases have been developed, notably cDNA microarrays. These arrays can detect thousands of mRNA transcripts present in the biological sample and can be used in a similar way as 2D-PAGE for detecting alterations in diseased versus normal tissues. The use of microarrays is now routine and due to the facile methodology can be used as a first step in detecting alterations in transcript levels in various pathologies. However, the correlation between mRNA levels and protein levels can be poor, so microarray data should be interpreted with caution and confirmed by looking

at protein levels by methods previously discussed. A recent example of the use of cDNA microarrays for the study of AD was reported by Colangelo et al. in which they profiled 12633 genes in human post-mortem hippocampal CA1 in normal versus AD patients (Colangelo, Schurr et al. 2002). They found decreases in transcript levels of 19 genes encoding metal ion-sensitive factors, transcription and neurotrophic factors, and signal-transduction elements involved in synaptic plasticity and organisation which correlated well with previously reported down-regulation of such function categories in AD. Also, they found increases in transcript levels of 19 genes involved in proapoptotic activities and inflammation, also consistent with AD pathology. Interpretation and integration of such data with that obtained from proteomic experiments is presently difficult as specific changes in transcripts and protein levels are not seen in many consistent sets of proteins/genes. Also, expression levels (mRNA/protein) seen in normal versus diseased experiments can be due to secondary effects and not central to the pathogenesis. However, consolidation of mRNA/protein expression data for a particular disease in a database, may uncover common disturbances in pathways relevant to the actual pathology and not due to secondary effects.

4.2 Networks and psychiatric disease

Complex diseases are ideally suited for study by proteomic approaches. Their complexity of causation requires a global approach and an example of one such study relates to the work described earlier on MASC.

The identification of MASC and its network properties provides a unique opportunity to explore human cognition and its disorders. Although it is generally accepted that cognitive mechanisms are conserved between mice and humans, it is unclear how much the rodent molecular studies map onto human psychiatric conditions. The possibility that MASC proteins may be involved with human psychiatric and neurological disorders was investigated and it was found that 46 MASC proteins implicated in mental illness in the literature (figure 6). Although all mental disorders were searched 26 were found in schizophrenia, 18 in mental retardation, 7 in Bipolar disorder and 6 in depressive illness. This apparent bias toward schizophrenia and mental retardation could be biologically relevant since they both have a major cognitive component to their primary symptoms unlike the affective disorders.

Appendix 9

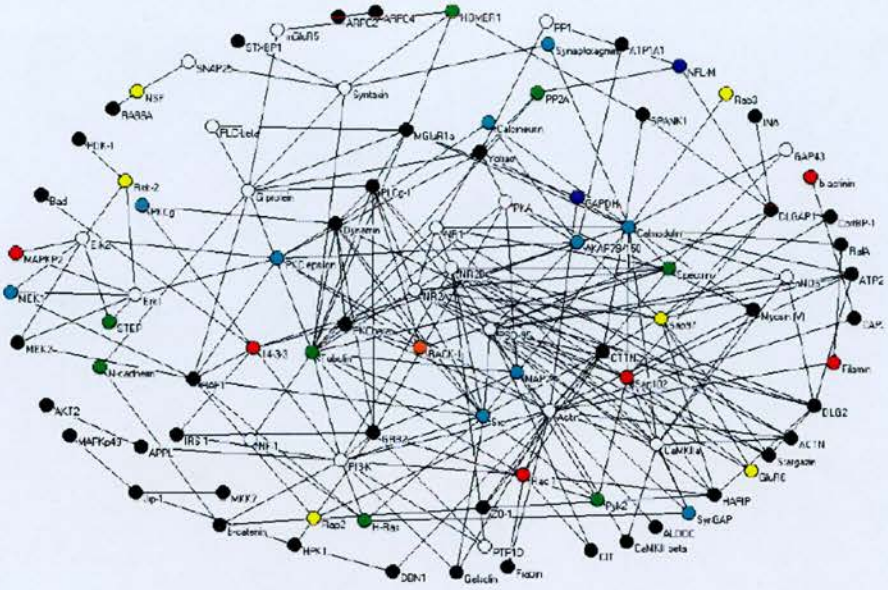


Figure 6. Network representation of MASC proteins with phenotypic annotation. 97 MASC proteins with known direct protein interactions to other MASC proteins are plotted. The NMDA receptor subunits (NR1, NR2A, NR2B) are centrally located. The association of each protein with plasticity, rodent behaviour or human psychiatric disorders is shown in the colour of the node. Key: Black – no known association; Red – Psychiatric Disorder; Green – Plasticity; Blue – Rodent Behaviour; Yellow – Psychiatric Disorders and Plasticity; Cyan – Plasticity and Rodent Behaviour; White – all three phenotypes. Orange for Psychiatric Disorders and Rodent Behaviour alone had no hits.

The network provides a common connection or association between these disease molecules. Network simulations also show that disruption of combinations of proteins produce more severe effects on the network. This suggests that combinations of proteins (or mutant alleles) underpin the polygenic nature of the disorder.

4.3 Future perspectives

The number of comparative 2D/MS studies of disease and disease models has increased over the last number of years. This has led to the accumulation of lists of proteins whose expression levels have been shown to be altered in the particular system, but with little integration of these lists to provide a global view. This is emerging problem in the study of complex diseases, not only at the level of proteomics but also at the level of study of individual protein function in relation to disease. The pathogenesis of neurodegenerative diseases such as AD is likely to consist of a number of initial pathways converging to form the distinctive pathology of amyloid

Appendix 9

plaques and NFT's (neurofibrillary tangles) characteristic of AD. Proteins implicated in this pathogenesis may also be characteristic of the stage of progression of the disease and may be transiently altered. The components and formation of amyloid plaques and NFT's are reasonably well understood but the many pathways leading to this end point are still elusive. It is clear that there is a need to integrate the large amount of data concerning proteins involved in the pathogenesis of AD into a coherent global view of the disease.

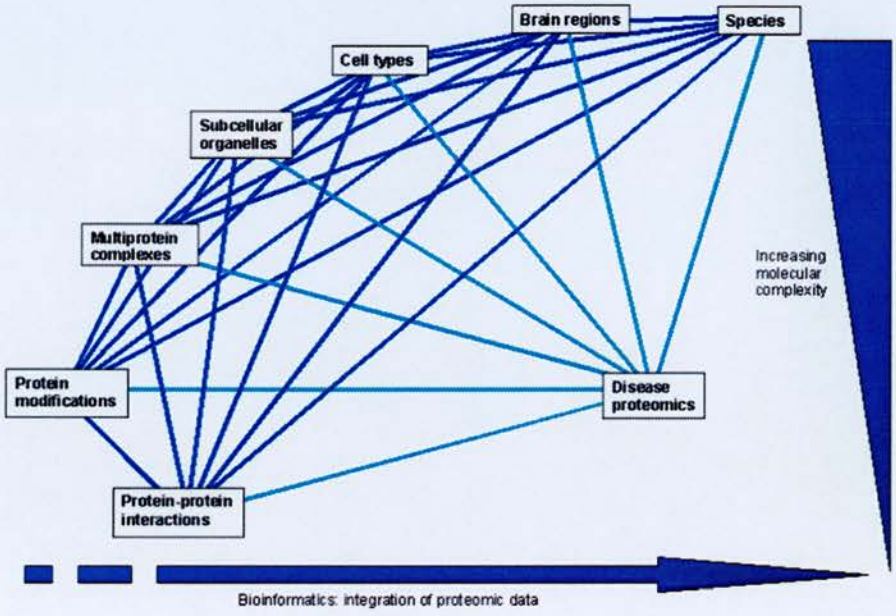


Figure 7. Overview of levels of proteomic complexity

It may be quite likely that an understanding of the fundamental components and mechanisms of the pathways leading to AD pathology will be gained from a holistic view that can be obtained by using a proteomic strategy combined with bioinformatics and network analysis (figure 7).

Appendix 9

5. BIOINFORMATICS: INTEGRATION OF PROTEOMIC DATA

It is evident that there is a growing need for bioinformatic tools and databases to deal with the wealth of data being generated by proteomic analyses. Increased image resolution and reproducibility in 2D-electrophoresis is allowing 2D maps to be generated for particular systems (http://us.expasy.org/ch2d/). The samples and the protocols used to generate these maps are being standardised to allow easy cross-referencing between gels. Comprehensive protein identification by MS allows assignment of protein spots on these master maps yielding a characteristic spot position for each protein species. This kind of approach, in which a proteome is defined on a 2D map, will allow rapid comparative analysis. This will be particularly useful in disease proteomics, as it will bypass the necessity for routine MS.

Integrative proteomic databases such as PADB (http://www.PPID.org) (Husi 2002) will be essential for the analysis of proteomic data. The ability to submit MS datasets to such databases and retrieve comprehensive protein information, including protein-protein interactions will greatly reduce the current bottleneck of manual analysis of large numbers of proteins. As discussed for the MASC and PSP, such assembly of proteins into networks has been facilitated by this tailored proteomic database and such insights into global basic biology would not be possible in its absence. Bioinformatic analysis of proteomes is in its infancy, it can be seen in figure 1 that the future objectives of the many proteomes that exist are ambitious, but necessary.

Proteomic data must be dealt with in a holistic way if any meaningful biological questions are to be answered. Basic information on protein-protein interactions, protein modifications and multiprotein complexes must be obtained for relevant subcellular structures in various cell types and brain regions and integrated with previously published functional information. Within this framework of biological knowledge, complex biological problems and pathologies can be addressed in a less biased and comprehensive way than ever possible. Integration of diverse proteomic data to provide the ultimate global view is a formidable long-term task, which will only be achieved by standardisation of data, very powerful bioinformatic tools and comprehensive databases.

Appendix 9

REFERENCES

Aebersold, R. and D. R. Goodlett (2001). "Mass spectrometry in proteomics." Chem Rev **101**(2): 269-95.

Andersson, L. and J. Porath (1986). "Isolation of phosphoproteins by immobilized metal (Fe3+) affinity chromatography." Anal Biochem **154**(1): 250-4.

Becamel, C., G. Alonso, et al. (2002). "Synaptic multiprotein complexes associated with 5-HT(2C) receptors: a proteomic approach." Embo J **21**(10): 2332-42.

Beranova-Giorgianni, S., F. Giorgianni, et al. (2002). "Analysis of the proteome in the human pituitary." Proteomics **2**(5): 534-42.

Berndt, P., U. Hobohm, et al. (1999). "Reliable automatic protein identification from matrix-assisted laser desorption/ionization mass spectrometric peptide fingerprints." Electrophoresis **20**(18): 3521-6. [pii].

Carlin, R. K., D. J. Grab, et al. (1980). "Isolation and characterization of postsynaptic densities from various brain regions: enrichment of different types of postsynaptic densities." J Cell Biol **86**(3): 831-45.

Colangelo, V., J. Schurr, et al. (2002). "Gene expression profiling of 12633 genes in Alzheimer hippocampal CA1: transcription and neurotrophic factor down-regulation and up-regulation of apoptotic and pro-inflammatory signaling." J Neurosci Res **70**(3): 462-73.

Fountoulakis, M., E. Schuller, et al. (1999). "Rat brain proteins: two-dimensional protein database and variations in the expression level." Electrophoresis **20**(18): 3572-9. [pii].

Friso, G. and L. Wikstrom (1999). "Analysis of proteins from membrane-enriched cerebellar preparations by two-dimensional gel electrophoresis and mass spectrometry." Electrophoresis **20**(4-5): 917-27. [pii].

Gauss, C., M. Kalkum, et al. (1999). "Analysis of the mouse proteome. (I) Brain proteins: separation by two- dimensional electrophoresis and identification by mass spectrometry and genetic variation." Electrophoresis **20**(3): 575-600.

Grimes, C. A. and R. S. Jope (2001). "The multifaceted roles of glycogen synthase kinase 3beta in cellular signaling." Prog Neurobiol **65**(4): 391-426.

Harwood, A. J. (2001). "Regulation of GSK-3: a cellular multiprocessor." Cell **105**(7): 821-4.

Hedou, G. and I. M. Mansuy (2003). "Inducible molecular switches for the study of long-term potentiation." Philos Trans R Soc Lond B Biol Sci **358**(1432): 797-804.

Henzel, W. J., T. M. Billeci, et al. (1993). "Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases." Proc Natl Acad Sci U S A **90**(11): 5011-5.

Husi, H., and Grant, S. G. (2002). "Construction of a Protein-Protein Interaction Database (PPID) for Synaptic Biology." In Neuroscience Databases: A practical Guide (Boston/Dordrecht/London, Kluwer Academic Publishers): pp. 51-62.

Appendix 9

Husi, H., Choudhary, J., Yu, L., Cumiskey, M., Blackstock, W., O'Dell, T.J., Visscher, P.M., Armstrong, J.D. & Grant, S.G.N. (2003). "Synapse proteomes show scale-free network properties underlying plasticity, cognition and mental illness." Submitted.

Husi, H. and S. G. Grant (2001). "Isolation of 2000-kDa complexes of N-methyl-D-aspartate receptor and postsynaptic density 95 from mouse brain." J Neurochem **77**(1): 281-91.

Husi, H., M. A. Ward, et al. (2000). "Proteomic analysis of NMDA receptor-adhesion protein signaling complexes." Nat Neurosci **3**(7): 661-9.

Kennedy, M. B. (1993). "The postsynaptic density." Curr Opin Neurobiol **3**(5): 732-7.

Kim, M., L. H. Jiang, et al. (2001). "Proteomic and functional evidence for a P2X7 receptor signalling complex." Embo J **20**(22): 6347-58.

Langen, H., P. Berndt, et al. (1999). "Two-dimensional map of human brain proteins." Electrophoresis **20**(4-5): 907-16. [pii].

Lesaicherre, M. L., M. Uttamchandani, et al. (2002). "Antibody-based fluorescence detection of kinase activity on a peptide array." Bioorg Med Chem Lett **12**(16): 2085-8.

Muszynska, G., L. Andersson, et al. (1986). "Selective adsorption of phosphoproteins on gel-immobilized ferric chelate." Biochemistry **25**(22): 6850-3.

Neubauer, G., A. Gottschalk, et al. (1997). "Identification of the proteins of the yeast U1 small nuclear ribonucleoprotein complex by mass spectrometry." Proc Natl Acad Sci U S A **94**(2): 385-90.

Pandey, A., A. V. Podtelejnikov, et al. (2000). "Analysis of receptor signaling pathways by mass spectrometry: identification of vav-2 as a substrate of the epidermal and platelet- derived growth factor receptors." Proc Natl Acad Sci U S A **97**(1): 179-84.

Rigaut, G., A. Shevchenko, et al. (1999). "A generic protein purification method for protein complex characterization and proteome exploration." Nat Biotechnol **17**(10): 1030-2.

Santoni, V., M. Molloy, et al. (2000). "Membrane proteins and proteomics: un amour impossible?" Electrophoresis **21**(6): 1054-70. [pii].

Schonberger, S. J., P. F. Edgar, et al. (2001). "Proteomic analysis of the brain in Alzheimer's disease: molecular phenotype of a complex disease process." Proteomics **1**(12): 1519-28.

Schulman, H. (1995). "Protein phosphorylation in neuronal plasticity and gene expression." Curr Opin Neurobiol **5**(3): 375-81.

Shevchenko, A., I. Chernushevich, et al. (1997). "Rapid 'de novo' peptide sequencing by a combination of nanoelectrospray, isotopic labeling and a quadrupole/time-of-flight mass spectrometer." Rapid Commun Mass Spectrom **11**(9): 1015-24.

Shevchenko, A., I. Chernushevich, et al. (2000). "De Novo peptide sequencing by nanoelectrospray tandem mass spectrometry using triple quadrupole and quadrupole/time-of-flight instruments." Methods Mol Biol **146**: 1-16.

Shevchenko, A., A. Loboda, et al. (2000). "MALDI quadrupole time-of-flight mass spectrometry: a powerful tool for proteomic research." Anal Chem **72**(9): 2132-41.

Shevchenko, A., M. Wilm, et al. (1996). "A strategy for identifying gel-separated proteins in sequence databases by MS alone." Biochem Soc Trans **24**(3): 893-6.

Taoka, M., A. Wakamiya, et al. (2000). "Protein profiling of rat cerebella during development." Electrophoresis **21**(9): 1872-9. [pii].

Appendix 9

Tilleman, K., I. Stevens, et al. (2002). "Differential expression of brain proteins in glycogen synthase kinase-3 transgenic mice: a proteomics point of view." Proteomics **2**(1): 94-104.

Tilleman, K., C. Van den Haute, et al. (2002). "Proteomics analysis of the neurodegeneration in the brain of tau transgenic mice." Proteomics **2**(6): 656-65.

Turner, K. M., R. D. Burgoyne, et al. (1999). "Protein phosphorylation and the regulation of synaptic membrane traffic." Trends Neurosci **22**(10): 459-64.

Walikonis, R. S., O. N. Jensen, et al. (2000). "Identification of proteins in the postsynaptic density fraction by mass spectrometry." J Neurosci **20**(11): 4069-80.

Wu, C. C., M. J. MacCoss, et al. (2003). "A method for the comprehensive proteomic analysis of membrane proteins." Nat Biotechnol **21**(5): 532-8.

Yamauchi, T. (2002). "Molecular constituents and phosphorylation-dependent regulation of the post-synaptic density." Mass Spectrom Rev **21**(4): 266-86.

Appendix 10

PROTOCOL

Robust Enrichment of Phosphorylated Species in Complex Mixtures by Sequential Protein and Peptide Metal-Affinity Chromatography and Analysis by Tandem Mass Spectrometry

Mark O. Collins,[1,2*] Lu Yu,[1] Holger Husi,[2] Walter P. Blackstock,[3] Jyoti S. Choudhary,[1] Seth G. N. Grant[1,2*]

(Published 23 August 2005)

[1]The Wellcome Trust Sanger Institute, Hinxton, CB10 1SA, UK. [2]Division of Neuroscience, University of Edinburgh, Edinburgh, EH8 9JE, UK. [3]Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield, S10 2TN, UK.

* Corresponding authors. E-mail, moc@sanger.ac.uk (M.O.C.); sg3@sanger.ac.uk (S.G.N.G.)

Appendix 10

PROTOCOL

Abstract

Reversible protein phosphorylation mediated by kinases, phosphatases, and regulatory molecules is an essential mechanism of signal transduction in living cells. Although phosphorylation is the most intensively studied of the several hundred known posttranslational modifications on proteins, until recently the rate of identification of phosphorylation sites has remained low. The use of tandem mass spectrometry has greatly accelerated the identification of phosphorylation sites, although progress was limited by difficulties in phosphoresidue enrichment techniques. We have improved upon existing immobilized metal-affinity chromatography (IMAC) techniques for capturing phosphopeptides, to selectively purify phosphoproteins from complex mixtures. Combinations of phosphoprotein and phosphopeptide enrichment were more effective than current single phosphopeptide purification approaches. We have also implemented iterative mass spectrometry-based scanning techniques to improve detection of phosphorylated peptides in these enriched samples. Here, we provide detailed instructions for implementing and validating these methods together with analysis by tandem mass spectrometry for the study of phosphorylation at the mammalian synapse. This strategy should be widely applicable to the characterization of protein phosphorylation in diverse tissues, organelles, and in cell culture.

Introduction

Protein phosphorylation is an essential regulator of protein function, and signaling pathways rely on this mechanism to regulate protein activity and protein-protein interactions. Phosphorylation can function as a reversible tag of activity by which information can be transduced from one protein to another in response to an appropriate stimulus. Alternatively, phosphorylation can serve to promote specific protein-protein interactions mediated by protein domains that recognize phosphorylated amino acid-containing sequences. Signal transduction is fundamentally important to many cellular processes and, consequently, is highly complex and intricately regulated. Signaling pathways are often initiated by multiprotein complexes that translate a stimulus (for example, receptor stimulation) to downstream effectors. Multiprotein signaling complexes, or "signaling machines," are responsible for orchestrating complex signaling events at glutamate receptors (*1*, *2*), ATP receptors (*3*), cytokine receptors (*4*), growth factor receptors (*5*), and heterotrimeric guanine nucleotide-binding protein (G protein)–coupled receptors (*6*). They seem to share the common organization of a membrane-bound receptor embedded in a protein scaffold with interacting signaling and cytoskeletal proteins. Phosphorylation is employed to modulate protein function and stability and to mediate phosphorylation-dependent protein-protein interactions, for example, binding phosphoinositide 3-kinase (PI3K) through its Src homology 2 (SH2) domain to the NR2B subunit of the *N*-methyl-D-aspartate (NMDA)–type glutamate receptor) (*7*). These phosphorylation-dependent interactions confer a higher-order level of regulation of such protein complexes than that imparted by changes in protein expression.

Historically, phosphorylation and its importance in regulating signal transduction and protein function have been studied at the level of single molecules (*8*, *9*), but new proteomic strategies lend themselves to the global characterization of the signaling properties of the proteome. The systematic characterization of kinase and phosphatase substrates is necessary to extend current knowledge of known signaling pathways, and it will inevitably reveal many more pathways. It is now apparent that the extent of phosphorylation in a proteome far exceeds original estimates, with the majority of proteins capable of being phosphorylated at multiple sites. Up to 20 to 30 phosphoresidues in a single protein have been reported (*10*, *11*). Of the estimated 100,000 possible phosphorylation sites in the human proteome, only 3 to 4% are presently known, highlighting the daunting task ahead (*12*). The use of mass spectrometry to characterize phosphorylation has rapidly increased, with a 5000% increase in phosphorylation sites defined by this approach from 1992 to 2002 (*13*). Although mass spectrometry–based strategies for phosphorylation site identification have clear advantages over traditional techniques [for a review, see Mann *et al.* (*14*)], it is not yet routine to analyze complex phosphoprotein samples. Analysis of phosphoproteins is not straightforward, for a number of reasons. First, the stoichiometry of phosphorylation is generally low. Second, the phosphorylated sites on proteins might vary, implying that any given phosphoprotein is heterogeneous. Third, many classes of proteins are present at low abundance within cells, mandating enrichment before analysis. Fourth, most analytical techniques used for studying protein phosphorylation have a limited dynamic range, meaning that minor sites might be difficult to identify. Finally, endogenous phosphatases could dephosphorylate proteins and peptides unless their activity is inhibited during sample preparation.

Sample preparation is crucial for the success of any mass spectrometry–based phosphoproteomic analysis, and numerous approaches have been developed to overcome the problems associated with sample preparation. These include phosphoresidue enrichment strategies such as affinity purification with phospho-specific antibodies or immobilized metal affinity chromatography (IMAC), labeling strategies such as ^{32}P-ATP, and chemical modification with subsequent biotin tagging [for a review, see Mann *et al.* (*14*)]. IMAC enrichment, which is based on the affinity of metals [usually Fe(III) or Ga(III)] to negatively charged phosphate groups, has become the most common approach for in vivo phosphorylation site identification (*13*). This is largely due to improvements made in the specificity of IMAC brought about by chemical derivatization of carboxylic acid groups to methyl esters, which decreases nonspecific binding of acidic peptides (*15*). Recent phosphoproteomic studies have utilized peptide IMAC protocols together with methyl-esterification to enrich for phosphopeptides prior to mass spectrometry. This approach has been used successfully to study phosphorylation in yeast (*15*), *Arabidopsis* (*16*), and cell lines (*17*); however, the use of these IMAC approaches on mammalian subcellular organelles has not proved as successful (*18*, *19*).

Appendix 10

PROTOCOL

The strategy employed for phosphorylation analysis depends on the type of mass spectrometer; the most commonly used are ESI mass spectrometers (usually nanoelectrospray) in conjunction with one- or two-dimensional liquid chromatography separation (*13*). Triple quadrupole mass spectrometers were the first to be widely used for phosphorylation studies, particularly with the nanoelectrospray source (*20*). This instrument was particularly well suited to linked scanning modes: Neutral loss and precursor ion scanning were useful for detecting phosphopeptides in the presence of much larger amounts of unmodified peptide. The triple quadrupole for proteomics work was largely superseded by the more sensitive quadrupole time-of-flight (Q-Tof) geometry, which could not perform linked scans by the very nature of the geometry. Postprocessing of the data could extract phosphopeptide information (and, indeed, information on many other types of posttranslational modification), and the increased sensitivity of the Q-Tof made this trade-off acceptable. A new generation of instruments based on linear traps offer new scanning possibilities for phosphopeptides similar to those originally available with triple quadrupoles. It should be noted that no one mass spectrometer type would cover all eventualities. In particular, we point out that matrix-assisted laser desorption ionization (MALDI) still has a very useful role in studying phosphopeptides, and with enrichment methods may be better suited to the detection of multiphosphorylated peptides (*21*). The work described here was carried out on either a Q-ToF Ultima (Waters) or a 4000 Q-Trap (AB). Specific mass spectrometry scanning modes, such as neutral loss and precursor ion scanning, have been developed to improve the detection of phosphorylation in highly complex peptide mixtures. Gas phase β elimination of phosphorylated serine and threonine residues produce a loss of phosphate that can be monitored in neutral loss mode, and sequence information can be obtained in the same experiment. When phosphoserine-, phosphothreonine-, or phosphotyrosine-containing peptides are fragmented by collision-induced dissociation (CID) in the negative ion mode, a characteristic PO_3^- ion ($m/z = -79$) can be detected by precursor ion scanning. Once this ion is detected, the mass spectrometer can be set to automatically switch to positive ion mode, to conduct tandem mass spectrometry (MS/MS) for sequencing.

In this Protocol, we describe improvements to traditional IMAC- and liquid chromatography tandem mass spectrometry (LC-MS/MS)-based strategies to study the synaptic phosphoproteome by implementing multistage proteome fractionation and phosphoproteome enrichment and iterative mass spectrometry-based scanning approaches (Fig. 1). This Protocol should be useful for researchers pursuing such large-scale analyses of phosphorylation in complex samples.

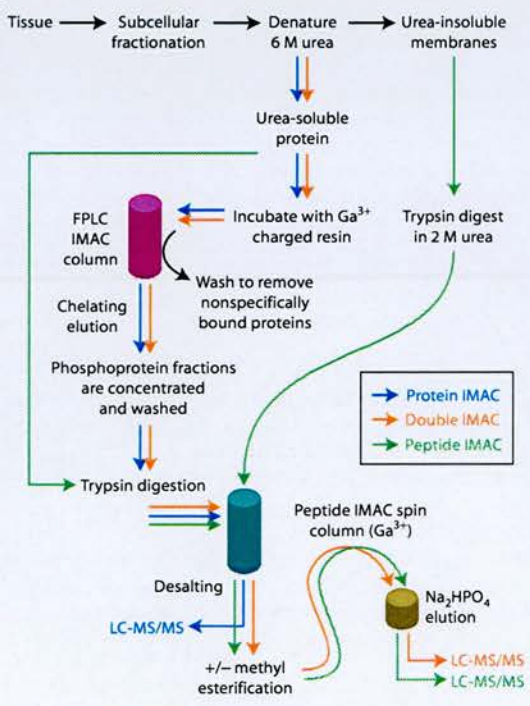


Fig. 1. Schematic overview of the methodology. The target proteome, in this case a subcellular fraction, is phase-partitioned by its differential solubility in 6 M urea. The soluble fraction can be processed in two ways. Traditionally, this fraction is digested in solution, desalted, applied to a peptide IMAC column (green arrows) with or without methyl-esterification, and analyzed by LC-MS/MS. In addition, we have developed a top-down enrichment step in which the soluble proteins are applied to a protein IMAC column (blue arrows), and the eluted phosphoproteins are digested and also analyzed by LC-MS/MS. A novel combination of these two approaches (orange arrow), in which a two-step enrichment of phosphoproteins and phosphopeptides and analysis by LC-MS/MS, results in improved identification of phosphorylation sites. The urea-insoluble fraction, which represents tightly bound or integral membrane proteins, can be analyzed by direct digestion in 2 M urea to release peptides from cytoplasmic and extracellular domains of these proteins, which then can be applied to a peptide IMAC column to isolate phosphopeptides. These combined strategies, when applied to the synapse phosphoproteome, yielded identification of over 650 phosphorylation events corresponding to 331 phosphorylation sites (*10*). This combined strategy of phase-partitioning and sequential enrichment of phosphoresidues allows comprehensive analysis of a complex sample.

Appendix 10

PROTOCOL

Materials

Chemicals

2-Mercaptoethanol (2-ME)

3-(Cyclohexylamino)-1-propanesulfonic acid (CAPS)

Acetic acid

Acetonitrile chromasolv (Fluka, #34851)

Acetyl chloride

Acrylamide

Ammonium bicarbonate

Ammonium persulfate

Aprotinin (Roche, #981532)

Bis-acrylamide

Bromophenol blue

Dimethyl sulfoxide (DMSO)

Dithiothreitol (DTT)

ECL Plus Western blotting detection reagents (Amersham Bioscience, #RPN2132)

Ethylene diamine tetraacetic acid (EDTA)

Ethylene glycol-bis(2-aminoethylether)-N,N,N′,N′-tetraacetic acid (EGTA)

Formic acid

Gallium chloride $GaCl_3$, 5-g vials (Sigma-Aldrich, #25,419-3, or Avocado Research Chemicals, #43879)

Glycerol

[Glu^1]-fibrinopeptide B (Sigma-Aldrich, #F-3261)

Glycine

H_2O_2 (30%)

Iso-butanol

KCl

Leupeptin (Roche, #1034626)

Liquid nitrogen

Luminol

Methanol

Na_2HPO_4

NaCl

p-Coumaric acid

Phenylmethanesulfonyl fluoride (PMSF)

Pro-Q Diamond Phosphoprotein gel stain kit (Molecular Probes, #M-33306)

Sodium dodecyl sulfate (SDS)

Sodium acetate

Sodium ortho-vanadate

Appendix 10

Sucrose

TEMED

Thiourea

Tris-acetate

Tris-base

Tween-20

Urea, ultra grade (Sigma, #U0631)

Enzymes and Proteins

Peppermint-Stick phosphoprotein molecular weight standards (Molecular Probes, #P-33350)

Trypsin Gold, Mass Spectrometry Grade (Promega, #V5280)

Plastics

1.5-ml microcentrifuge tubes

38-ml polycarbonate centrifuge tubes (Nalgene, #306/0258/08)

Columns and Resins

Analytical column: PepMap C18 3 μm 100 Å, 75 μm i.d. × 15 cm (LC Packings, #160321)

Note: Other column types of similar performance may also be used.

Bio-Spin chromatography columns (BioRad, #732-6025)

C 10/10 chromatography column (Amersham Bioscience, #19-5001-01)

Fast-flow chelating sepharose with iminodiacetic acid (IDA) (Amersham Biosciences, #17-0575-01)

RESOURCE 15RPC 1 ml reversed phase chromatography column (Amersham Bioscience, #17-1181-01)

Self Pack POROS 20 MC media (Applied Biosystems, #1-5428-02)

Trap column: BetaMax Neutral 5 μm, 180 μm i.d. × 30 mm (Thermo Hypersil-Keystone, #95005-030215)

Trap column: PepMap C18 3 μm, 100 Å, 300 μm i.d. × 5 mm (LC Packings)

Vivaspin 6 PES membrane spin column (Vivascience)

Vivaspin 20 PES membrane 5000 MWCO (Vivascience, #VS2011)

Other

Hyperfilm ECL (Amersham Bioscience, #RPN2103K)

Micro BCA Protein Assay Kit (Pierce, #23235)

Totallab software (Nonlinear Dynamics)

Appendix 10

PROTOCOL

Equipment

37°C incubator

4000 Q TRAP (Applied Biosystems)

ÄKTA fast-performance liquid chromatography (FPLC; Amersham Bioscience, #18-1900-256)

Dounce homogenizer (VWR, #406/0315/10)

Liquid nitrogen dewar flask

Mascot V2.0 (Matrix Science) server

Mass spectrometer: Q-Tof Ultima API with nanospray source (Micromass)

Note: Other types of Q-Tof API or tandem mass spectrometers that give similar performance can also be used.

Microcentrifuge

Mini-PROTEAN II Multi-Screen apparatus (BioRad, #170-4017)

Nano flow HPLC system with an autosampler and a valve switch unit, such as the Ultimate gradient pump equipped with FAMOS autosampler and Switchos II (LC Packings), or CapLC with nano stream selected module (Waters)

Peristaltic pump (Amersham Bioscience, #18-1110-91)

Picofuge

Roller shaker

Savant SpeedVac (Thermo)

Sorvall Discovery 100SE ultracentrifuge with swinging bucket rotor AH-629 (36 ml) or equivalent

Typhoon Scanner (Amersham Bioscience, #9410-PC)

Vortexer

X-ray film developer

Recipes

Recipe 1: Synaptosome Buffer A

50 mM Tris-acetate, pH 7.4 (from 0.5 M stock)

10% (w/v) sucrose

5 mM EDTA (from 0.5 M stock)

1 mM PMSF

2 μg/ml aprotinin (from 2 mg/ml stock)

2 μg/ml leupeptin (from 2 mg/ml stock)

1 mM sodium ortho-vanadate

Add H_2O to 50 ml. Combine components in a 50-ml tube, rotate to dissolve sucrose, and chill at 4°C.

Appendix 10

PROTOCOL

Recipe 2: Synaptosome Buffer B

5 mM Tris-acetate, pH 8.1 (from 0.5 M stock)

1 mM PMSF

2 μg/ml aprotinin (from 2 mg/ml stock)

2 μg/ml leupeptin (from 2 mg/ml stock)

1 mM sodium ortho-vanadate

Add H_2O to 50 ml. Combine components in a 50-ml tube and chill at 4°C.

Recipe 3: Synaptosome Buffer C

50 mM Tris-acetate, pH 7.4 (from 0.5 M stock)

10% (w/v) sucrose

Combine components in a 50-ml tube, rotate to dissolve sucrose, and chill at 4°C.

Recipe 4: Synaptosome Buffer D

50 mM Tris-acetate, pH 7.4 (from 0.5 M stock)

28.5% (w/v) sucrose

Combine components in a 50-ml tube, rotate to dissolve sucrose, and chill at 4°C.

Recipe 5: EDTA-EGTA Solution

Make a solution of 100 mM EDTA and 100 mM EGTA in H_2O.

Recipe 6: 100 mM $GaCl_3$

Carefully break off the glass seal of the vial of $GaCl_3$ powder (5 g) and immerse in 284 ml of distilled H_2O in a glass beaker. Once dissolved, store in at 4°C.

Note: Caution must be exercised when dissolving $GaCl_3$ in water, as it is an exothermic reaction, which sputters and evolves HCl gas. Wear protective clothing and perform in a fume hood.

Recipe 7: Protein IMAC Wash Buffer

6 M urea

50 mM Tris-acetate, pH 7.4

Adjust volume to 1 liter and pass through a 0.2-micron filter.

Recipe 8: Protein IMAC Elution Buffer

6 M urea

50 mM Tris-acetate, pH 7.4

100 mM EDTA

100 mM EGTA

NaOH pellets

Dissolve the urea in approximately 800 ml of H_2O by stirring on a heated magnetic stirrer. Add the EDTA and EGTA powder and

Appendix 10

Tris-acetate stock (100 ml of 0.5 M) to the dissolved urea and add NaOH pellets to raise the pH to 7.5 to dissolve the EDTA and EGTA. Adjust volume to 1 liter with H_2O and pass through a 0.2-micron filter.

Recipe 9: 5× Reducing Sample Buffer

0.5 M Tris-HCl, pH 6.8 1 ml

Glycerol 1 ml

10% SDS 1.6 ml

2-ME 0.4 ml

0.1% bromophenol blue 0.2 ml

H_2O 4 ml

Note: 20 mM DTT (0.32 ml of 0.5 M stock) can be used in place of the 2-ME.

Recipe 10: SDS-PAGE Gel Solutions

Acrylamide/Bis

Acrylamide 146 g

Bis-acrylamide 4 g

Add H_2O to 500 ml. Store at 4°C or at room temperature.

1.5 M Tris, pH 8.8

Tris-base 90.75 g

Prepare in 400 ml of H_2O, adjust pH to 8.8 with 1 M HCl, and add H_2O to 500 ml. Store at room temperature.

0.5 M Tris-HCl pH 6.8

Tris-base 30 g

Prepare in 400 ml of H_2O, adjust pH to 6.8 with 1 M HCl, and add H_2O to 500 ml. Store at room temperature.

10% SDS (w/v)

SDS (sodium salt) 10 g

Add H_2O for a final volume of 100 ml. Store at room temperature.

10% APS (w/v)

Ammonium persulfate 0.1 g

Add H_2O to a final volume of 1 ml. Store at room temperature and discard after 1 week.

Recipe 11: 10× Running Buffer

Tris-base 30 g

Glycine 144 g

SDS 10 g

Add H_2O to 1 liter final volume.

Note: Buffer should be 8.3; if it is not, adjust as needed.

Recipe 12: Phosphoprotein Gel Fixing Solution

Prepare a solution of 50% methanol and 10% acetic acid.

Appendix 10

PROTOCOL

Recipe 13: Phosphoprotein Gel Destain Solution

1 M sodium acetate, pH 4.0 50 ml

Ultrapure water 750 ml

Acetonitrile 200 ml

Combine and mix thoroughly.

Recipe 14: SYPRO Wash Solution

Prepare a 100-ml solution of 10% methanol and 7% acetic acid.

Recipe 15: Transfer Solution

Make a solution of 10% (v/v) methanol in 10 mM CAPS, pH 11.0 at 4°C.

Recipe 16: 20× PBS

NaCl 160 g

KCl 4 g

Na_2HPO_4 (anhydrous) 12.2 g

KH_2PO_4 4 g

Make in H_2O at a final volume of 1 liter. Store at 4°C or room temperature.

Note: At 4°C, a precipitate may form. Warm and dissolve before use if necessary.

Recipe 17: PBS-Tween

Prepare a 1× solution of PBS [use 20× PBS (Recipe 16)] supplemented with 0.1% Tween-20.

Recipe 18: Chemiluminescence Solutions

Luminol Stock

Make a solution of 250 mM luminol (44.3 g/l) in DMSO. Store at –20°C.

p-Coumaric Acid Stock

Make a solution of 90 mM *p*-coumaric acid (14.8 g/l) in DMSO. Store at –20°C.

Solution A

Luminol stock 5 ml

p-Coumaric acid stock 5 ml

1.5 M Tris, pH 8.8 12.5 ml

0.5 M Tris, pH 6.8 12.5 ml

Add 215 ml of H_2O. Store at 4°C in the dark.

Solution B

250 µl 30% H_2O_2

250 ml H_2O

Store at 4°C in the dark. Mix Solutions A and B at a ratio of 1:1 just prior to blot development.

Appendix 10

PROTOCOL

Recipe 19: Tryptic Digest Buffer

1 M urea

0.125 M thiourea

25 mM NH_4HCO_3, pH 8

Use this solution to dilute protein or phosphoprotein sample.

Note: Trypsin is reconstituted with 50 mM acetic acid and is added (in a 1:20 ratio to protein or phosphoprotein) to this solution.

Recipe 20: Reversed Phase Mobile Phase A

94.9% H_2O

5% acetonitrile

0.1% TFA

Note: H_2O must be HPLC gradient grade.

Recipe 21: Reversed Phase Mobile Phase B

94.9% acetonitrile

5% H_2O

0.1% TFA

Note: H_2O must be HPLC gradient grade.

Recipe 22: 2 M Methanolic HCl

Add 160 μl of acetyl chloride slowly with stirring to 1 ml of methanol in a dropwise fashion.

Note: Caution must be exercised during the addition of acetyl chloride, because this reaction generates heat. Perform operation in fume hood using safety glasses and gloves.

Recipe 23: High Organic Loading Buffer

1:1:1 of acetonitrile, methanol and H_2O

pH should be 2.5 to 3.

Note: H_2O must be HPLC gradient grade.

Recipe 24: Aqueous Loading Buffer

0.3% acetic acid

pH should be 2.5 to 3.

Note: H_2O must be HPLC gradient grade.

Recipe 25: Peptide IMAC Elution Buffer

200 mM Na_2HPO_4

Appendix 10

PROTOCOL

Recipe 26: Standard for Q-Tof Optimization

25% acetonitrile

74.9% H_2O

0.1% formic acid

Add [Glu^1]-fibrinopeptide B to a 500 fmol/μl concentration.

Recipe 27: Standard for Q-Tof Ultima Optimization

25% acetonitrile

74.9% H_2O

0.1% formic acid

Add [Glu^1]-fibrinopeptide B to a 100 fmol/μl concentration.

Recipe 28: 0.1% Formic Acid

99.9% H_2O

0.1% formic acid

Recipe 29: Tryptic Digest LC-MS/MS Mobile Phase A

94.9% H_2O

5% acetonitrile

0.1% formic acid

Note: Both acetonitrile and H_2O must be HPLC gradient grade.

Recipe 30: Tryptic Digest LC-MS/MS Mobile Phase B

95% acetonitrile

4.9% H_2O

0.1% formic acid

Recipe 31: 0.5% Formic Acid

99.5% H_2O

0.5% formic acid

Recipe 32: IMAC LC-MS/MS Mobile Phase A

99.9% H_2O

0.1% formic acid

Recipe 33: IMAC LC-MS/MS Mobile Phase B

70% acetonitrile

29.9% H_2O

0.1% formic acid

Appendix 10

PROTOCOL

Instructions

Purification of Synaptosomes by Subcellular Fractionation

This section briefly describes purification of synaptosomes from mouse forebrains. Although this method has been well documented (*22*), we have made several modifications.

1. Rapidly dissect mouse forebrains and freeze immediately in liquid nitrogen, and store at –80°C.
2. Cool all Synaptosome Buffers (Recipes 1 through 4) and the homogenizer to 4°C.
3. Add 16 mouse forebrains to Synaptosome Buffer A (Recipe 1) using 9 ml of buffer per gram of tissue in a 50 ml Dounce homogenizer.
4. Homogenize with 15 strokes.
5. Collect the extract and centrifuge at 800*g* for 20 min.

   Note: All centrifugation steps in this section should be performed at 4°C.
6. Centrifuge the supernatant at 16,000*g* for 30 min.
7. Resuspend the pellet in Synaptosome Buffer B (Recipe 2) in 5 ml of buffer per gram of original tissue starting material.
8. Homogenize with three strokes.
9. Incubate the extract in the homogenizer at 4°C for 45 min.
10. Homogenize with 10 strokes.
11. Add sucrose to 34% (w/v).
12. Prepare sucrose gradients by addition of 10 ml of Synaptosome Buffer C [Recipe 3; 10% sucrose (w/v)] to 38-ml polycarbonate centrifuge tubes, followed by careful introduction of 10 ml of Synaptosome Buffer D [Recipe 4; 28.5% sucrose (w/v)] from the bottom of the tube.
13. Add 10 ml of extract from step 11 [in 34% sucrose (w/v)] to the bottom of the tube, such that the gradients are 10%, 28.5%, and 34% sucrose (w/v) from the top to the bottom.
14. Centrifuge at 60,000*g* for 120 min.
15. Collect the protein-containing band that forms between the 34% and 28.5% sucrose gradients, and dilute with 50 mM Tris-acetate pH 7.4 to 10% sucrose.
16. Centrifuge at 48,000*g* for 30 min.
17. Discard the supernatant.
18. Resuspend this synaptosome pellet in 50 mM Tris-acetate, pH 7.4, remove a small aliquot (2 μl) for subsequent protein concentration determination (MicroBCA kit, Pierce) and store at –80°C.

    Note: This is the synaptosomal protein sample.

Protein IMAC Sample Preparation

In this protein IMAC Protocol, a cell lysate or subcellular fraction is solubilized in 6 M urea and is incubated with a metal-charged sepharose-based resin. Specific elution is achieved by a stopped-flow elution with chelating buffer, and eluted phosphoproteins can be concentrated and washed. When performed with an FPLC system, this method permits large-scale purification, which is necessary for subsequent phosphopeptide analysis by mass spectrometry (Fig. 2). Batch purification in microcentrifuge

Appendix 10

tubes can also be performed, which can be useful for rapid purification of smaller quantities of phosphoprotein for gel or Western blot analysis. We recommend a working concentration of approximately 1 mg/ml for the protein extract and a ratio of 10:1 of extract to resin to ensure specific binding of phosphoproteins.

1. Strip the fast-flow chelating sepharose with IDA of any bound metal by incubation with EDTA-EGTA Solution (Recipe 5) for 1 hour with gentle agitation.
2. Clamp a C 10/10 chromatography column to a retort stand and connect to a peristaltic pump.
3. Add 5 ml (bed volume) of the resin from step 1 to the column.

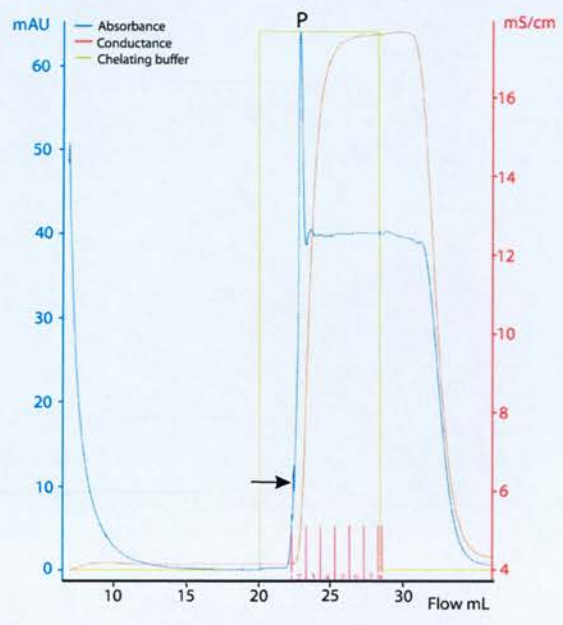


Fig. 2. FPLC chromatogram of a protein IMAC purification. Unbound protein is washed off the IMAC column with 20 column volumes of Protein IMAC Wash Buffer (Recipe 7) (that is, set pump A to 100% Protein IMAC Wash Buffer) at a flow rate of 1 ml/min until the A280 > 0.25 mAU. Then the flow is switched to Protein IMAC Elution Buffer (Recipe 8) (set pump B to 100% Protein IMAC Elution Buffer) and run at 0.5 ml/min. A stopped-flow elution with a chelating buffer is performed to elute the phosphoproteins in a discrete peak (P). The arrow indicates the point of stopped-flow.

4. Wash the resin with 10 column volumes of H_2O using the pump at a flow rate of 1 ml/min.
5. Remove the resin, add to 50 ml of 100 mM $GaCl_3$ (Recipe 6), and incubate at room temperature on a roller-mixer overnight.
6. Wash the charged resin with 10 resin volumes of H_2O using the pump at a flow rate of 1 ml/min.
7. Store the charged resin in an equal volume of H_2O (50% H_2O) at 4°C and use within 1 week.
8. Solubilize the synaptosomal protein sample in 6 M urea (12.5 mg protein/10 ml) in 1.5-ml microfuge tubes for 20 min with mixing.
9. Centrifuge at 16,000*g* for 5 min and retain the supernatant.
10. Load 1 ml of washed, gallium-charged resin into a C 10/10 column and equilibrate the resin by washing the column with five bed volumes of Protein IMAC Wash Buffer (Recipe 7).
11. Incubate the urea-soluble supernatant (about 9.5 ml) with 1 ml of resin for 1 hour at room temperature on a roller shaker.
12. Equilibrate an FPLC (ÄKTA) system with 5 ml of Protein IMAC Elution Buffer (Recipe 8) and then 5 ml of Protein IMAC Wash Buffer (Recipe 7) at room temperature.
13. Pack the resin-protein mixture into a C 10/10 column and attach to the FPLC, avoiding introduction of air bubbles into the system.
14. Wash unbound protein from the column with 20 column volumes of Protein IMAC Wash Buffer (Recipe 7) (that is, set pump A to 100% Protein IMAC Wash Buffer) at a flow rate of 1 ml/min until the A280 > 0.25 mAU.

Appendix 10

15. Set pump A to 0% Protein IMAC Wash Buffer (Recipe 7) and switch the flow to 100% Protein IMAC Elution Buffer (Recipe 8) (pump B) and run at 0.5 ml/min, and start to collect 1-ml fractions.
16. Once the A280 reaches ~ 10 mAU, change pump B to 0% Protein IMAC Elution Buffer (Recipe 8) (no flow) for 10 min.
17. Switch pump B to 100% Protein IMAC Elution Buffer (Recipe 8) and run at 0.5 ml/min and start to collect 1 ml fractions.

    Note: This stopped flow elution results in a discrete elution peak; see Fig. 2.
18. Pool four or five fractions surrounding the phosphoprotein peak and wash with Protein IMAC Elution Buffer (Recipe 8) using a 1:4 dilution of the sample to Protein IMAC Wash Buffer (Recipe 7) in a Vivaspin 6 PES membrane spin column.
19. Concentrate the sample to 200 µl and repeat this wash step to further reduce concentration of EDTA and EGTA in the sample.

Validation of Phosphoprotein Purification

This section describes validation of the protein IMAC enrichment, which is advisable before further purification or analysis by mass spectrometry. Validation can be achieved by phosphoprotein staining of protein IMAC-purified phosphoproteins on an SDS-polyacrylamide gel (Fig 3). Alternatively, Western blotting with phospho-specific antibodies or antibodies to known phosphoproteins should show clear enrichment in the protein IMAC-purified sample (Fig. 3). Because elements of these two techniques are routine, they will be described briefly.

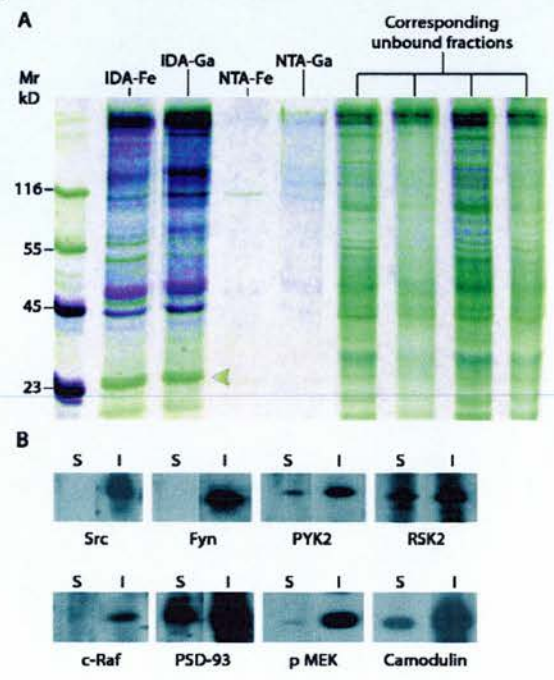


Fig. 3. Validation of IMAC phosphoprotein isolation strategy. (**A**) Overlay of Pro-Q diamond phosphoprotein and Ruby total protein stains of an SDS-PAGE gel for different resin and metal IMAC combinations used to selectively isolate phosphoproteins from a urea-soluble synaptosomal extract. The corresponding unbound fractions are included for comparison. Twenty percent of eluted protein from each purification was separated on a 12% SDS-PAGE gel, stained, and imaged sequentially. Phosphorylated protein appears pink to black, whereas unphosphorylated protein appears green. Lane 1 (far left) shows Peppermint Stick molecular weight markers in which proteins of molecular weight 23 and 45 kD are phosphorylated and those of 55 and 116 kD are not phosphorylated, thus serving as an internal control for selective staining of phosphorylated proteins. A 25-kD band (green arrow) consists of ribosomal proteins, which were retained on the IMAC resins due to ion-exchange effects. (**B**) Western blots of equal amounts of IMAC enriched (I) versus synaptosomal (S) protein for eight synaptic phosphoproteins. A substantial enrichment for all phosphoproteins was observed in the protein IMAC sample, and a specific enrichment for the phosphorylated form of mitogen-activated protein kinase kinase (pMEK) was seen using a phospho-specific antibody.

Electrophoresis and phosphoprotein staining of SDS-PAGE gels

1. Boil 10% (20 µl) of the eluted phosphoprotein and molecular weight markers in tightly closed microcentrifuge tubes in 5× Reducing Sample Buffer (Recipe 9).
2. Separate on an SDS-PAGE (Recipes 10 and 11) minigel for 50 min at 200 V.

   Note: All of the following incubation steps are performed at room temperature with gentle agitation.

Appendix 10

3. Fix the gel by incubating twice for 30 min, each time in 100 ml of Phosphoprotein Gel Fixing Solution (Recipe 12).
4. Wash the gel three times for 10 min, each time in 100 ml of H_2O.
5. Stain the gel with 50 ml of Pro-Q Diamond phosphoprotein gel stain in the dark for 90 min.
6. Destain the gel three times for 30 min, each time with 100 ml of Phosphoprotein Gel Destain Solution (Recipe 13).
7. Wash the gel twice with H_2O for 5 min each.
8. Scan the gel with a Typhoon scanner using an excitation maximum at 555 nm and an emission maximum at 580 nm.
9. Rinse the gel twice with H_2O for 5 min each.
10. Incubate the gel in 50 ml SYPRO Ruby Gel Stain Solution overnight.
11. Wash in SYPRO Wash Solution (Recipe 14) for 30 min and then twice with H_2O for 5 min each.
12. Scan the gel with a Typhoon scanner using an excitation maximum at 280/450 nm and an emission maximum at 610 nm.
13. Overlay the phosphoprotein and total protein images of the gel using Totallab software.

Western blotting

Western blotting of protein IMAC-enriched samples and whole synaptosome samples for a selected set of phosphoproteins can be performed using a Mini-PROTEAN II Multi-Screen apparatus (Fig. 3). This apparatus allows screening of up to 40 different antibodies on a single membrane and requires much less of each individual antibody compared to traditional Western blotting. This offers a high-throughput approach for validating a set of proteins, but a conventional Western blotting apparatus will suffice in its absence.

1. Adapt an SDS-PAGE gel-casting comb to cast a single well across the gel.

   Note: This can be done by taping over the teeth or by using a specialized comb consisting of a single tooth.
2. Separate the samples on a 12% SDS-PAGE minigel and transfer to PVDF membrane at 4°C for 90 min at 75 V in Transfer Solution (Recipe 15).
3. Block the membrane with 1% BSA in PBS-Tween (Recipe 17) overnight at 4°C.
4. Wash the membrane three times for 5 min, each time with 20 ml of PBS-Tween (Recipe 17).
5. Place the membrane onto the base of the Multi-Screen apparatus; tighten screws to clamp.
6. Use optimal dilutions of primary antibodies in PBS-Tween (Recipe 17), 550 µl per well, and incubate for 2 hours with gentle agitation at room temperature.

   Note: Antibody concentrations must be empirically determined or, for commercially available antibodies, use the recommended concentration for Western blotting.
7. Wash each well three times for 5 min, each time with 0.5 ml of PBS-Tween (Recipe 17).
8. Add 0.6 ml of peroxidase-linked secondary antibody diluted 1:5000 in 1% BSA with PBS-Tween (Recipe 17) and incubate for 2 hours at room temperature.
9. Wash each well three times with 0.5 ml of PBS-Tween (Recipe 17) for 5, 10, and 15 min each.
10. Detect the antibodies by enhanced chemiluminescence using Chemiluminescence Solutions A and B (Recipe 18); just before developing the blot, mix enough of solutions A and B at a 1:1 ratio to cover.

    Note: ECL Plus Western blotting detection reagents (Amersham Biosciences) may be used instead of Recipes 17 and 18.

In-Solution Protein Digestion

1. Dilute the 3 mg of the protein IMAC-enriched sample in Tryptic Digest Buffer (Recipe 19) to a final volume of 400 µl.

   Note: 6 M Urea insoluble membranes can be digested by increasing the urea concentration in the digestion buffer to 2 M. This will result in reduced activity of trypsin but will allow better solubility and thus digestion of these proteins.

Appendix 10

2. Add Trypsin Gold in a ratio of 1:20 in 50 mM acetic acid.
3. Incubate at 37°C for 3 hours.

   Note: We recommend mass spectrometry grade Trypsin Gold, because it has high activity and good specificity. Sequencing-grade trypsin (Roche) can also be used, but an overnight incubation is necessary for complete proteolysis. We have also used sequencing grade modified porcine trypsin (Promega); in our experience, it produces semi-tryptic ends, but this can be accommodated by specifying "semi-trypsin" in Mascot's search parameters.
4. Terminate the reaction by freezing in liquid nitrogen, then lyophilize.

Reversed Phase Desalting of Protein Digests

Digested protein samples, whole lysate, or protein IMAC-enriched that are to be further processed to peptide IMAC samples are desalted on a 1-ml bed volume RESOURCE 15RPC reversed phase chromatography column. Peptides for analysis by mass spectrometry are desalted on-line, as described in the LC-MS/MS section to follow.

1. Wash and equilibrate the column first with 5 column volumes of Reversed Phase Mobile Phase A (Recipe 20), then 5 column volumes of Reversed Phase Mobile Phase B (Recipe 21), and bring back to baseline with Reversed Phase Mobile Phase A again.
2. Load the tryptic digest into the injection loop and introduce into the column at 0.5 ml/min.
3. Wash with Reversed Phase Mobile Phase A (Recipe 20) until baseline is reached, about 10 column volumes.
4. Elute desalted peptides with a 0 to 95% gradient of Reversed Phase Mobile Phase B (Recipe 21) over 5 min.
5. Collect fractions corresponding to the peptide peak (absorbance at 215 nm), pool, and lyophilize.

Peptide IMAC Sample Preparation

Peptide IMAC was carried out as described (*15*) with some modifications. A number of conditions in this Protocol can affect the yield and population of phosphopeptides obtained. First, the choice of metal used to perform peptide IMAC can affect specificity and recovery. The most commonly used metals are Fe^{3+} and Ga^{3+} (*15*, *23*, *24*), and these display different efficiencies of capture and release of phosphopeptides (*23*). However, we found that Ga^{3+} performed the best in protein IMAC experiments (*10*), so we recommend using Ga^{3+} in the peptide IMAC step of a sequential phosphoprotein and phosphopeptide purification to minimize losses due to differences in specificity. Acidic peptides affect the specificity of peptide IMAC purifications, and methyl-esterification of peptide side chains improves the specificity (*15*). We have observed a 30% reduction in the number of acidic amino acids in phosphopeptides that have been detected in peptide IMAC of methyl-esterified samples compared to unmodified samples (*25*). However, other studies have found that contaminating peptides do not have any different physiochemical properties than phosphopeptides in the sample, but were actually from highly abundant proteins in the sample (*23*).

In addition, peptide chemical modification reactions do not go to completion, and we have observed that methyl-esterification in particular never occurs on 100% of peptides. This has the effect of diluting any particular sample into populations of modified and unmodified peptides, which effectively reduces the amount of any given peptide and increases the sample complexity overall. We recommend performing peptide IMAC on both methyl-esterified and unmodified versions of a sample to increase overall phosphopeptide coverage.

Next, the buffer in which peptides are solubilized and applied to the column affects the population of phosphopeptides recovered. In sequential protein and peptide IMAC experiments, we used a high-organic (equal volumes of acetonitrile, methanol, and water, pH 2.5 to 3) and an aqueous (0.3% acetic acid, pH 2.5 to 3) buffer to solubilize and load peptides onto the peptide IMAC column. Phosphopeptides found in both conditions represented a small fraction of the total phosphopeptides detected, with 37% and 60% of the double IMAC dataset found in the high organic and aqueous experiments, respectively (*25*). The reason for this differential appearance based on the solvent can be explained by looking at the grand average hydrophobicity (GRAVY) (*26*) values of these peptides (Fig. 4). Phosphopeptides found in aqueous-solubilized samples are overall less hydrophobic than those found in the high-organic-solvent-solubilized samples. This would indicate that in previous studies, such as that of the yeast phosphoproteome (*15*), in which only a high-organic loading buffer was used, a substantial subpopulation of phosphopeptides may have been lost at this step. It is clear that in any complex sample, there are populations of peptides that favor different solubilization conditions depending on their hydrophobicity, so it is recommended that different conditions be used to achieve greater coverage of the peptide population.

Appendix 10

PROTOCOL

Methyl-esterification of peptide digests

1. For methyl-esterification, reconstitute the lyophilized desalted peptide digest with 0.5 mg/ml of 2 M Methanolic HCl (Recipe 22).
2. Incubate with gentle agitation for 2 hours at room temperature.
3. Quench with two volumes of H_2O and lyophilize.

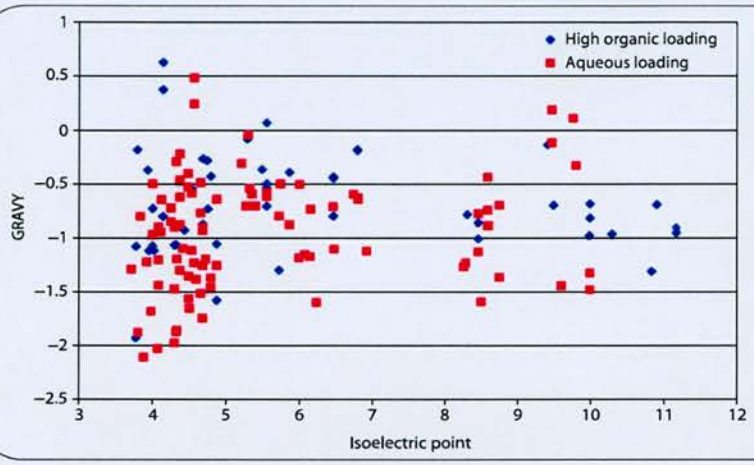


Fig. 4. Comparison of peptide IMAC solubilization and loading buffers. Grand average of hydrophobicity (GRAVY) versus isoelectric point for phosphopeptides found in two peptide IMAC experiments are plotted. In one double IMAC experiment, peptides were solubilized and loaded onto a peptide IMAC column in a high organic buffer (1:1:1 of acetonitrile:methanol:H_2O) (data indicated by blue diamonds), and, in the other, peptides were solubilized and loaded in an aqueous buffer (0.3% acetic acid) (data indicated by red squares). These two buffers enable identification of two populations of phosphopeptides differing in their hydrophobicity.

IMAC enrichment

1. Incubate 0.5 ml (bed volume) of POROS MC with 100 mM EDTA and 100 mM EGTA (Recipe 5) on a roller mixer overnight to remove any trace metals from the resin.
2. Transfer the resin to a C 10/10 column and wash with 10 volumes of H_2O.
3. Incubate with 5 ml of 100 mM $GaCl_3$ (Recipe 6) at room temperature on a roller mixer overnight.
4. Wash the charged resin with 10 bed volumes of H_2O.

   Note: The charged resin (in 50% H_2O) may be stored at 4°C and should be used within 1 week.
5. Reconstitute the peptide sample (methyl-esterified peptide digest or unmodified desalted peptide digest) in either a High Organic Loading Buffer (Recipe 23) or an Aqueous Loading Buffer (Recipe 24).

   Note: For phosphoprotein digests, apply 3 mg of digested phosphoprotein in 1 ml of High Organic Loading Buffer (Recipe 23) to a 100-µl bed volume of gallium-charged POROS MC. A higher concentration of whole cell lysate may be applied to a similar resin volume, because the percentage phosphopeptide content will be lower. For example, 6 mg of digested whole synaptosomes may be applied in 1 ml of High Organic Loading Buffer to a 100-µl bed volume of gallium-charged POROS MC.
6. Centrifuge the peptide solution at 13,000 rpm for 10 min in a microcentrifuge to remove any unsolubilized material.
7. Incubate the solubilized peptides with 200 µl of pre-equilibrated gallium-charged POROS MC slurry (100-µl bed volume) for 1 hour at room temperature with gentle agitation.
8. Load the peptide-resin slurry into a pre-equilibrated spin column and centrifuge in a picofuge briefly to elute the unbound peptides.
9. Wash by addition and removal in a picofuge with 10 resin volumes of the loading buffer (Recipe 23 or Recipe 24) used to load the peptide onto the column.
10. Elute the phosphopeptides with 100 µl of Peptide IMAC Elution Buffer (Recipe 25) by incubating the peptide-resin slurry for 5 min, centrifuging briefly in a picofuge, and then collecting the eluted phosphopeptides.

Appendix 10

11. Repeat the elution (step 10) and pool eluted phosphopeptides.

    Note: The concentration of Na_2HPO_4 used to elute phosphopeptides from the peptide IMAC column may be reduced to 50 mM, if volume reduction by partial lyophilization is necessary.

12. Rapidly freeze at −80°C.

Double IMAC Sample Preparation

This Protocol is a combination of protein and peptide IMAC steps described above and provides the obvious advantage of two rounds of enrichment for phosphorylated molecules. The application of a single protein IMAC strategy to the synaptic phosphoproteome resulted in the detection of 105 phosphorylation events, but selective enrichment of the phosphopeptides derived from these purified phosphoproteins improved the detection of phosphorylation threefold, with the detection of 308 phosphorylation events (*10*).

1. Digest 3 mg of protein IMAC purified phosphoprotein with 1:20 ratio of Trypsin Gold for 3 hours in Tryptic Digest Buffer (Recipe 19).
2. Desalt the digest as described in the Reversed Phase Desalting of Protein Digests section.
3. Perform methyl-esterification if desired (see Peptide IMAC Section).
4. Reconstitute in either a High Organic Loading Buffer (Recipe 23) or an Aqueous Loading Buffer (Recipe 24). Incubate 1 ml of this peptide mixture with 200 µl of pre-equilibrated POROS-Ga slurry (100 µl bed volume) for 1 hour at room temperature.
5. Load the resin into a spin-column and wash with 10 volumes of the same Loading Buffer used to load the peptide onto the column.
6. Elute the phosphopeptides with 100 µl of Peptide IMAC Elution Buffer (Recipe 25) by incubating the peptide-resin slurry for 5 min, centrifuging briefly in a picofuge, and then collecting the eluted phosphopeptides.
7. Repeat and pool the eluted phosphopeptides.

LC-MS/MS Analysis

For LC-MS/MS analysis, a nanoflow HPLC system with valve switch unit is needed, so that the trap column can be placed for loading and desalting the sample efficiently. This is particularly true for desalting phosphopeptide samples in high salt [for example, the 200 mM Na_2HPO_4 in Peptide IMAC Elution Buffer (Recipe 25)] with a large volume size (100 to 200 µl). Using a flow at 10 µl/min, we found it took up to 1 hour to desalt the sample. We used an Ultimate nanoflow gradient pump equipped with FAMOS autosampler and Switchos II (LC Packings), or CapLC with nanostream selected module (Waters) coupled to Q-Tof hybrid tandem mass spectrometer. Mass accuracy in MS/MS spectra in assigning the phosphorylation sites was important.

Before sample analysis on LC-MS/MS, we always perform a systematic check; this process is described.

LC-MS/MS systematic check

1. Optimize the nanospray position and instrument parameters with 500 fmol/µl $[Glu1]^1$-fibrinopeptide B in Standard for Q-Tof Optimization (Recipe 26) or 100 fmol/µl in Standard for Q-Tof Ultima Optimization (Recipe 27).
2. Calibrate the instrument with the $[Glu1]^1$-fibrinopeptide B solution prepared in step 1.
3. Systematically perform LC-MS/MS system checks with a tryptic digest of BSA, such as chromatographic resolution, mass accuracy, and sequence coverage of BSA. For Q-Tof, use either 100 or 200 fmol BSA; for Q-Tof Ultima, 100 fmol.

   Note: If the analytical sample is a small quantity, then the system should be tested using BSA at 10 fmol or lower.

Appendix 10

PROTOCOL

LC-MS/MS analysis of a tryptic digest of IMAC-purified phosphoproteins on a Q-Tof Ultima

MS acquisition starts at t = 3.0 min when desalting is finished. The instrument is operated in automated data-dependent acquisition (DDA) mode using MassLynx. In DDA mode, the instrument continues to scan at Tof-MS survey and only switches to Tof-MS/MS scan when multiply charged ions (precursor ions) in the Tof-MS survey spectrum meet the criterion for fragmentation. This is because majority of ions generated from tryptic peptides is at a charge status of at least 2. Details for the instrument settings are described.

1. Load the protein digest through the autosampler to the trap column on the valve switch unit and desalt with 0.1% Formic Acid (Recipe 28) at a flow rate of 10 or 20 μl/min (depending on the column).

   Note: When using a PepMap C18 3 μm 100 Å, 300 μm i.d. × 5 mm (LC Packings) trap column, the flow rate can be at 20 μl/min. If the trap is BetaMax Neutral 5 μm, 180 μm i.d. × 30 mm (Thermo Hypersil-Keystone), the flow rate should be set at 10 μl/min.

2. Switch valve after 3 min to connect the trap column on-line with the analytical column.

3. Run a gradient as shown in Table 1 to reverse elute the peptides from the trap and separate on the analytical column using Tryptic Digest LC-MS/MS Mobile Phases A and B (Recipes 29 and 30) (run 1). The flow rate through the columns should be 250 to 300 nl/min.

Table 1. Liquid chromatography conditions for analysis of phosphoprotein digest.

Time (min)	Percent A (Recipe 29)	Percent B (Recipe 30)	Comments
0.1	94	6	Trap and analytical column off-line; sample loading and desalting
3.0	94	6	
180.0	70	30	Trap and analytical column on-line; gradient separation; MS acquisition in this section
240.0	50	50	
242.0	20	80	
247.0	20	80	
248.0	94	6	Trap and analytical column off-line to equilibrate separately
300.0	94	6	

4. Acquire Tof-MS survey over the mass range *m/z* 400 to 1500 at 1 s per scan.

   Note: The switching criterion from Tof-MS to Tof-MS/MS is dependent on precursor ion intensity and charge state (+2 and +3 selected), each for 5 s. The collision energy used to perform Tof-MS/MS is automatically varied according to the m/z and the charge state of the precursor ion.

5. Acquire data for Tof-MS/MS over the mass range *m/z* 50 to 2000.

6. Perform Tof-MS/MS on up to three multiply charged ions of the highest intensity with a dynamic exclusion time of 120 s.

   Note: Depending on the complexity of the sample, more precursor ions can be selected from the Tof-MS survey spectrum to subject to Tof-MS/MS. For a highly complex sample, two LC-MS/MS runs with the same LC gradient may be used.

7. After the first standard LC-MS/MS run is finished (run 1), examine the survey data.

8. Select up to five multiply charged ions above the intensity threshold from each survey spectrum that had not been subjected to MS/MS in run 1 and incorporate into an inclusion list for the second run (run 2).

9. Perform the run 2 analysis using the same conditions as in run 1, except that the MS switching to MS/MS should be directed to analyze only multiply charged ions from the "including list."

   Note: We term this LC-MS/MS approach, incorporating a primary analysis followed by a repeat experiment based on an inclusion list, targeted repeat analysis (TRA).

10. Process both LC-MS/MS datasets to pkl files using Proteinlynx in Masslynx, and then merge to a single mgf file with the software (merge) provided by Matrix Science and submit results to Mascot, a search engine that searches protein sequence databases to match peptide spectra to protein sequences (http://www.matrixscience.com).

Appendix 10

PROTOCOL

LC-MS/MS analysis of a tryptic digest of IMAC-purified phosphoproteins on a 4000 Q TRAP

Although we have found that the majority of proteins in protein IMAC samples are phosphorylated, the molar ratio of phosphorylated to nonphosphorylated species is reduced upon tryptic digestion. An alternative to performing a second round of enrichment of phosphopeptides is to selectively monitor phosphopeptides in the mass spectrometer. The specific behavior of phosphopeptides in a mass spectrometer can be exploited in a number of ways for analysis of very complex peptide mixtures with a low abundance of phosphorylated peptides: (i) Neutral loss of 98 in MS/MS for those phosphopeptides containing phosphoserine and phosphothreonine in positive electrospray ionization mode; (ii) 216 fragment ion for phosphopeptides containing phosphotyrosine in positive electrospray ionization mode; and (iii) 79 fragment ion for phosphopeptides in negative electrospray ionization mode. By setting the mass spectrometer to specifically monitor precursor ions of the fragment ions, phosphopeptides can be selectively identified. We found that monitoring the precursor of 79 fragment ion in the negative mode was the most efficient approach, and the use of the 4000 Q TRAP, which is capable of fast auto-switching between negative and positive ionization mode, is well suited for this.

Two precursor-scanning approaches based on monitoring the precursor of 79 fragment ion can be adopted on the 4000 Q TRAP to selectively analyze phosphopeptides. The LC conditions are the same as that used for IMAC-purified phosphoprotein digest, described above.

Approach 1: Conventional precursor of 79 fragment ion scanning

1. Set the instrument in negative ion mode (–ESI) to scan for the precursors of m/z 79 (PO_3^-).
2. When a precursor ion for *m/z* 79 is detected, the instrument is automatically switched to positive mode (+ESI) and MS/MS fragmentation, the detected precursors proceeds.
3. Reset the instrument to –ESI and scan for precursors of m/z 79 (PO_3^-).
4. Upon detection of a precursor ion for *m/z* 79, the instrument automatically switches to +ESI, and MS/MS fragmentation proceeds.

Note: This process is repeated until all precursor ions for m/z 79 have been scanned.

Approach 2: TRA precursor of 79 fragment ion scanning

1. Analyze one-third of the sample in negative mode (–ESI) to detect precursors of *m/z* 79.
2. Examine the data to generate the inclusion list for the second run in positive ion mode (+ESI).
3. Analyze the remaining two-thirds of the sample in information-dependent acquisition (IDA) mode with the above inclusion list.

LC-MS/MS analysis of IMAC-enriched phosphopeptides

This procedure is performed in a similar way to that described above for digested IMAC-purified phosphoproteins, with a few exceptions. The majority of phosphopeptides elute at very low organic percentage compared to those isolated by liquid chromatography separations of a phosphoprotein digests (Fig. 5).

1. Load phosphopeptides to the trap column and desalt with 0.5% Formic Acid (Recipe 31) at a flow rate of 10 μl/minute for 30 to 60 min, depending on salt amount in the sample.

Note: We suggest using the trap column BetaMax Neutral 5 μm, 180 μm i.d. ? 30 mm (Thermo Hypersil-Keystone).

2. Separate phosphopeptides with the gradient shown in Table 2, using IMAC LC-MS/MS Mobile Phases A and B (Recipes 32 and 33) using a flow rate of 150 nl/min.

Note: The high flow rate is due to the high pressure when the trap and analytical columns are set up on-line.

Appendix 10

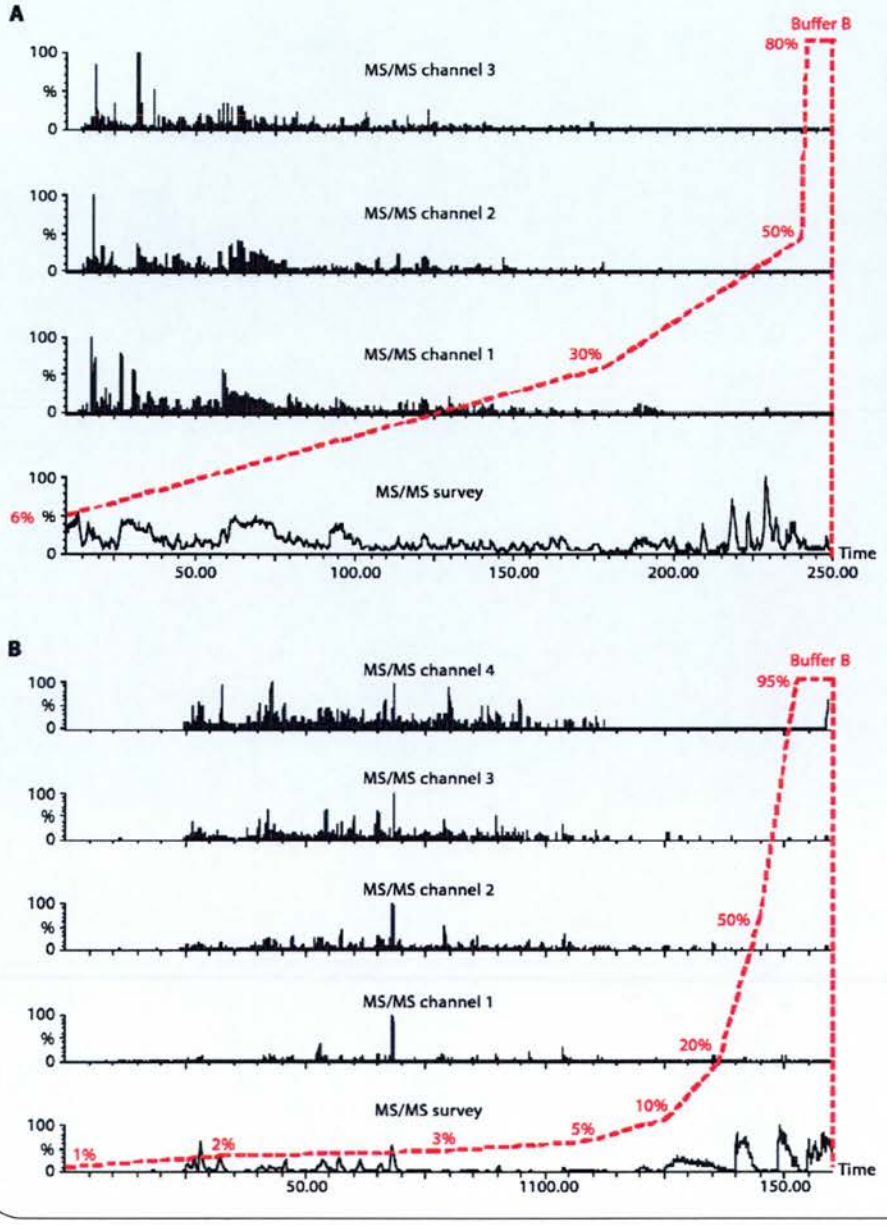


Fig. 5. Base peak intensity (BPI) chromatographs of run 1 for a protein IMAC digest (**A**) and for peptide IMAC purified phosphopeptides (**B**). The red dashed line represents the gradient plot in each chromatogram. A standard gradient was used for the protein IMAC digest (A), with the majority of peptides eluted by 30% Buffer B (32% organic). As most phosphopeptides isolated in a peptide IMAC are multiply phosphorylated and hence quite hydrophilic, a shallow gradient of Buffer B is used (B). It can be seen that the majority of phosphopeptides are eluted before 5% B (3.5% organic).

Appendix 10

PROTOCOL

Data Analysis

We recommend a Mascot V2.0 (Matrix Science) server for iterative searching of a custom, nonidentical, combined human, mouse, and rat protein sequence [International Protein Index (IPI) database (http://www.ebi.ac.uk/IPI/)]. We will next describe the parameters and settings for using the Mascot V2.0. An example of a mass spectrum of a methyl-esterified, diphosphorylated peptide is shown in Fig. 6.

1. Select trypsin with a maximum of two miscleavage sites when using Trypsin Gold (Promega) or sequencing-grade trypsin (Roche); select semi-trypsin with maximum of two miscleavage sites for Promega's sequencing-grade modified trypsin, because this trypsin produces some semi-tryptic fragments.
2. Select correct database with the appropriate taxonomy.
3. Set following variable modifications:

   Carbamyl (N-term)

   Carbamyl (K)

   Methyl ester (C-term) – when peptides are methyl esterified

   Methyl ester (DE) – when peptides are methyl esterified

   Oxidation (M)

   Phospho (ST)

   Phospho (Y)
4. Select monoisotopic mass.
5. Set peptide tolerance at 0.4 Da and MS/MS tolerance at 0.3 Da

   Note: The majority of the molecular weight errors should be less than 0.2 dalton if calibration is correct.
6. Select data format as Micromass (.pkl) file when data file is generated by Proteinlynx in Masslynx; for merged file as described previously, select "Mascot generic."
7. Verify the phosphorylation sites output by Mascot (protein identifications with peptides matching details) by manually checking the spectra using PEAKS software.

Table 2. Liquid chromatography conditions for analysis of a phosphopeptide mixture isolated by a peptide IMAC purification.

Time (min)	Percent A (Recipe 32)	Percent B (Recipe 33)	Comments
0	99	1	Trap and analytical column on-line for gradient separation; MS acquisition in this section
30	98	2	
75	97	3	
105	95	5	
125	90	10	
135	80	20	
145	50	50	
150	5	95	
160	5	95	
161	99	1	Trap and analytical column off-line to equilibrate separately
180	99	1	

Appendix 10

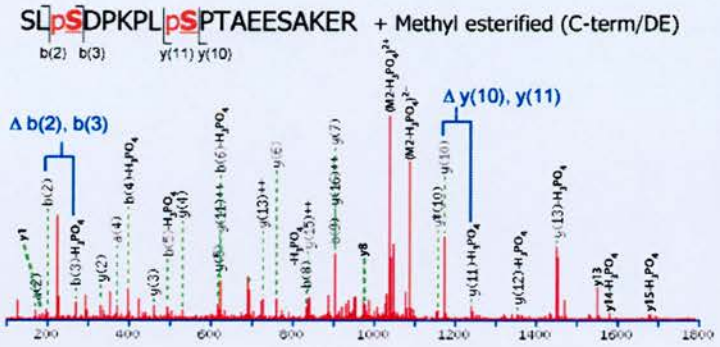


Fig. 6. Mass spectrum of a methyl-esterified phosphopeptide derived from Bassoon protein (Swiss-Prot accession O88737) showing phosphorylated serine residues 2860 and 2866. Phosphorylation of serine 2860 was identified by the mass difference between the b(2) and b(3) ions caused by neutral loss of phosphate in the mass spectrometer. 167 daltons (phosphoserine) minus 98 daltons (H_3PO_4) results in a mass difference of 69 daltons, which is characteristic of dehydroalanine, and therefore phosphoserine prior to neutral loss. Similarly, phosphorylation of serine 2866 was identified by the mass difference between the y(10) and y(11) ions that was caused by neutral loss of phosphate, resulting in a mass difference of 69 daltons.

Troubleshooting

Low Yield in Protein or Peptide IMAC Preparation

In our experience, the protein IMAC Protocol is very reproducible, but a number of parameters should be checked if problems are encountered. The most likely reason for failure or low yields in protein or peptide IMAC purifications is sample or buffer contamination. Small amounts of chelating chemicals, such as EDTA or EGTA, will strip the metal from the resin, and contamination of the buffer with the metal used as the affinity reagent will result in reduced or no phosphoprotein yield. In the event of a low-yield IMAC purification, all buffers should be replaced, and the efficiency of purification monitored by Western blotting or phosphoprotein staining.

Nonphosphorylated Protein Contamination

Reduction of the protein extract concentration or extract to resin ratio may result in nonspecific binding and thus contamination with nonphosphorylated proteins. In addition, contamination by nonspecific protein binding may be observed if the phosphoproteins are eluted from the resin by boiling (as opposed to specific chelating elution) in at 97°C for 5 min in 5× Reducing Sample Buffer (Recipe 9).

Integral or Membrane-Associated Proteins

The applicability of protein IMAC to integral membrane or membrane-associated proteins is limited by their solubility in 6 M urea. We recommend that partial proteolysis with trypsin or endoproteinase Lys-C be performed to release the cytoplasmic and extracellular domains of such proteins for subsequent protein and/or peptide IMAC purification.

Nonspecific Contaminants in the Peptide IMAC Preparation

Traditional peptide IMAC suffers to varying degrees from nonspecific binding of acidic peptides. This can be reduced by the conversion of peptides to corresponding methyl-esters. However, this reaction does not usually go to completion and introduces other problems, such as dilution of a peptide species into modified and unmodified forms. Nonspecific interactions due to hydrophobic and ionic effects can be reduced by performing IMAC at pH 2.5 to 3, at which most glutamic and aspartic acid residues are protonated, and by including up to 30% acetonitrile in the loading buffer, which decreases nonspecific interactions and helps peptide solubilization. We recommend these low pH, acetonitrile-containing conditions with and without methyl-esterification for greatest coverage of a phosphopeptide population in a sample.

Appendix 10

PROTOCOL

Notes and Remarks

It is generally accepted that, currently, no single approach for enrichment and analysis of phosphoproteins or phosphopeptides will enable identification of all phosphorylation sites (*27*). We support the idea that a multifaceted approach such as that we have described in this Protocol is necessary to achieve higher phosphoproteome coverage. It is also clear that a portion of the phosphoproteome is hidden from analysis in basal state conditions. These phosphoproteomic approaches, which have been developed in the last number of years, should be applied to numerous activation and treatment states in a given system to reveal more of the low-abundance and transient phosphoproteome.

References and Notes

1. H. Husi, M. A. Ward, J. S. Choudhary, W. P. Blackstock, S. G. Grant, Proteomic analysis of NMDA receptor-adhesion protein signaling complexes. *Nat. Neurosci.* **3**, 661–669 (2000).
2. C. D. Farr, P. R. Gafken, A. D. Norbeck, C. E. Doneanu, M. D. Stapels, D. F. Barofsky, M. Minami, J. A. Saugstad, Proteomic analysis of native metabotropic glutamate receptor 5 protein complexes reveals novel molecular constituents. *J. Neurochem.* **91**, 438–450 (2004).
3. M. Kim, L. H. Jiang, H. L. Wilson, R. A. North, A. Surprenant, Proteomic and functional evidence for a P2X7 receptor signalling complex. *EMBO J.* **20**, 6347–6358 (2001).
4. T. Bouwmeester, A. Bauch, H. Ruffner, P. O. Angrand, G. Bergamini, K. Croughton, C. Cruciat, D. Eberhard, J. Gagneur, S. Ghidelli, C. Hopf, B. Huhse, R. Mangano, A. M. Michon, M. Schirle, J. Schlegl, M. Schwab, M. A. Stein, A. Bauer, G. Casari, G. Drewes, A. C. Gavin, D. B. Jackson, G. Joberty, G. Neubauer, J. Rick, B. Kuster, G. Superti-Furga, A physical and functional map of the human TNF-alpha/NF-kappa B signal transduction pathway. *Nat. Cell Biol.* **6**, 97–105 (2004).
5. B. Blagoev, S. E. Ong, I. Kratchmarova, M. Mann, Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nat. Biotechnol.* **22**, 1139–1145 (2004).
6. C. Becamel, G. Alonso, N. Galeotti, E. Demey, P. Jouin, C. Ullmer, A. Dumuis, J. Bockaert, P. Marin, Synaptic multiprotein complexes associated with 5-HT(2C) receptors: A proteomic approach. *EMBO J.* **21**, 2332–2342 (2002).
7. C. Hisatsune, H. Umemori, M. Mishina, T. Yamamoto, Phosphorylation-dependent interaction of the N-methyl-D-aspartate receptor epsilon 2 subunit with phosphatidylinositol 3-kinase. *Genes Cells* **4**, 657–666 (1999).
8. P. Greengard, F. Valtorta, A. J. Czernik, F. Benfenati, Synaptic vesicle phosphoproteins and regulation of synaptic function. *Science* **259**, 780–785 (1993).
9. J. D. Sweatt, E. R. Kandel, Persistent and transcriptionally-dependent increase in protein phosphorylation in long-term facilitation of *Aplysia* sensory neurons. *Nature* **339**, 51–54 (1989).
10. M. O. Collins, L. Yu, M. P. Coba, H. Husi, I. Campuzano, W. P. Blackstock, J. S. Choudhary, S. G. Grant, Proteomic analysis of in vivo phosphorylated synaptic proteins. *J. Biol. Chem.* **280**, 5972–5982 (2005).
11. C. Zhang, E. H. Williams, Y. Guo, L. Lum, P. A. Beachy, Extensive phosphorylation of Smoothened in Hedgehog pathway activation. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 17900–17907 (2004).
12. H. Zhang, X. Zha, Y. Tan, P. V. Hornbeck, A. J. Mastrangelo, D. R. Alessi, R. D. Polakiewicz, M. J. Comb, Phosphoprotein analysis using antibodies broadly reactive against phosphorylated motifs. *J. Biol. Chem.* **277**, 39379–39387 (2002).
13. K. M. Loyet, J. T. Stults, D. Arnott, Mass spectrometric contributions to the practice of phosphorylation site mapping through 2003: A literature review. *Mol. Cell. Proteomics* **4**, 235–245 (2005).
14. M. Mann, S. E. Ong, M. Gronborg, H. Steen, O. N. Jensen, A. Pandey, Analysis of protein phosphorylation using mass spectrometry: Deciphering the phosphoproteome. *Trends Biotechnol.* **20**, 261–268 (2002).
15. S. B. Ficarro, M. L. McCleland, P. T. Stukenberg, D. J. Burke, M. M. Ross, J. Shabanowitz, D. F. Hunt, F. M. White, Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat. Biotechnol.* **20**, 301–305 (2002).
16. T. S. Nuhse, A. Stensballe, O. N. Jensen, S. C. Peck, Large-scale analysis of in vivo phosphorylated membrane proteins by immobilized metal ion affinity chromatography and mass spectrometry. *Mol. Cell. Proteomics* **2**, 1234–1243 (2003).
17. H. Shu, S. Chen, Q. Bi, M. Mumby, D. L. Brekken, Identification of phosphoproteins and their phosphorylation sites in the WEHI-231 B lymphoma cell line. *Mol. Cell. Proteomics* **3**, 279–286 (2004).
18. J. A. DeGiorgis, H. Jaffe, J. E. Moreira, C. G. Carlotti Jr., J. P. Leite, H. C. Pant, A. Dosemeci, Phosphoproteomic analysis of synaptosomes from human cerebral cortex. *J. Proteome Res.* **4**, 306–315 (2005).
19. J. C. Trinidad, A. Thalhammer, C. G. Specht, R. Schoepfer, A. L. Burlingame, Phosphorylation state of postsynaptic density proteins. *J. Neurochem.* **92**, 1306–1316 (2005).
20. M. Wilm, M. Mann, Analytical properties of the nanoelectrospray ion source. *Anal. Chem.* **68**, 1–8 (1996).
21. M. R. Larsen, T. E. Thingholm, O. N. Jensen, P. Roepstorff, T. J. Jorgensen, Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol. Cell. Proteomics* **4**, 873–886 (2005).
22. R. K. Carlin, D. J. Grab, R. S. Cohen, P. Siekevitz, Isolation and characterization of postsynaptic densities from various brain regions: enrichment of different types of postsynaptic densities. *J. Cell Biol.* **86**, 831–845 (1980).
23. C. E. Haydon, P. A. Eyers, L. D. Aveline-Wolf, K. A. Resing, J. L. Maller, N. G. Ahn, Identification of novel phosphorylation sites on Xenopus laevis Aurora A and analysis of phosphopeptide enrichment by immobilized metal-affinity chromatography. *Mol. Cell. Proteomics* **2**, 1055–1067 (2003).
24. M. C. Posewitz, P. Tempst, Immobilized gallium(III) affinity chromatography of phosphopeptides. *Anal. Chem.* **71**, 2883–2892 (1999).
25. M. O. Collins, L. Yu, H. Husi, W. P. Blackstock, J. S. Choudhary, S. G. N. Grant, unpublished results.
26. J. Kyte, R. F. Doolittle, A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–132 (1982).
27. H. Steen, A. Pandey, J. S. Andersen, M. Mann, Analysis of tyrosine phosphorylation sites in signaling molecules by a phosphotyrosine-specific immonium ion scanning method. *Sci. STKE* **2002**, PL16 (2002).

28. This work was supported by grants from the Wellcome Trust (to H.H., J.S.C., L.Y., M.O.C., and S.G.N.G.) and the Biotechnology and Biological Sciences Research Council (to M.O.C.).

Citation: M. O. Collins, L. Yu, H. Husi, W. P. Blackstock, J. S. Choudhary, S. G. N. Grant, Robust enrichment of phosphorylated species in complex mixtures by sequential protein and peptide metal-affinity chromatography and analysis by tandem mass spectrometry. *Sci. STKE* **2005**, pl6 (2005).

Appendix 11

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 280, No. 7, Issue of February 18, pp. 5972–5982, 2005
Printed in U.S.A.
© 2005 by The American Society for Biochemistry and Molecular Biology, Inc.

Proteomic Analysis of *in Vivo* Phosphorylated Synaptic Proteins*[S]

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 EH8 9JZ, 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

In the nervous system, protein phosphorylation is an essential feature of synaptic function. Although protein phosphorylation is known to be important for many synaptic processes and in disease, little is known about global phosphorylation of synaptic proteins. Heterogeneity and low abundance make protein phosphorylation analysis difficult, particularly for mammalian tissue samples. Using a new approach, combining both protein and peptide immobilized metal affinity chromatography and mass spectrometry data acquisition strategies, we have produced the first large scale map of the mouse synapse phosphoproteome. We report over 650 phosphorylation events corresponding to 331 sites (289 have been unambiguously assigned), 92% of which are novel. These represent 79 proteins, half of which are novel phosphoproteins, and include several highly phosphorylated proteins such as MAP1B (33 sites) and Bassoon (30 sites). An additional 149 candidate phosphoproteins were identified by profiling the composition of the protein immobilized metal affinity chromatography enrichment. All major synaptic protein classes were observed, including components of important pre- and postsynaptic complexes as well as low abundance signaling proteins. Bioinformatic and *in vitro* phosphorylation assays of peptide arrays suggest that a small number of kinases phosphorylate many proteins and that each substrate is phosphorylated by many kinases. These data substantially increase existing knowledge of synapse protein phosphorylation and support a model where the synapse phosphoproteome is functionally organized into a highly interconnected signaling network.

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)[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 diverse 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), *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 analyzes are being explored (13, 14).

Here, we describe the use of a combination of cellular frac-

* This work was spported by the Wellcome Trust (to H. H., J. S. C., L. Y., M. O. C., M. P. C., and S. G. N. G.) and the Biotechnology and Biological Sciences Research Council (to M. O. C.). 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.jbc.org) contains an additional figure and four tables.

** To whom correspondence should be addressed: Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire CB10 1SA, United Kingdom. Tel.: 44-1223-494-908; Fax: 44-1233-494-919; E-mail: sg3@sanger.ac.uk.

[1] The abbreviations used are: MS, mass spectrometry; IMAC, immobilized metal affinity chromatography; IDA, iminodiacetic acid; TRA, targeted repeat analysis; NMDA, *N*-methyl-D-aspartate; NRC, NMDA receptor complex; IP_3, inositol trisphosphate; 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.

This paper is available on line at http://www.jbc.org

Appendix 11

tionation 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 × *g*, and the resulting supernatant was centrifuged again for 30 min at 16,000 × *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 × *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 × *g*. The pellet 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 nitrilotriacetic acid (Qiagen) chelating groups were charged with $GaCl_3$ or $FeCl_3$. Synaptosomal proteins (12.5 mg) were solubilized in 6 M 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 M urea, 50 mM Tris acetate) to base line, and the phosphoproteins were specifically eluted with buffer B (6 M 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 μg of protein was diluted with buffer C (1 M urea, 0.125 M thiourea, 5% (v/v) CH_3CN, and 0.1 M NH_4HCO_3) to a final volume of 400 μl. Trypsin (sequence grade; Roche Applied Science) was added to the sample in a ratio of 1:20, and the mixture was incubated at 37 °C for 2 h and then dried in a SpeedVac (Thermo Life Science). The sample was reconstituted in 50 mM NH_4HCO_3, and one-quarter was injected for each on-line LC-MS/MS analysis.

Double IMAC of Urea-soluble Synaptosomal Fraction—3 mg of protein IMAC purified sample was digested with modified porcine or gold trypsin (Promega) in a ratio of 1:20 in buffer D (1 M urea and 25 mM NH_4HCO_3) at 37 °C for 4 h. The resultant digest was desalted and dried, and methyl esterification was performed with 2 M methanolic HCl (10) when needed. Self Pack POROS® 20 MC media (Applied Biosystems) for phosphopeptide purification was charged with $GaCl_3$ as described above for the IDA/nitrilotriacetic acid resins. Peptide digests (with and without methyl esterification) were reconstituted in buffer E (equal volumes of acetonitrile, methanol, and water, pH 2.5–3). 1 ml of this peptide mixture was incubated with 200 μl of POROS-Ga slurry for 1 h at room temperature. The resin was then loaded into a spin column and washed with 10 volumes of buffer E. Phosphopeptides were eluted with 2 × 100 μl of 200 mM Na_2HPO_4.

Peptide IMAC of Whole and Urea-insoluble Synaptosomal Fractions—2.5 mg of the 6 M urea insoluble pellet left after removal of the supernatant was digested with trypsin in buffer F (2 M urea and 25 mM NH_4HCO_3). The resultant digest was desalted, methyl-esterified, and subjected to a peptide IMAC step as described previously. Similarly, 6 mg of whole synaptosomal proteins were digested in buffer F, desalted, methyl-esterified, and subjected to a peptide IMAC step.

Phosphoprotein Staining and Image Analysis—IMAC-enriched phosphoprotein samples were 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). *Peppermint Stick* 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 (Nonlinear Dynamics).

TABLE I
LC gradient for separation of phosphoprotein digest
Solvent A was 95% H_2O, 5% ACN, 0.1% formic acid; solvent B was 95% ACN, 5% H_2O, 0.1% formic acid. Flow rate through the columns in the separation was 200–250 nl/min.

Solvent	0 min	180 min	240 min	242 min	247 min
A (%)	94	30	50	20	20
B (%)	6	70	50	80	80

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 with 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 (PO_3^-) with automatic switching to positive mode (+ESI) for MS/MS 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—Jerini 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 immobilized to the glass slide surface. Each peptide was present in triplicate on the chip, and seven full-length proteins that are capable of being phosphorylated were also included. Negative control peptides for each phosphorylation 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 (Abgene 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 dithiothreitol, 100 μM

Appendix 11

TABLE II
LC gradient for IMAC-enriched phosphopeptide analysis

Solvent A was 100% H_2O, 0.1% formic acid; solvent B was 70% ACN, 30% H_2O, 0.1% formic acid. Flow rate through the columns in the separation was 150 nl/min.

Solvent	0	30	75	105	125	135	145	150	160
A (%)	99	98	97	95	90	80	50	5	5
B (%)	1	2	3	5	10	20	50	95	95

ATP, 15 mM $MgCl_2$, and 10 μM [γ-^{32}P]ATP). Recombinant active kinases (Upstate Biotechnology, Inc., Lake Placid, NY) (3 μg of PKA catalytic subunit (with 2 μM cAMP), 2 μg of Akt1 (Δ 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, 0.5 mM $CaCl_2$, 0.1 mg/ml phosphatidylserine, 25 μg/ml diacylglycerol).

After a 45-min incubation at 32 °C, the peptide microarrays were washed six times, alternating between 0.1 M phosphoric acid and distilled water. γ-^{32}P incorporation in the immobilized peptide spots was detected on a Typhoon 8600 PhosphorImager (50-μm resolution; Amersham Biosciences). Image analysis and signal quantification was carried out using ImageQuant TL (Amersham Biosciences), and positive signals were defined after background subtraction.

Bioinformatic Analysis—All proteins detected in this study were classified according to Swiss-Prot keywords (available on the World Wide Web at ca.expasy.org/sprot/). Known phosphoproteins in Supplementary Table I were annotated by extensive literature mining of PubMed. Scansite (available on the World Wide Web at scansite.mit.edu/) was used in high stringency mode to predict if proteins in Supplementary Table I were likely to be phosphorylated, and responsible kinases are indicated. TMHMM Server version 2.0 (available on the World Wide Web at www.cbs.dtu.dk/services/TMHMM/) was used to predict the occurrence of transmembrane helices. Scansite and NetPhosK (available on the World Wide Web at www.cbs.dtu.dk/services/NetPhos/) were both used for predicting phosphorylation sites by cognate kinases and on those peptides where the phosphorylation sites could not be unambiguously assigned by MS data (Supplemental Table III). Scansite was also used to check the presence of phosphorylation sites in Pfam protein domains and to predict whether phosphorylation sites were localized in phosphodependent interaction domains (Supplemental Table III). Graphical illustrations of the location of identified sites in relation to protein domain structure, three-dimensional structure, and information on homologs and paralogs of identified phosphorylation sites can be found on the World Wide Web at www.ppo4.org/phospho/synaptosome.html.

Network Construction—Protein-protein interactions for the *N*-methyl-D-aspartate (NMDA) receptor complex proteins were obtained from PPID (available on the World Wide Web at www.PPID.org). Each protein (node) that had interaction (edge) information was used to plot a network graph using InterViewer, a network graphing program. Kinase-substrate edges were manually superimposed onto the protein-protein interaction network.

RESULTS

Several approaches capable of identifying large numbers of *in vivo* phosphorylated proteins have been used to functionally annotate the mouse synaptic proteome. A summary of this is shown in Fig. 1*A*. In the first approach, a mouse synaptosomal preparation was digested with trypsin and phosphopeptides captured by peptide IMAC. The enriched sample was analyzed by nano-flow LC-MS/MS. Despite the enrichment, the majority of identifications corresponded to unphosphorylated peptides. In the second approach, the synaptosomes were first divided into urea soluble and insoluble fractions, the latter containing mostly membrane-bound components. Peptide IMAC and MS was performed on the insoluble fraction in a comparable manner to the first approach. For the urea-soluble fraction, a new strategy was adopted, entailing an additional stage of enrichment by capture of the phosphoproteins. A two-pronged route was taken for subsequent analysis of the trypsin digest of this enriched fraction. The first route is based on rigorous MS using (*a*) LC-MS/MS peptide identification based on a TRA approach described under "Experimental Procedures" and (*b*) precursor ion scanning to identify phosphopeptides prior to sequencing. The second route used a further peptide IMAC followed by LC-MS/MS. Using these combined approaches and implementing a previously unreported strategy of sequential protein and peptide IMAC, we have identified 228 potential synaptic phosphoproteins (Supplementary Table I) and characterized over 350 phosphopeptides containing 331 sites of phosphorylation (Supplemental Tables II and III).

Synaptic Phosphoprotein Analysis

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-nitrilotriacetic acid (a quadradentate ligand) and two metal ions, $FeCl_3$ and $GaCl_3$ (Fig. 1*B*), 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 Ga^{3+} resin showed more effective depletion of phosphoprotein from the unbound fraction together with stronger phosphostaining of the purified sample (Fig. 1*C*) 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 (Supplementary Tables I and II). TRA extended this by an additional 30 protein identifications and 21 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. 2*B* 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. 2*A* 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

Appendix 11

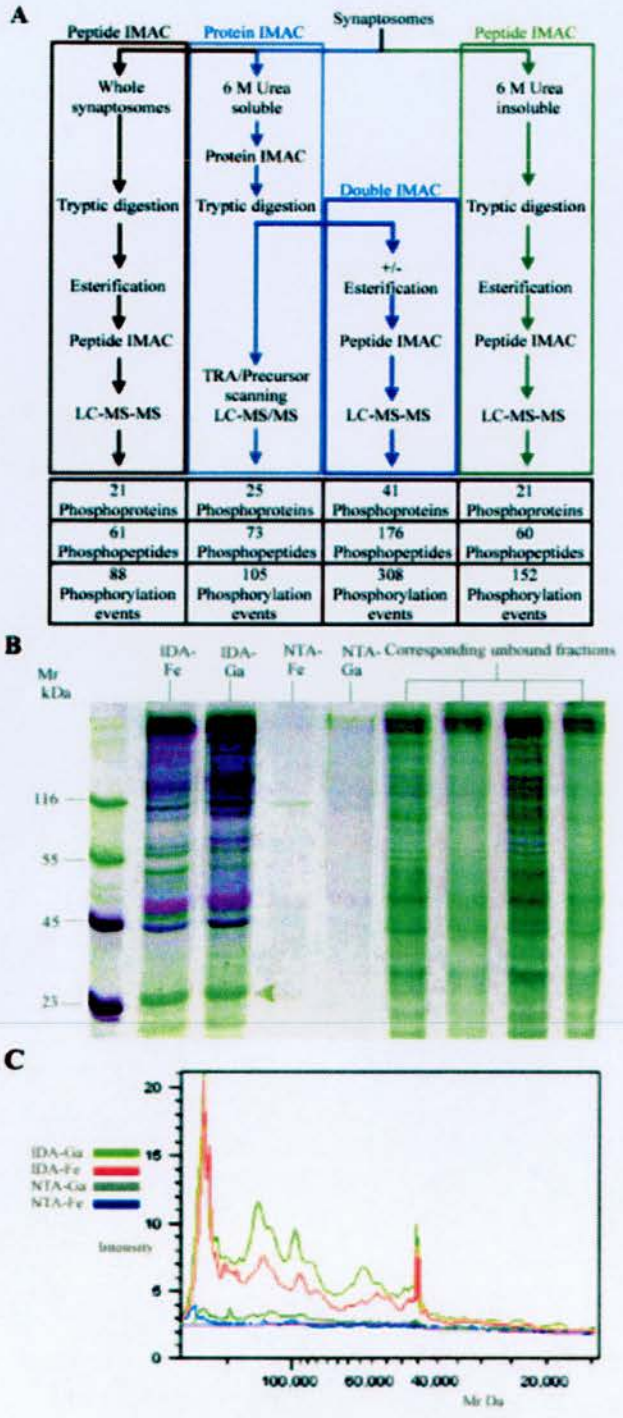


FIG. 1. *A*, schematic representation of the protein and peptide purification strategy used. Alternative starting material was derived from the differential solubility of synaptosomes in 6 M urea. Large scale protein identification was achieved by TRA LC-MS/MS analysis of the IMAC-purified phosphoproteins from the urea-soluble fraction with some of the most abundant phosphopeptides characterized at this step. Further phosphopeptide identification was achieved by sequential protein and peptide IMAC purification with esterification and single peptide IMAC with esterification of the urea-insoluble fraction and of whole synaptosomal digests. *B*, development of IMAC phosphoprotein isolation strategy. Overlay of Pro-Q diamond phosphoprotein/Ruby total protein stains of resin/metal IMAC combinations used to selectively isolate phosphoproteins from urea-soluble synaptosomal extract. The corresponding unbound fractions are included for comparison. 20% 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 resins 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.

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. 1*A*). 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. 3*A*).

Complementarity of Analytical Approaches—Inspection of the summarized data (Figs. 2*A* 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. 1*A*). 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

Appendix 11

A

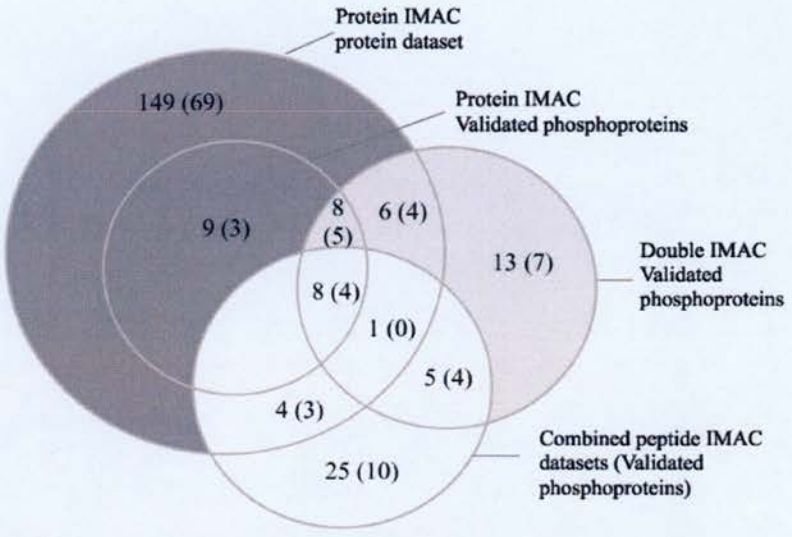


FIG. 2. *A*, Venn diagram of phosphoproteins detected in the protein, double, and peptide IMAC protocols. Data from peptide IMAC experiments on whole synaptosomal digests and urea-insoluble synaptosomal digests were combined to simplify the comparison. *Values* in *brackets* indicate the number of phosphoproteins 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.

B

Protein class	PSD Proteins	NRC Proteins	Validated Phosphoproteins	Total
Channels/receptors	4	0	7	7
Transporters/transporting ATPases	5	2	9	12
Scaffolding and adaptors	17	10	17	21
Kinases	14	5	17	19
Phosphatases	4	1	2	4
Small G-proteins and modulators	12	0	8	18
Other signaling molecules	6	5	7	10
Other enzymes	12	5	8	20
Cell adhesion and cytoskeletal	55	16	51	69
Nuclear/transcriptional	0	0	6	10
Translational	2	0	2	3
Novel/uncharacterised proteins	3	0	14	35
Total	**134**	**44**	**148**	**228**

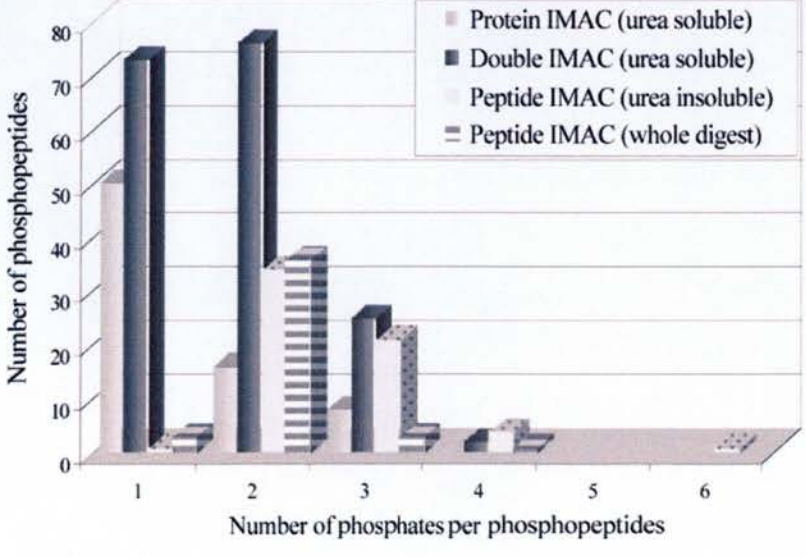


FIG. 3. **Distribution of the number of phosphates present per phosphopeptide identified.** The values are representative of all phosphopeptides detected by each approach and therefore contain some peptides covering overlapping sequences in proteins but with differing numbers of phosphates attached.

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

Appendix 11

representative coverage, with very similar preferences for mono- and diphosphorylated 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 multiphosphorylated 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 multiphosphorylated 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 multiphosphorylated 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 149 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. 2*A*) 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 ^{44}IG(S)(T)(T)NPFLDIPHDPNAAVYK^{64} 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 Ser^{46}, 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. 1*A*) 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. 1*B*). 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

Appendix 11

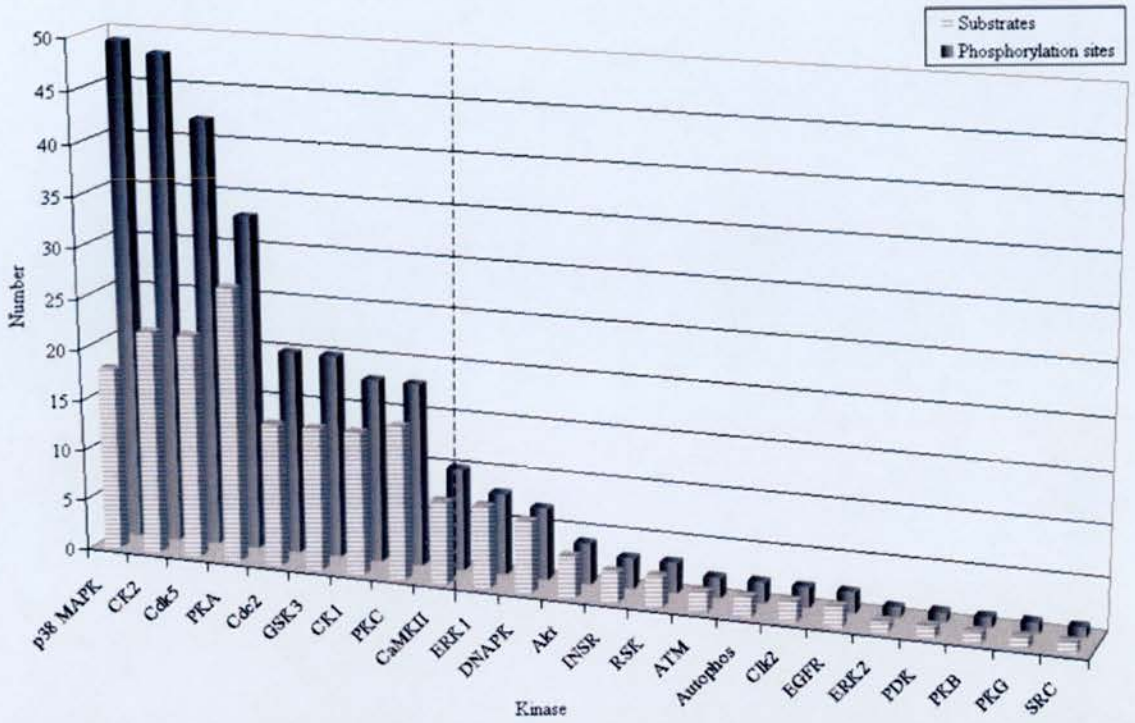


FIG. 4. **Predicted kinase multiplicity for 289 defined synaptic phosphorylation sites.** Scansite and NetPhosK kinase predictions and in a few cases data from the literature were used to evaluate the number of phosphorylation sites and distinct substrates for 23 kinases on 289 sites in synapse proteins. The total number of predicted sites (*filled bars*) and substrate proteins (*cross-hatched bars*) are shown on the *x* axis for a given kinase. The *dashed vertical line* between CamKII and extracellular signal-regulated kinase divides the kinases into two groups; the *left-hand set* has kinase/substrate ratio of >1, and the *right-hand set* has a kinase/substrate ratio of 1.

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 Ca^{2+} 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 (Ser^{414}), which was also found in MS analyses of purified NRC complexes.[2] 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 β_1 sheet. Structural analysis of the SH3 and guanylate kinase domains of PSD-95 (27, 28) showed that there is an intermo-

[2] H. Husi, J. S. Choudhary, W. P. Blackstock, and S. G. N. Grant, unpublished results.

Appendix 11

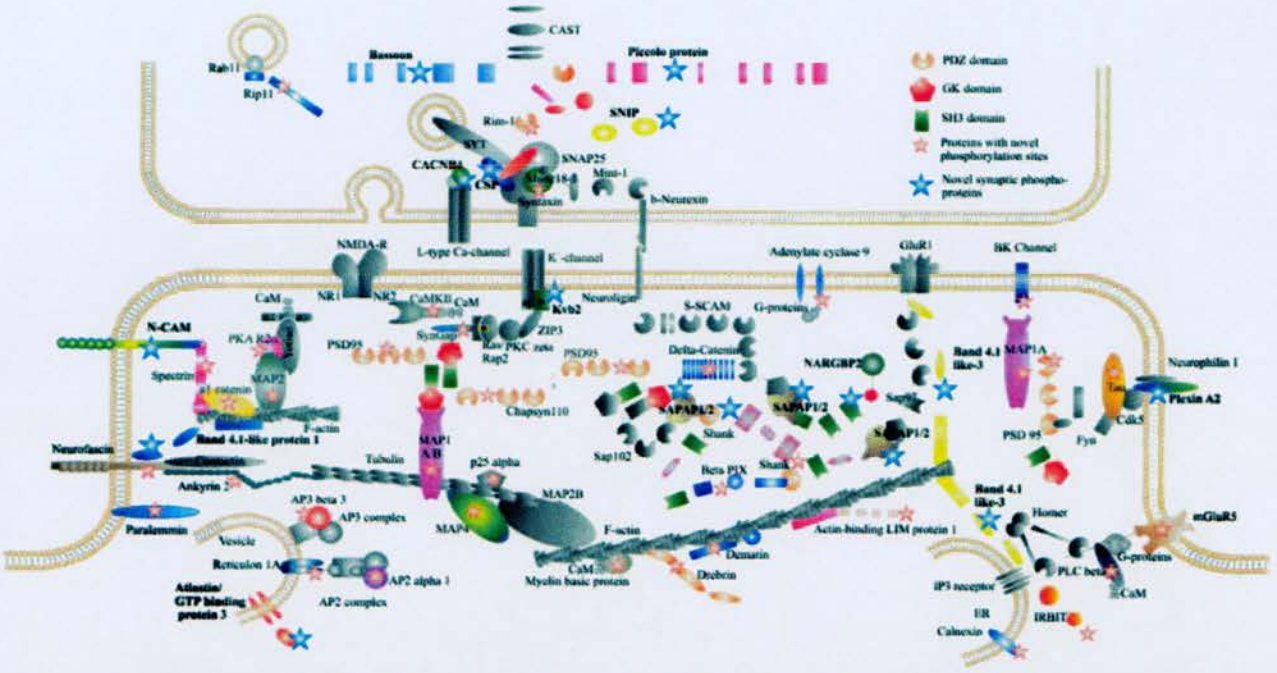


FIG. 5. **Schematic illustration of phosphoproteins and their interaction partners at the synapse.** Selected phosphoproteins characterized in this study were used to construct a synaptic interaction diagram. Protein-protein interaction information was gathered from PPID (available on the World Wide Web at www.ppid.org). All phosphoproteins are *colored*, and proteins *shaded in gray* are known synaptic proteins and are included for protein-protein interaction completeness. Proteins with novel phosphorylation sites are marked with a *red star*, and novel phosphoproteins are marked with a *blue star*.

lecular 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. Tyr^{432} is located in this hydrophobic core mediating this interaction and is itself interacting with His^{461}. Introduction of a negative charge by phosphorylation of Tyr^{432} 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 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.

Appendix 11

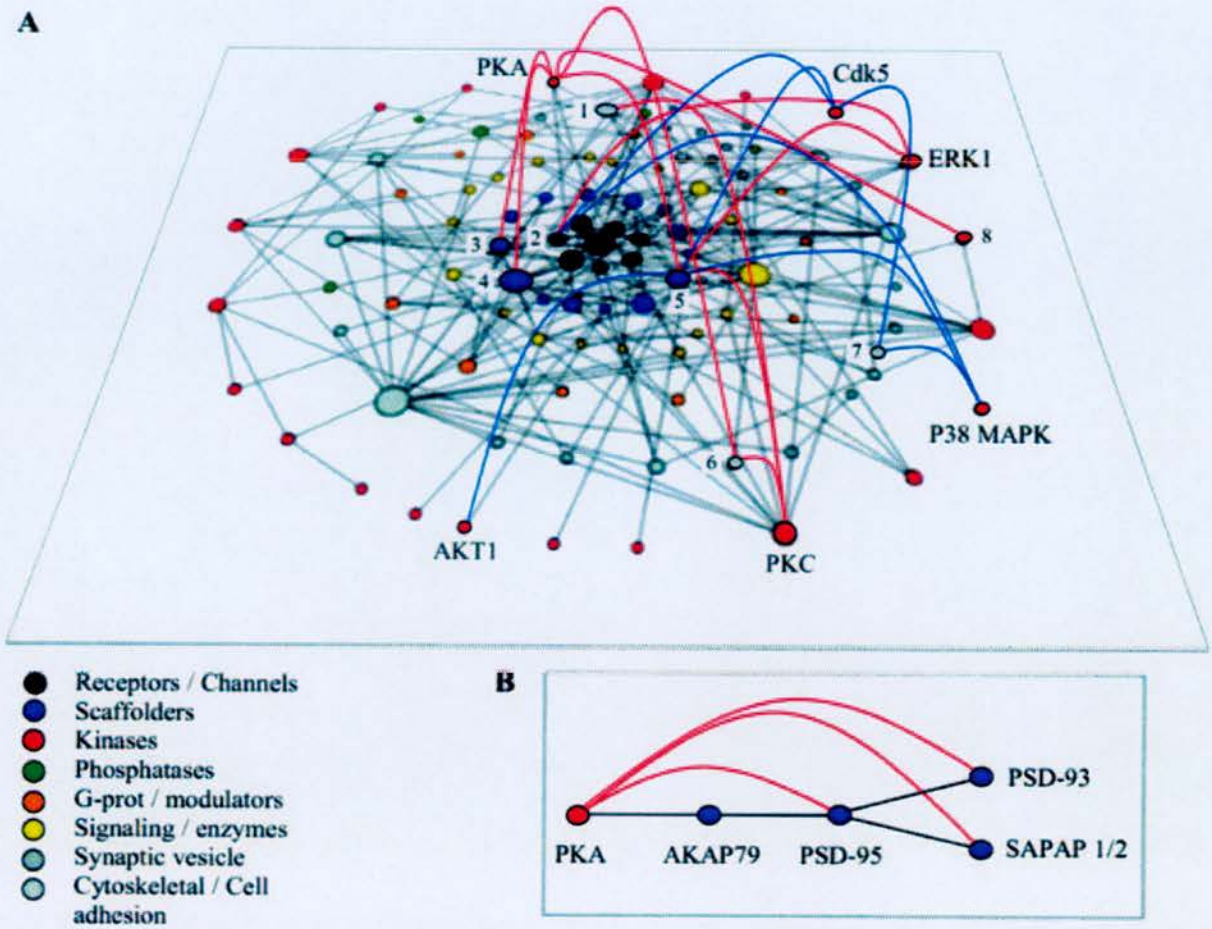


FIG. 6. *A*, network illustration of protein-protein interactions and kinase-substrate interactions in the NRC. Protein-protein interaction data (248 binary interactions from PPID) for 104 NRC proteins was used to construct a protein-protein interaction network. Each node (*colored circle* representing a protein) is *color-coded* according to its protein class. Kinase-substrate data from MS-identified phosphorylation sites and peptide array experiments, for eight NRC proteins, were superimposed onto the protein-protein interaction network. *Red network edges* (*lines* connecting two proteins) represent kinases present within the NRC, which phosphorylate NRC substrates, and *blue network edges* represent kinases not presently known to be present within the NRC, which phosphorylate NRC substrates. Substrates are numbered as follows: *1*, MAP2; *2*, mGlur5; *3*, PSD-93/Chapsyn110; *4*, PSD-95; *5*, SAPAP1 and -2; *6*, myelin basic protein; *7*, Bassoon; *8*, CaMK2B. *B*, illustration of path lengths from PKA to its NRC substrates (PSD-95, PSD-93/Chapsyn110, and SAPAP1 and -2). These PKA substrates are important synaptic scaffolders, which are closely associated with each other, and their proximity to this kinase within the NRC would imply a physiological context for their phosphorylation.

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 Ser^{493} 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 Ser^{144} 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,

Appendix 11

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. 6*A*). 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. 6*B*). 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 synaptic 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 schizophrenia 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. Peng, J., Kim, 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., Casetta, B., Jimenez, C. R., Gouwenberg, Y., Gundelfinger, E. D., Smalla, K. H., and Smit, 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. Hisatsune, C., Umemori, H., Mishina, M., and Yamamoto, T. (1999) *Genes Cells* **4,** 657–666
7. Greengard, P., Valtorta, F., Czernik, 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., Shabanowitz, 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. C. (2003) *Mol. Cell Proteomics* **11,** 1234–1243
12. Shu, H., Chen, S., Bi, Q., Mumby, M., and Brekken, D. L. (2004) *Mol. Cell Proteomics* **3,** 279–286
13. 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) *Proc. Natl. Acad. Sci. U. S. A.* **101,** 12130–12135
14. Chang, E. J., Archambault, V., McLachlin, D. T., Krutchinsky, A. N., and Chait, B. T. (2004) *Anal. Chem.* **76,** 4472–4483
15. Carlin, R. K., Grab, D. J., Cohen, R. S., and Siekevitz, P. (1980) *J. Cell Biol.* **86,** 831–845
16. Wenschuh, H., Volkmer-Engert, R., Schmidt, M., Schulz, M., Schneider-Mergener, J., and Reineke, U. (2000) *Biopolymers* **55,** 188–206
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,** 2394–2405
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., Zhai, R. G., Dresbach, T., Bresler, T., Torres, V. I., Gundelfinger, E. D., Ziv, N. E., and Garner, C. C. (2003) *Neuron* **38,** 237–252
22. Lonart, G., Schoch, S., Kaeser, P. S., Larkin, C. J., Sudhof, T. C., and Linden, D. J. (2003) *Cell* **115,** 49–60
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 Borg, J. P. (2003) *Science's STKE* http://stke.sciencemag.org/cgi/content/full/OC_sigtrans;2003/179/re7
26. Nada, S., Shima, T., Yanai, H., Husi, H., Grant, S. G., Okada, M., and Akiyama, T. (2003) *J. Biol. Chem.* **278,** 47610–47621
27. McGee, A. W., Dakoji, S. R., Olsen, O., Bredt, D. S., Lim, W. A., and Prehoda, K. E. (2001) *Mol. Cell* **8,** 1291–1301
28. Tavares, G. A., Panepucci, E. H., and Brunger, A. T. (2001) *Mol. Cell* **8,** 1313–1325
29. Colledge, M., Dean, R. A., Scott, G. K., Langeberg, L. K., Huganir, R. L., and Scott, J. D. (2000) *Neuron* **27,** 107–119
30. Gardoni, F., Mauceri, D., Fiorentini, C., Bellone, C., Missale, C., Cattabeni, F.,

Appendix 11

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. (1990) *J. Biol. Chem.* **265,** 21561–21566
32. Farr, C. D., Gafken, P. R., Norbeck, A. D., Doneanu, 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., Tsutsumi, R., Itahana, H., Hayashibe, S., Yamaguchi, T., and Okada, M. (1998) *J. Biol. Chem.* **273,** 17381–17385
34. Ando, H., Mizutani, A., Matsu-ura, T., and Mikoshiba, K. (2003) *J. Biol. Chem.* **278,** 10602–10612
35. Tu, J. C., Xiao, B., Yuan, J. P., Lanahan, A. A., Leoffert, 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., Kschischo, M., Dong, 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 Tsai, L. H. (2003) *Neuron* **40,** 471–483
41. Kishimoto, A., Nishiyama, K., Nakanishi, 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

Journal of Neurochemistry, 2005 doi:10.1111/j.1471-4159.2005.03507.x

SPECIAL ISSUE

Molecular characterization and comparison of the components and multiprotein complexes in the postsynaptic proteome

Mark O. Collins,*,‡ Holger Husi,‡ Lu Yu,† Julia M. Brandon,* Chris N. G. Anderson,* Walter P. Blackstock,†,§ Jyoti S. Choudhary† and Seth G. N. Grant*,‡

**Genes to Cognition and †Proteomic Mass Spectrometry, The Wellcome Trust Sanger Institute, Hinxton, UK*

‡Division of Neuroscience, University of Edinburgh, Edinburgh, UK

§Molecular Biology and Biotechnology, University of Sheffield, Sheffield, UK

Abstract

Characterization of the composition of the postsynaptic proteome (PSP) provides a framework for understanding the overall organization and function of the synapse in normal and pathological conditions. We have identified 698 proteins from the postsynaptic terminal of mouse CNS synapses using a series of purification strategies and analysis by liquid chromatography tandem mass spectrometry and large-scale immunoblotting. Some 620 proteins were found in purified postsynaptic densities (PSDs), nine in AMPA-receptor immuno-purifications, 100 in isolates using an antibody against the NMDA receptor subunit NR1, and 170 by peptide-affinity purification of complexes with the C-terminus of NR2B. Together, the NR1 and NR2B complexes contain 186 proteins, collectively referred to as membrane-associated guanylate kinase-associated signalling complexes. We extracted data from six other synapse proteome experiments and combined these with our data to provide a consensus on the composition of the PSP. In total, 1124 proteins are present in the PSP, of which 466 were validated by their detection in two or more studies, forming what we have designated the Consensus PSD. These synapse proteome data sets offer a basis for future research in synaptic biology and will provide useful information in brain disease and mental disorder studies.

Keywords: α-Amino-3-hydroxy-5-methylisoxazole-4-propionate, mass spectrometry, *N*-methyl-D-aspartate; postsynaptic density, synapse.

J. Neurochem. (2005) 10.1111/j.1471-4159.2005.03507.x

An important goal of cognitive science is to identify neural mechanisms for processing information. The synapse both transmits information between cells and processes it by detecting specific patterns of neural activity and converting this electrical activity into intracellular biochemical events that change the properties of the neuron (Greengard 2001; Kandel 2001). In recent years it has become clear that synapses, like other cell–cell interactions in metazoans, utilize signal transduction complexes and pathways with a high degree of molecular complexity and cross-talk (Pawson and Nash 2003; Sheng and Kim 2002). A major challenge is to devise strategies that extract emergent physiological properties and simple biological principles from highly complex genomic and proteomic data sets, and shed light on both mechanisms and disease processes. The synapse proteome is a suitable prototype to explore these general issues for several reasons. First, it contains a highly localized set of proteins found in dendritic spines. Second, an important role for signalling complexes and pathways has

Received June 24, 2005; revised manuscript received August 5, 2005; accepted August 26, 2005.

Address correspondence and reprint requests to Seth G. N. Grant, The Wellcome Trust Sanger Institute, Hinxton, CB10 1SA, UK.
E-mail: sg3@sanger.ac.uk

Abbreviations used: AMPA, α-amino-3-hydroxy-5-methylisoxazole-4-propionate; ARC, AMPA receptor complex; cPSD, Consensus postsynaptic density; ICAT, Isotope-coded affinity tag; LC–MS/MS, liquid chromatography tandem mass spectrometry; MAGUK, membrane-associated guanylate kinase; MASC, MAGUK-associated signalling complex; (m)GluR, (metabotropic) glutamate receptor; MS, mass spectrometry; NR, NMDA receptor; NRC, NMDA receptor complex; PSD, postsynaptic density; PSP, postsynaptic proteome; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis.

© 2005 International Society for Neurochemistry,

Appendix 12

been established. Third, signalling can be studied using patterns of action potentials and, finally, genetic and pharmacological perturbations result in behavioural changes. However, the molecular complexity and global organization is not well understood.

The postsynaptic density (PSD) is a structural component of the synapse that can be observed under the electron microscope, beneath the postsynaptic membrane (Cotman *et al.* 1974). Recent proteomic efforts have given an insight into the complexity of the PSD and its subcomponents (Husi and Grant 2001; Walikonis *et al.* 2000; Yamauchi 2002; Yoshimura *et al.* 2004; Jordan *et al.* 2004; Peng *et al.* 2004; Li *et al.* 2004; Collins *et al.* 2005). Advances in mass spectrometry (MS) techniques have enabled large-scale investigations of the protein content of organelles with similar complexity to the PSD (for reviews see Jeong *et al.* 2000; Taylor *et al.* 2003).

The neurotransmitter glutamate activates synaptic plasticity primarily via the ionotropic NMDA receptor and metabotropic (mGluR) receptors. This leads to an increased Ca^{2+} level in the dendritic spine and signal transduction to α-amino-3-hydroxy-5-methylisoxazole-4-propionate (AMPA) receptors and other effector mechanisms. The cytoplasmic C-termini of NMDA receptor subunits (NR2A/ε1, NR2B/ε2) bind PDZ domains of PSD-95, a membrane-associated guanylate kinase (MAGUK) protein. Previous proteomic analysis of NMDA receptor–PSD-95 complexes identified 77 proteins including mGluR receptors, whereas AMPA receptors were found in different complexes (Husi *et al.* 2000). These complexes are embedded in the postsynaptic terminal of excitatory synapses with other postsynaptic proteins.

Here we present a large-scale proteomic analysis of the postsynaptic proteome (PSP) of mouse brain. We have performed a MS-based analysis of MAGUK-associated signalling complexes (MASCs), AMPA receptor complexes (ARCs) and isolated PSDs, and validated a set of MASC and PSD proteins by large-scale immunoblotting. In addition to generating novel data, we have systematically compiled data from publicly available sources to produce a comprehensive data set of the composition of the PSP. This includes data on the PSD and various subcomponents, namely NMDA receptor complexes (NRCs), ARCs and MASCs.

Experimental procedures

Biochemistry

MASCs were isolated using a peptide affinity method modeled on the PDZ binding domain of NR2 subunits as reported elsewhere (Husi *et al.* 2000). ARCs were isolated using immuno-precipitation of AMPA-receptor subunits with an anti-GluR2 antibody (BD Biosciences, Oxford, UK). PSD fractions were prepared as described previously (Carlin *et al.* 1980). Protein samples were separated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and stained with colloidal Coomassie blue. The entire gel lane was excised into 42 individual protein bands, reduced, alkylated and digested with trypsin.

MS

The resulting peptide mixtures were analysed by on-line liquid chromatography tandem MS (LC–MS/MS) to generate peptide sequence information (Link *et al.* 1999). Chromatographic separations of the peptide mixture were performed on a 180-μm (intensely stained bands) and 75-μm (weakly stained bands) PepMap column using an Ultimate LC system (LC Packings, Amsterdam, The Netherlands) delivering a gradient of formic acid (0.05%) and acetonitrile. The LC column was connected directly to a micro-electrospray interface on a Q-TOF mass spectrometer (Micromass Ltd, Manchester, UK). The mass spectrometer operated in automated function switching mode, selecting precursor ions based on intensity for peptide sequencing by tandem MS. Several hundred MS/MS spectra could be generated per run allowing the analysis of complex mixtures without any prior interpretation. The search engine used in this work was Mascot (http://www.matrixscience.com/). Data were reduced to MassLynx 'PKL' format peak lists before searching. Proteins were identified based on matching the MS/MS data with mass values calculated for selected ion series of a peptide. The searches on an in-house non-identical protein database were performed without applying any constraints on molecular weight or species of origin. Most proteins were identified with several peptide matches, although a few proteins were assigned on the basis of a single peptide provided that a near-complete peptide sequence had been obtained.

Immunoblotting

Protein samples were subjected to reducing SDS–PAGE and transferred to polyvinyldifluoride membrane (Bio-Rad, Hercules, CA, USA) at 4°C for 90 min at 75 V in 10% (v/v) methanol, 10 mM 3-(cyclohexylamino)-1-propane-sulfonic acid (CAPS), pH 11.0. Dilution of primary antibodies was between 1 : 100 and 1 : 1000, depending on the quality of the IgGs. Detection of signals was carried out using peroxidase-linked secondary IgGs and enhanced chemiluminescence.

Results

Composition of the PSP

We isolated MASCs using a previously described peptide affinity method (Husi and Grant 2001). We used a peptide corresponding to the six C-terminal residues of NR2B that binds MAGUK proteins (including PSD-95), as a specific affinity ligand to isolate NMDA receptor–MAGUK-containing complexes. Purified complexes were separated by SDS–PAGE, and bands were excised and digested and analysed by MS. Western blotting of peptide-purified complexes for candidate proteins was also performed. These complexes are similar in composition to NRCs and MAGUK complexes isolated with antibodies (Husi and Grant 2001; Husi *et al.* 2000). Of the 100 proteins previously identified in the NRC

© 2005 International Society for Neurochemistry, *J. Neurochem.* (2005) 10.1111/j.1471-4159.2005.03507.x

Appendix 12

by analysis of anti-NR1 immuno-purifications, 84 were also found by the peptide approach. In addition, 86 new proteins were identified by the peptide approach, many of which were in the molecular weight range obscured by the contaminating antibody chains in the immuno-purification strategy. Because of the large overlap, we combined data sets of proteins found in both the NRC and the MAGUK-pep (NR2B peptide-affinity purification) under the term MASCs. Ninety-five MASC proteins were detected by western blotting of purified complexes and 35 of these were validated by MS data. The MASC set represents a comprehensive account of NMDA receptor–MAGUK complexes identified by large-scale western blotting and MS-based analyses of both antibody and peptide affinity-based purifications.

We performed immuno-precipitation of the ARC with an antibody directed against the GluR2 subunit and found a significantly smaller complex of nine proteins by MS analysis, which included all four AMPA receptor subunits.

In order to gain an overall view of the PSP and to validate the components of the MASCs and ARCs, we performed a systematic analysis of purified PSDs. PSD proteins were separated by SDS–PAGE (Fig. 1) and 37 gel slices were analysed by LC–MS/MS. Some 7402 peptides were identified, corresponding to 620 non-redundant protein identifications (hereafter referred to as the G2C PSD data set), with an average of 34 proteins identified per gel slice (corresponding to an average of 16.7 unique proteins per gel slice) (Supplementary Table 1). These data sets are available at http://www.genes2cognition.org/science/proteomics/synapse.html.

Integration of PSP data sets

To evaluate the coverage of our PSD data set, we compiled data from six published proteomic studies of the PSD and 119 individual papers reporting proteins localized to the PSD. A number of different sequence databases were used in these studies for searching peptide mass spectra searching and, as a result, we decided to use UniProt (http://www.expasy.uniprot.org/) as our standard protein identifier. In a few cases protein identifiers could not be mapped to UniProt, so their original identifiers were retained. We observed redundancy in data sets, with the same protein accession

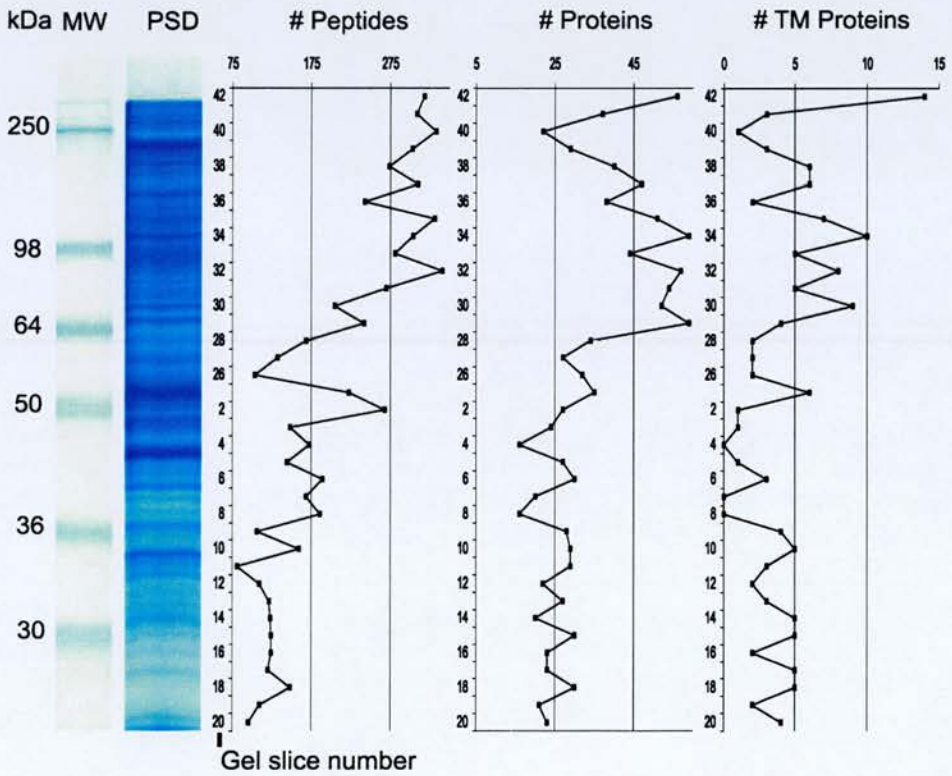


Fig. 1 Distribution of peptide and protein identifications of PSD proteins by MS analysis of gel slices. Coomassie-stained SDS–polyacrylamide gel lane containing purified mouse forebrain PSD proteins. The numbers of peptides, proteins and transmembrane proteins identified in each gel slice by LC–MS/MS analysis are shown. One-dimensional SDS–PAGE offers effective protein separation of complex mixtures with an average of 34 proteins identified per gel slice and an average of 12 peptides characterized for each protein. The number of peptides detected per gel slice increases proportionally with molecular weight. One-dimensional SDS–PAGE is very compatible with membrane samples as membrane proteins can enter the gel easily and the distribution of proteins with transmembrane (TM) regions is shown.

© 2005 International Society for Neurochemistry, *J. Neurochem.* (2005) 10.1111/j.1471-4159.2005.03507.x

Appendix 12

Table 1 Summary of PSD protein classifications

	Total PSD	G2C PSD	cPSD
Channels and receptors	80	60	39
MAGUKs/adaptors/scaffolders	54	31	33
Serine/threonine kinases	46	23	20
Tyrosine kinase	3	3	2
Protein phosphatases	18	13	11
G proteins and modulators	77	40	34
Signalling molecules and enzymes	278	167	98
Transcription and translation	119	36	41
Cytoskeletal and cell adhesion molecules	153	99	97
Synaptic vesicles and protein transport	159	119	76
Novel	107	29	11
Other	30	0	4
Summary	1124	620	466

Proteins found in the Total PSD data set, our PSD data set and the cPSD data set are classified functionally to show protein class enrichment. The distributions of our PSD data set and cPSD protein classes are similar except for signalling molecules and enzymes, which are particularly enriched in our PSD data set. In addition, the main contribution of low-abundance proteins such as tyrosine kinases and phosphatases to the Total PSD data set is from our PSD data set. In fact, our study was the only one to identify tyrosine kinases in the PSD, further supporting the increased sensitivity of our approach. Details for individual proteins and further subclassifications are found in Supplementary Table 2.

numbers present numerous times, as well as many accession numbers corresponding to the same gene product. We then mapped UniProt accession numbers to Unigene (http://www.ncbi.nlm.nih.gov/) via EMBL and Entrez Gene to cluster the merged Total PSD data set to remove redundancy. This process enabled grouping of numerous entries into Unigene clusters and resulted in reduction of the dataset to 1124 non-redundant entries (Supplementary Tables 2 and 4). An evaluation of the relative coverage of each component PSD data set is presented in Supplementary Table 3. This matrix of percentage coverage shows that the G2C data set has the highest coverage of each data set in the group, indicating that this set of proteins is the most representative and complete single data set.

We addressed the issue of bias arising from the use of different methodologies by comparing the distribution of protein identifications. The analytical approaches used in these six studies varied from use of one- and two-dimensional electrophoresis for protein separation, ICAT (Isotope-coded affinity tag) for sample simplification, to a number of peptide chromatographic methods and MS methods (Supplementary Table 3). We compared our data set (one-dimensional SDS–PAGE and LC–MS/MS); the data set of Yoshimura *et al.* (2004) (multi dimensional liquid chromatography–MS/MS) and the data set of Li *et al.* (2004) (two-dimensional gel/ICAT and matrix-assisted laser desorption ionization–time of flight, LC–MS/MS) by generating a two-dimensional scatter plot of log Mr versus pI (Supplementary Fig. 1). It can be seen that the two-dimensional LC–MS/MS approach used by Yoshimura *et al.* (2004) has the benefit of identifying many low molecular weight proteins as this approach does not utilize any gel electrophoresis step in which these proteins would be lost. Our data set is biased towards integral membrane proteins, with this class representing 24% of our data set. Both two-dimensional LC and two-dimensional gel electrophoresis are poorer approaches for identifying membrane proteins, with a representation of 17 and 7% transmembrane proteins respectively.

The distribution of protein identifications in the Total PSD data set shows that the majority of proteins were detected only once (58%) (Fig. 2a) and that 198 proteins (18%) were detected twice. The number of proteins detected five times or

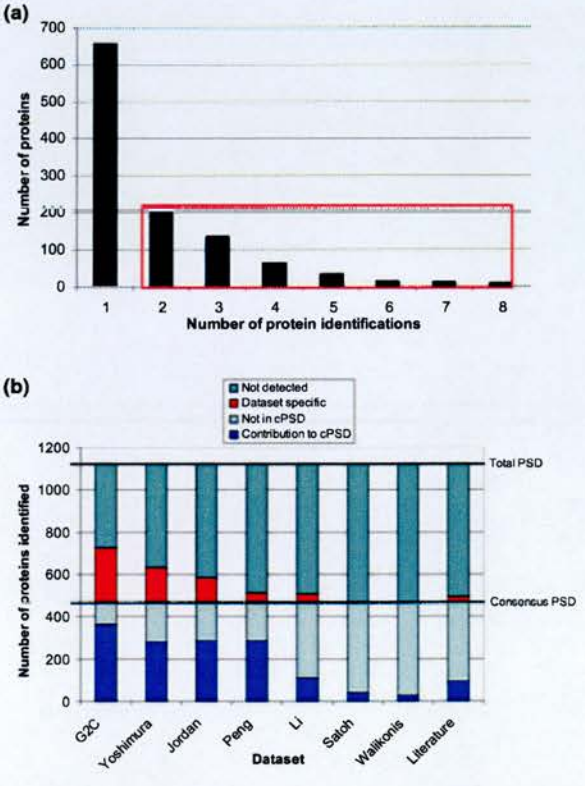


Fig. 2 Definition of a cPSD set of proteins by integration of multiple synaptic data sets. (a) Analysis of the distribution of proteins identified reveals that the majority of PSD proteins were identified only once. Proteins identified in two or more studies may represent a higher-confidence and, most likely, a high-abundance cPSD data set. (b) Contributions of each PSD data set to the cPSD (red) and Total PSD (blue) data sets. The G2C PSD analysis has the highest coverage of the cPSD at 78% and also contributes the most unique proteins to the Total PSD data set.

© 2005 International Society for Neurochemistry, *J. Neurochem.* (2005) 10.1111/j.1471-4159.2005.03507.x

Appendix 12

more was just 72 (6%), although this value is limited by the smaller data sets included in the Total PSD data set. In order to define a set of higher-confidence proteins we placed proteins that were identified two or more times (466 proteins) in a group termed the Consensus PSD (cPSD). We reasoned that the majority these proteins are abundant, so are more likely to be detected in multiple analyses, and that the set of proteins detected once are perhaps of lower abundance or may reflect differences in sample preparation. The coverage of each individual data set with respect to the consensus and Total PSD data sets is illustrated in Fig. 2(b). Some 363 proteins in our data set have been validated by their presence in one or more of the other data sets, representing 78% coverage of the cPSD. In addition, our data set contains 257 unique proteins, representing 39% of proteins detected once.

Verification of PSP proteins

We sought to verify the subcellular location of a selected set of PSD and MASC proteins by large-scale immunoblotting of whole extract, synaptosomes and PSDs from mouse forebrains. Forty-eight commercially available antibodies were tested on these three fractions to demonstrate the presence and enrichment of postsynaptic proteins (Supplementary Table 5). Overall, we found that 21 proteins (44%) were clearly enriched, eight (17%) showed similar abundance, 13 (27%) were present but not enriched and six (12%) were undetectable in isolated PSDs compared with other brain fractions (Supplementary Figs 2 and 3). Thirty-three of the 48 proteins tested are cPSD proteins and 15 of these displayed enrichment in the PSD fraction. Proteins contained in the MASC data set were identified in purifications from whole-brain extract and, although 144 of of 186 have been verified by their identification in the PSD fraction, the remainder of this data set requires validation. Fifteen antibodies against MASC proteins were used to confirm their presence in the PSD; of these eight were not shown to be PSD proteins, by MS, and seven were found in the PSD once by MS. We confirmed the PSD localization of 14 of the 15 MASC proteins tested (Supplementary Figs 2 and 3), indicating that the majority of MASC proteins are probably *bona fide* PSD proteins.

In addition to verification by immunoblotting, we performed extensive searching in PubMed for PSD localization data for these 1124 proteins (Supplementary Table 2). We found that 119 (10.6%) of these proteins have been previously reported to be localized in PSD. As shown in Supplementary Table 3, we identified more known PSD proteins (66% of PSD proteins from the literature) in our analysis of the PSD than did other similar analyses, indicating the high coverage of our data set.

Classification of PSP proteins

The initial approach to characterizing the organization of the PSP was to classify the proteins into categories that describe their known functional properties (Table 1). Within the overall PSP, representatives of many classes of proteins representing a broad range of cell biological functions were observed: membrane-bound receptors, adhesion proteins and channels; a plethora of signalling proteins and adaptors; and proteins involved with regulation of transport, RNA metabolism, translation and transcription. The cPSD contains similar numbers of proteins in most classes as our PSD data set, with the exception of channels and receptors, signalling proteins and synaptic vesicle proteins. This becomes evident when the cPSD is compared with the Total PSD; the reduced presence of these protein classes in the cPSD is most probably due to the lower abundance of these proteins, and thus a decreased likelihood of detection. This supports the idea that the cPSD represents a consensus of protein identifications which, on one hand, decreases individual data set bias for contaminating proteins (nuclear or mitochondrial) but, on the other, has limited coverage of some protein classes because of abundance issues. It is worth noting that the G2C PSD data set contributed all three tyrosine kinases found in the PSD, an indication of the sensitivity of our approach as tyrosine kinases are of relatively low abundance. The G2C PSD data set also has more coverage of most protein classes than individual data sets that constitute the Total PSD.

Protein domain analysis

Do synapse proteomes contain particular functional subsets of proteins? We examined their protein domain profiles using the InterPro resource (http://www.ebi.ac.uk/interpro). All domains in three PSD data sets [G2C PSD (594 proteins with 1709 domains), the Total PSD (1032 proteins with 2786 domains and the cPSD (447 proteins with 1298 domains)] were identified and ranked by frequency (% of proteins in each PSD data set with a specific domain) and compared with the UniProt mouse proteome (http://www.ebi.ac.uk/integr8). The enrichment of the top 10 domains found in the cPSD was compared with the Total and our PSD data set (Supplementary Fig. 4). The greatest enrichment in these data sets is for protein interaction domains and, in particular, the cPSD has a striking enrichment of PDZ and SH3 domains (7.7- and 7.2-fold respectively). This is in agreement with the notion that the cPSD contains abundant proteins such as synaptic scaffolders containing PDZ and SH3 domains. The most common domains in all three data sets are associated with kinase function and the enrichment profile for these domains is similar across these data sets. This enrichment in kinase domains correlates with enrichment in other domains commonly found in protein kinases, which allow them to interact in networks (e.g. SH3, PDZ interaction domains (Manning *et al.* 2002).

Importantly, domains involved with Ca^{2+}-dependent signalling, a major feature of postsynaptic signal transduction, were also highly abundant (C2, C2 calcium/lipid-binding).

© 2005 International Society for Neurochemistry, *J. Neurochem.* (2005) 10.1111/j.1471-4159.2005.03507.x

Appendix 12

Similarly, members of the Ras GTPase superfamily were enriched by over 5-fold in the Total PSD data set. Conversely, some domains that are highly abundant in the mouse proteome were absent from all PSD data sets, such as olfactory receptor homeobox domains (second and 19th most abundant domains respectively). The same signalling domains enriched at the synapse are known to have expanded in the genome with the evolutionary step from single to multicellular organisms (Manning *et al.* 2002). Some 62% of domains found in proteins contained in the Total PSD are present in *Saccharomyces cerevisiae*, indicating that a large proportion of the basic building blocks of mammalian synaptic proteins are conserved from before the metazoan expansion. The complement of these domains increases after this expansion with *Caenorhabditis elegans*, *Danio rerio* and *Drosophila melanogaster* having 85, 86 and 87% of Total PSD protein domains respectively. This indicates that after this expansion there was an increase in the order of 20–30% in the appearance of new domains as well as domain shuffling to produce new proteins.

Multiprotein complexes in the PSD

Previous proteomic studies showed that NMDA receptor–PSD-95 complexes were 2–3-MDa particles, which included mGluR receptors, and that AMPA receptors were in different complexes (Husi and Grant 2001; Garry *et al.* 2003). The mGluR5 receptor complex has recently been characterized in a similar fashion to the MASC and was found to contain 76 proteins, including the NMDA receptor subunit NR2A (Farr *et al.* 2004). The majority of proteins in all three glutamate receptor complexes have been independently identified as PSD proteins (Fig. 3). These complexes also contain proteins that have not yet been detected in PSD analyses, presumably owing to abundance issues in analyses of PSD preparations compared with enriched protein complexes. We have validated the presence of seven such proteins in the PSD by immunoblotting and this has been taken into account in the numbers shown in Fig. 3. It is interesting to note that the majority of the components of these complexes have been validated as PSD proteins and that overlap between these complexes occurs in the cPSD (Fig. 3), supporting the notion that the cPSD is an important subset of the Total PSD data set. This is consistent with a postsynaptic organization in which the MASC is connected to multiple cell biological effector mechanisms organized into their respective complexes. In addition to MASCs and ARCs, components of other complexes, such as cell adhesion, growth factor, cytoskeletal, transport and ribosomal complexes, were detected.

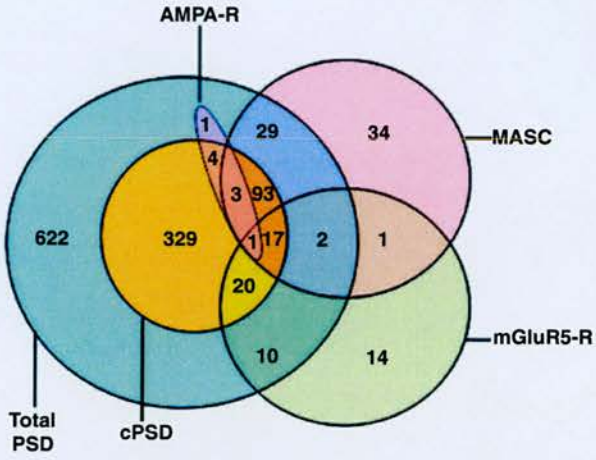


Fig. 3 Multiprotein complexes at the PSD. Venn diagram illustrating the overlap of three glutamate receptor complexes with the Total PSD and cPSD data sets. Most of the overlap between components of these complexes and the PSD occurs in the high-confidence cPSD data set. Proteins detected in these multiprotein complexes that were not found in any of the PSD data sets are generally of low abundance and enriched in immuno-purifications of complexes compared with whole-PSD analyses.

Discussion

We have presented a comprehensive analysis of isolated PSDs and embedded multiprotein complexes associated with NMDA and AMPA receptors. We have combined existing synapse proteomic data with that reported here to provide the most complete picture of the PSP to date. To our knowledge, the PSP is one of the most complicated subcellular structures in the eukaryotic cell but, unlike other subcellular structures, the PSP appears to hold a high degree of cellular autonomy. The composition of the PSP in terms of functional protein classes is diverse, with the necessary molecular machinery for not only classical synaptic processes and a myriad of signalling pathways, but for a host of supporting activities such as protein synthesis and degradation, transport and metabolism. Moreover, 10% of PSP proteins are novel and may participate in synaptic processes that are currently unknown.

It is clear that there are a large number of proteins at the postsynaptic membrane that coordinate a wide variety of cellular processes. Is there evidence of functional organization at the postsynaptic membrane? We have observed organization at many levels; from the classes of proteins present at the PSD, the enrichment of scaffolding and signalling domains in PSD proteins, to macromolecular organization in terms of multiprotein complexes. These levels of organization contribute to, or are dependent on, protein–protein interactions and so mapping and understanding protein–protein interactions in the PSP will be very useful. Protein interactions are essential to the functionality of proteomes because proteins rarely act alone, but rather in complexes. Large-scale mapping of protein interactions in the yeast proteome reveals highly complex networks of protein interactions (Fromont-Racine *et al.* 1997; Gavin *et al.* 2002). To facilitate analysis of the organization of the

© 2005 International Society for Neurochemistry,

Appendix 12

PSP at the level of protein interactions we have constructed an in-house database (http//:http://www.PPID.org). This database is a mammalian (mouse, rat, human) Protein–Protein Interaction Database (PPID) describing approximately 8000 biochemically defined protein interactions for more than 2000 proteins. So far we have performed literature mining for PSP protein interactions and have found 650 protein–protein interactions for 281 PSP proteins (of which the majority are MASC proteins). We are using these data to perform network analyses to investigate protein interaction network architecture in the PSP (Grant 2003a). In addition to literature mining, we are performing immuno-affinity purifications on PSP proteins to map additional multiprotein complexes at the synapse.

The large number of proteins found at the PSD, and the presence of many ribosomal, mitochondrial and nuclear proteins, has prompted speculation over the purity of isolated PSD fractions used in proteomic analyses. There are numerous examples of such proteins being validated as true PSD proteins. Some proteins, such as voltage-dependent anion channel 1, which was originally found in the outer mitochondrial membrane, is present in the PSD (Moon *et al.* 1999). Nuclear proteins such as AIDA-1 A, mKIAA0417 and hnRNP localize to both the nucleus and the PSD (Jordan *et al.* 2004). It is apparent that many proteins do not exist in discrete locations or indeed possess the same functions in different cellular contexts.

The data presented and compiled in this study represent a draft of an average isolated PSD. The starting material for the analyses incorporated into our draft PSP are from different species (mouse and rat) and from different brain regions (forebrain and whole brain) and encompass hundreds of neuronal cell types. In addition, it is probable that some components of the MASC are interacting with NMDA receptors during their trafficking, localization and internalization. It is now is becoming feasible to compare different types of synapse proteomes. This could be achieved by subtractive proteomics (Andersen *et al.* 2003), a quantitative approach whereby synaptosomes and isolated PSDs could be isotopically tagged (ICAT or iTRAQ (isotopic tags for relative and obsolute quantification)) and the resulting quantitative MS data would indicate enrichment or depletion of proteins in the PSD compared with synaptosomes. This type of analysis would provide confident assignment of true PSD proteins, highlight contaminating proteins and would validate the PSD localization of proteins currently considered to be restricted to other cellular compartments.

In addition, isolated PSDs or protein complexes from different brain regions or cultured cell types could be compared quantitatively using isotopic labelling. Alternatively, peptides derived from SILAC (stable isotope labelling with amino acids in cell culture)-labelled cells could be used quantitatively to compare the expression of proteins in many tissue samples (Ishihama *et al.* 2005). Using these approaches, proteins enriched in a particular cell type or brain region would be identified as being potentially more relevant to the physiology of that particular cell or tissue type. It is at this stage that systematic immuno-cytochemistry of these enriched proteins would be feasible and would provide complementary information regarding their specific protein localization at the PSD and throughout the cell type in question. Current projects, such as Gene Expression Nervous System Atlas (GENSAT), in which large-scale *in situ* hybridization and Bacteriol Artificial Chromosome (BAC) transgenic reporter genes are being used to map gene expression in neuronal cell types, will provide complementary information to the proteomic experiments suggested here.

This draft PSP data set constitutes a candidate gene list for the Genes to Cognition research programme in which the functions of a large number of these genes are being studied by systematic targeted mutations in mice (Grant 2003b). Comprehensive phenotypic analysis of these mutants in terms of biochemistry, electrophysiology, cell biology and behaviour is allowing functional annotation of PSP components. This approach, together with single-nucleotide polymorphism screening in humans for all genes in the PSP, will produce systematic data relevant to synaptic physiology and disease. Establishment of this PSP data set (available at http://www.genes2cognition.org/science/proteomics/synapse.html) is essential for the progression of synapse proteomics and, with new quantitative proteomic strategies that may be applied to activated or diseased states of the synapse, will also provide a benchmark for analysis of dynamic aspects of the synapse proteome.

Acknowledgements

This work was supported by the Genes to Cognition programme (Wellcome Trust) (to MOC, HH, LY, JMB, CNGA, JSC and SGNG) and the Biotechnology and Biological Sciences Research Council (to MOC). For detailed contributions of authors to this work see http://www.genes2cognition.org.

Supplementary Material

The following supplementary material is available for this article online:

Figure S1. Comparison of analytical approaches by physiochemical properties.

Figure S2. Immunoblotting of proteins enriched in the PSD fraction.

Figure S3. Immunoblotting of proteins present or absent in the PSD fraction

Figure S4. Enrichment of protein domains in PSD datasets

Table S1. Overview of G2C PSP datasets

Table S2. Integration of PSP datasets

Table S3. Summary of integrated PSD and multiprotein complex datasets

Table S4. Percentage overlap between PSD datasets

© 2005 International Society for Neurochemistry, *J. Neurochem.* (2005) 10.1111/j.1471-4159.2005.03507.x

Appendix 12

Table S5. Summary of immunoblotting analysis of PSP components

This material is available as part of the online article from http://www.blackwell-synergy.com.

References

Andersen J. S., Wilkinson C. J., Mayor T., Mortensen P., Nigg E. A. and Mann M. (2003) Proteomic characterization of the human centrosome by protein correlation profiling. *Nature* **426**, 570–574.

Carlin R. K., Grab D. J., Cohen R. S. and Siekevitz P. (1980) Isolation and characterization of postsynaptic densities from various brain regions: enrichment of different types of postsynaptic densities. *J. Cell Biol.* **86**, 831–845.

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.

Cotman C. W., Banker G., Churchill L. and Taylor D. (1974) Isolation of postsynaptic densities from rat brain. *J. Cell Biol.* **63**, 441–455.

Farr C. D., Gafken P. R., Norbeck A. D., Doneanu C. E., Stapels M. D., Barofsky D. F., Minami M. and Saugstad J. A. (2004) Proteomic analysis of native metabotropic glutamate receptor 5 protein complexes reveals novel molecular constituents. *J. Neurochem.* **91**, 438–450.

Fromont-Racine M., Rain J. C. and Legrain P. (1997) Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. *Nat. Genet.* **16**, 277–282.

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) Neuropathic sensitization of behavioral reflexes and spinal NMDA receptor/CaM kinase II interactions are disrupted in PSD-95 mutant mice. *Curr. Biol.* **13**, 321–328.

Gavin A. C., Bosche M., Krause R. *et al.* (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141–147.

Grant S. G. (2003a) Synapse signalling complexes and networks: machines underlying cognition. *Bioessays* **25**, 1229–1235.

Grant S. G. (2003b) Systems biology in neuroscience: bridging genes to cognition. *Curr. Opin. Neurobiol.* **13**, 577–582.

Greengard P. (2001) The neurobiology of slow synaptic transmission. *Science* **294**, 1024–1030.

Husi H. and Grant S. G. (2001) Isolation of 2000-kDa complexes of *N*-methyl-D-aspartate receptor and postsynaptic density 95 from mouse brain. *J. Neurochem.* **77**, 281–291.

Husi H., Ward M. A., Choudhary J. S., Blackstock W. P. and Grant S. G. (2000) Proteomic analysis of NMDA receptor-adhesion protein signaling complexes. *Nat. Neurosci.* **3**, 661–669.

Ishihama Y., Sato T., Tabata T., Miyamoto N., Sagane K., Nagasu T. and Oda Y. (2005) Quantitative mouse brain proteomics using culture-derived isotope tags as internal standards. *Nat. Biotechnol.* **23**, 617–621.

Jeong H., Tombor B., Albert R., Oltvai Z. N. and Barabasi A. L. (2000) The large-scale organization of metabolic networks. *Nature* **407**, 651–654.

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.

Kandel E. R. (2001) The molecular biology of memory storage: a dialogue between genes and synapses. *Science* **294**, 1030–1038.

Li K. W., Hornshaw M. P., Van Der Schors R. C. *et al.* (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.

Link A. J., Eng J., Schieltz D. M., Carmack E., Mize G. J., Morris D. R., Garvik B. M., Yates J. R. III. (1999) Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol.* **17**, 676–682.

Manning G., Whyte D. B., Martinez R., Hunter T. and Sudarsanam S. (2002) The protein kinase complement of the human genome. *Science* **298**, 1912–1934.

Moon J. I., Jung Y. W., Ko B. H., De Pinto V., Jin I. and Moon I. S. (1999) Presence of a voltage-dependent anion channel 1 in the rat postsynaptic density fraction. *Neuroreport* **10**, 443–447.

Pawson T. and Nash P. (2003) Assembly of cell regulatory systems through protein interaction domains. *Science* **300**, 445–452.

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.

Sheng M. and Kim M. J. (2002) Postsynaptic signaling and plasticity mechanisms. *Science* **298**, 776–780.

Taylor S. W., Fahy E. and Ghosh S. S. (2003) Global organellar proteomics. *Trends Biotechnol.* **21**, 82–88.

Walikonis R. S., Jensen O. N., Mann M., Provance D. W. Jr, Mercer J. A. and Kennedy M. B. (2000) Identification of proteins in the postsynaptic density fraction by mass spectrometry. *J. Neurosci.* **20**, 4069–4080.

Yamauchi T. (2002) Molecular constituents and phosphorylation-dependent regulation of the post-synaptic density. *Mass Spectrom. Rev.* **21**, 266–286.

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.

© 2005 International Society for Neurochemistry, *J. Neurochem.* (2005) 10.1111/j.1471-4159.2005.03507.x

REFERENCES

Aebersold R, Goodlett DR. 2001. Mass spectrometry in proteomics. Chem Rev 101:269-295.

Alsberg PALaCL. 1906. The cleavage products of vitellin. J. Biol. Chem 2:127–133.

Andersen JS, Lam YW, Leung AK, Ong SE, Lyon CE, Lamond AI, Mann M. 2005. Nucleolar proteome dynamics. Nature 433:77-83.

Andersen JS, Wilkinson CJ, Mayor T, Mortensen P, Nigg EA, Mann M. 2003. Proteomic characterization of the human centrosome by protein correlation profiling. Nature 426:570-574.

Andersson L, Porath J. 1986. Isolation of phosphoproteins by immobilized metal (Fe3+) affinity chromatography. Anal Biochem 154:250-254.

Ando H, Mizutani A, Matsu-ura T, Mikoshiba K. 2003. IRBIT, a novel inositol 1,4,5-trisphosphate (IP3) receptor-binding protein, is released from the IP3 receptor upon IP3 binding to the receptor. J Biol Chem 278:10602-10612.

Bagshaw RD, Mahuran DJ, Callahan JW. 2005. A proteomic analysis of lysosomal integral membrane proteins reveals the diverse composition of the organelle. Mol Cell Proteomics 4:133-143.

Ballif BA, Villen J, Beausoleil SA, Schwartz D, Gygi SP. 2004. Phosphoproteomic analysis of the developing mouse brain. Mol Cell Proteomics.

Bardoni B, Castets M, Huot ME, Schenck A, Adinolfi S, Corbin F, Pastore A, Khandjian EW, Mandel JL. 2003. 82-FIP, a novel FMRP (fragile X mental retardation protein) interacting protein, shows a cell cycle-dependent intracellular localization. Hum Mol Genet 12:1689-1698.

Barghorn S, Davies P, Mandelkow E. 2004. Tau paired helical filaments from Alzheimer's disease brain and assembled in vitro are based on beta-structure in the core domain. Biochemistry 43:1694-1703.

Barria A, Muller D, Derkach V, Griffith LC, Soderling TR. 1997. Regulatory phosphorylation of AMPA-type glutamate receptors by CaM-KII during long-term potentiation. Science 276:2042-2045.

Bayer KU, De Koninck P, Leonard AS, Hell JW, Schulman H. 2001. Interaction with the NMDA receptor locks CaMKII in an active conformation. Nature 411:801-805.

Beausoleil SA, Jedrychowski M, Schwartz D, Elias JE, Villen J, Li J, Cohn MA, Cantley LC, Gygi SP. 2004. Large-scale characterization of HeLa cell nuclear phosphoproteins. Proc Natl Acad Sci U S A 101:12130-12135.

Becamel C, Alonso G, Galeotti N, Demey E, Jouin P, Ullmer C, Dumuis A, Bockaert J, Marin P. 2002. Synaptic multiprotein complexes associated with 5-HT(2C) receptors: a proteomic approach. Embo J 21:2332-2342.

Bell AW, Ward MA, Blackstock WP, Freeman HN, Choudhary JS, Lewis AP, Chotai D, Fazel A, Gushue JN, Paiement J, Palcy S, Chevet E, Lafreniere-Roula M, Solari R, Thomas DY, Rowley A, Bergeron JJ. 2001. Proteomics characterization of abundant Golgi membrane proteins. J Biol Chem 276:5152-5165.

Bennett MK, Scheller RH. 1993. The molecular machinery for secretion is conserved from yeast to neurons. Proc Natl Acad Sci U S A 90:2559-2563.

Beranova-Giorgianni S, Giorgianni F, Desiderio DM. 2002. Analysis of the proteome in the human pituitary. Proteomics 2:534-542.

Bhardwaj N, Lu H. 2005. Correlation between gene expression profiles and protein-protein interactions within and across genomes. Bioinformatics 21:2730-2738.

Blagoev B, Ong SE, Kratchmarova I, Mann M. 2004. Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. Nat Biotechnol 22:1139-1145.

Blencowe BJ, Bauren G, Eldridge AG, Issner R, Nickerson JA, Rosonina E, Sharp PA. 2000. The SRm160/300 splicing coactivator subunits. Rna 6:111-120.

Blondeau F, Ritter B, Allaire PD, Wasiak S, Girard M, Hussain NK, Angers A, Legendre-Guillemin V, Roy L, Boismenu D, Kearney RE, Bell AW, Bergeron JJ, McPherson PS. 2004. Tandem MS analysis of brain clathrin-coated vesicles reveals their critical involvement in synaptic vesicle recycling. Proc Natl Acad Sci U S A 101:3833-3838.

Bouwmeester T, Bauch A, Ruffner H, Angrand PO, Bergamini G, Croughton K, Cruciat C, Eberhard D, Gagneur J, Ghidelli S, Hopf C, Huhse B, Mangano R, Michon AM, Schirle M, Schlegl J, Schwab M, Stein MA, Bauer A, Casari G, Drewes G, Gavin AC, Jackson DB, Joberty G, Neubauer G, Rick J, Kuster B, Superti-Furga G. 2004. A physical and functional map of the human TNF-alpha/NF-kappa B signal transduction pathway. Nat Cell Biol 6:97-105.

Browning M, Bennett WF, Kelly P, Lynch G. 1981. Evidence that the 40,000 Mr phosphoprotein influenced by high frequency synaptic stimulation is the alpha subunit of pyruvate dehydrogenase. Brain Res 218:255-266.

Browning M, Dunwiddie T, Bennett W, Gispen W, Lynch G. 1979. Synaptic phosphoproteins: specific changes after repetitive stimulation of the hippocampal slice. Science 203:60-62.

Bruner J, Tauc L. 1965. [Synaptic Plasticity Involved In The Habituation Process In Aplysia.]. J Physiol (Paris) 57:230-231.

Burnett G, Kennedy EP. 1954. The enzymatic phosphorylation of proteins. J Biol Chem 211:969-980.

Cao W, Garcia-Blanco MA. 1998. A serine/arginine-rich domain in the human U1 70k protein is necessary and sufficient for ASF/SF2 binding. J Biol Chem 273:20629-20635.

Cao W, Jamison SF, Garcia-Blanco MA. 1997. Both phosphorylation and dephosphorylation of ASF/SF2 are required for pre-mRNA splicing in vitro. Rna 3:1456-1467.

Carlin RK, Grab DJ, Cohen RS, Siekevitz P. 1980. Isolation and characterization of postsynaptic densities from various brain regions: enrichment of different types of postsynaptic densities. J Cell Biol 86:831-845.

Chang EJ, Archambault V, McLachlin DT, Krutchinsky AN, Chait BT. 2004. Analysis of protein phosphorylation by hypothesis-driven multiple-stage mass spectrometry. Anal Chem 76:4472-4483.

Chung HJ, Huang YH, Lau LF, Huganir RL. 2004. Regulation of the NMDA receptor complex and trafficking by activity-dependent phosphorylation of the NR2B subunit PDZ ligand. J Neurosci 24:10248-10259.

Chung HJ, Xia J, Scannevin RH, Zhang X, Huganir RL. 2000. Phosphorylation of the AMPA receptor subunit GluR2 differentially regulates its interaction with PDZ domain-containing proteins. J Neurosci 20:7258-7267.

Coghlan VM, Perrino BA, Howard M, Langeberg LK, Hicks JB, Gallatin WM, Scott JD. 1995. Association of protein kinase A and protein phosphatase 2B with a common anchoring protein. Science 267:108-111.

Cohen P, Burchell A, Foulkes JG, Cohen PT. 1978. Identification of the Ca2+-dependent modulator protein as the fourth subunit of rabbit skeletal muscle phosphorylase kinase. FEBS Lett 92:287-293.

Cole AR, Knebel A, Morrice NA, Robertson LA, Irving AJ, Connolly CN, Sutherland C. 2004. GSK-3 phosphorylation of the Alzheimer epitope within collapsin response mediator proteins regulates axon elongation in primary neurons. J Biol Chem 279:50176-50180.

Colledge M, Dean RA, Scott GK, Langeberg LK, Huganir RL, Scott JD. 2000. Targeting of PKA to glutamate receptors through a MAGUK-AKAP complex. Neuron 27:107-119.

Collins MO, Yu L, Coba MP, Husi H, Campuzano I, Blackstock WP, Choudhary JS, Grant SG. 2005. Proteomic analysis of in vivo phosphorylated synaptic proteins. J Biol Chem 280:5972-5982.

Collins MO, Yu L, Husi H, Blackstock WP, Choudhary JS, Grant SG. 2005. Robust enrichment of phosphorylated species in complex mixtures by sequential protein and peptide metal-affinity chromatography and analysis by tandem mass spectrometry. Sci STKE 2005:pl6.

Colwill K, Pawson T, Andrews B, Prasad J, Manley JL, Bell JC, Duncan PI. 1996. The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. Embo J 15:265-275.

Cotman CW, Banker G, Churchill L, Taylor D. 1974. Isolation of postsynaptic densities from rat brain. J Cell Biol 63:441-455.

Cotter D, Guda P, Fahy E, Subramaniam S. 2004. MitoProteome: mitochondrial protein sequence database and annotation system. Nucleic Acids Res 32:D463-467.

Craig TJ, Evans GJ, Morgan A. 2003. Physiological regulation of Munc18/nSec1 phosphorylation on serine-313. J Neurochem 86:1450-1457.

Cruz JC, Tseng HC, Goldman JA, Shih H, Tsai LH. 2003. Aberrant Cdk5 activation by p25 triggers pathological events leading to neurodegeneration and neurofibrillary tangles. Neuron 40:471-483.

Dabrowska R, Sherry JM, Aromatorio DK, Hartshorne DJ. 1978. Modulator protein as a component of the myosin light chain kinase from chicken gizzard. Biochemistry 17:253-258.

Dajani R, Fraser E, Roe SM, Young N, Good V, Dale TC, Pearl LH. 2001. Crystal structure of glycogen synthase kinase 3 beta: structural basis for phosphate-primed substrate specificity and autoinhibition. Cell 105:721-732.

DeGiorgis JA, Jaffe H, Moreira JE, Carlotti CG, Jr., Leite JP, Pant HC, Dosemeci A. 2005. Phosphoproteomic analysis of synaptosomes from human cerebral cortex. J Proteome Res 4:306-315.

Dermine JF, Duclos S, Garin J, St-Louis F, Rea S, Parton RG, Desjardins M. 2001. Flotillin-1-enriched lipid raft domains accumulate on maturing phagosomes. J Biol Chem 276:18507-18512.

Ducret A, Van Oostveen I, Eng JK, Yates JR, 3rd, Aebersold R. 1998. High throughput protein characterization by automated reverse-phase chromatography/electrospray tandem mass spectrometry. Protein Sci 7:706-719.

Dyson HJ, Wright PE. 2005. Intrinsically unstructured proteins and their functions. Nat Rev Mol Cell Biol 6:197-208.

Farr CD, Gafken PR, Norbeck AD, Doneanu CE, Stapels MD, Barofsky DF, Minami M, Saugstad JA. 2004. Proteomic analysis of native metabotropic glutamate receptor 5 protein complexes reveals novel molecular constituents. J Neurochem 91:438-450.

Ficarro SB, McCleland ML, Stukenberg PT, Burke DJ, Ross MM, Shabanowitz J, Hunt DF, White FM. 2002. Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. Nat Biotechnol 20:301-305.

Fischer EH, Krebs EG. 1955. Conversion of phosphorylase b to phosphorylase a in muscle extracts. J Biol Chem 216:121-132.

Foster LJ, De Hoog CL, Mann M. 2003. Unbiased quantitative proteomics of lipid rafts reveals high specificity for signaling factors. Proc Natl Acad Sci U S A 100:5813-5818.

Fountoulakis M, Schuller E, Hardmeier R, Berndt P, Lubec G. 1999. Rat brain proteins: two-dimensional protein database and variations in the expression level. Electrophoresis 20:3572-3579. [pii].

Friso G, Wikstrom L. 1999. Analysis of proteins from membrane-enriched cerebellar preparations by two-dimensional gel electrophoresis and mass spectrometry. Electrophoresis 20:917-927. [pii].

Fromont-Racine M, Rain JC, Legrain P. 1997. Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. Nat Genet 16:277-282.

Gagnon E, Duclos S, Rondeau C, Chevet E, Cameron PH, Steele-Mortimer O, Paiement J, Bergeron JJ, Desjardins M. 2002. Endoplasmic reticulum-mediated phagocytosis is a mechanism of entry into macrophages. Cell 110:119-131.

Gardoni F, Mauceri D, Fiorentini C, Bellone C, Missale C, Cattabeni F, Luca MD. 2003. CaMKII-dependent Phosphorylation Regulates SAP97/NR2A Interaction. J Biol Chem 278:44745-44752.

Garin J, Diez R, Kieffer S, Dermine JF, Duclos S, Gagnon E, Sadoul R, Rondeau C, Desjardins M. 2001. The phagosome proteome: insight into phagosome functions. J Cell Biol 152:165-180.

Garry EM, Moss A, Delaney A, O'Neill F, Blakemore J, Bowen J, Husi H, Mitchell R, Grant SG, Fleetwood-Walker SM. 2003. Neuropathic sensitization of behavioral reflexes and spinal NMDA receptor/CaM kinase II interactions are disrupted in PSD-95 mutant mice. Curr Biol 13:321-328.

Gauss C, Kalkum M, Lowe M, Lehrach H, Klose J. 1999. Analysis of the mouse proteome. (I) Brain proteins: separation by two- dimensional electrophoresis and identification by mass spectrometry and genetic variation. Electrophoresis 20:575-600.

Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon AM, Cruciat CM, Remor M, Hofert C, Schelder M, Brajenovic M, Ruffner H, Merino A, Klein K, Hudak M, Dickson D, Rudi T, Gnau V, Bauch A, Bastuck S, Huhse B, Leutwein C, Heurtier MA, Copley RR, Edelmann A, Querfurth E, Rybin V, Drewes G, Raida M, Bouwmeester T, Bork P, Seraphin B, Kuster B, Neubauer G, Superti-Furga G. 2002. Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature 415:141-147.

Gavin AC, Superti-Furga G. 2003. Protein complexes and proteome organization from yeast to man. Curr Opin Chem Biol 7:21-27.

Ge H, Liu Z, Church GM, Vidal M. 2001. Correlation between transcriptome and interactome mapping data from Saccharomyces cerevisiae. Nat Genet 29:482-486.

Goehler H, Lalowski M, Stelzl U, Waelter S, Stroedicke M, Worm U, Droege A, Lindenberg KS, Knoblich M, Haenig C, Herbst M, Suopanki J, Scherzinger E, Abraham C, Bauer B, Hasenbank R, Fritzsche A, Ludewig AH, Bussow K, Coleman SH, Gutekunst CA, Landwehrmeyer BG, Lehrach H, Wanker EE. 2004. A protein interaction network links GIT1, an enhancer of huntingtin aggregation, to Huntington's disease. Mol Cell 15:853-865.

Gomez LL, Alam S, Smith KE, Horne E, Dell'Acqua ML. 2002. Regulation of A-kinase anchoring protein 79/150-cAMP-dependent protein kinase postsynaptic targeting by NMDA receptor activation of calcineurin and remodeling of dendritic actin. J Neurosci 22:7027-7044.

Goode BL, Denis PE, Panda D, Radeke MJ, Miller HP, Wilson L, Feinstein SC. 1997. Functional interactions between the proline-rich and repeat regions of tau enhance microtubule binding and assembly. Mol Biol Cell 8:353-365.

Grant SG, O'Dell TJ, Karl KA, Stein PL, Soriano P, Kandel ER. 1992. Impaired long-term potentiation, spatial learning, and hippocampal development in fyn mutant mice. Science 258:1903-1910.

Grant SG. 2003. Synapse signalling complexes and networks: machines underlying cognition. Bioessays 25:1229-1235.

Grant SG. 2003. Systems biology in neuroscience: bridging genes to cognition. Curr Opin Neurobiol 13:577-582.

Graveley BR. 2000. Sorting out the complexity of SR protein functions. Rna 6:1197-1211.

Graveley BR, Maniatis T. 1998. Arginine/serine-rich domains of SR proteins can function as activators of pre-mRNA splicing. Mol Cell 1:765-771.

Gray PA, Fu H, Luo P, Zhao Q, Yu J, Ferrari A, Tenzen T, Yuk DI, Tsung EF, Cai Z, Alberta JA, Cheng LP, Liu Y, Stenman JM, Valerius MT, Billings N, Kim HA, Greenberg ME, McMahon AP, Rowitch DH, Stiles CD, Ma Q. 2004. Mouse brain organization revealed through direct genome-scale TF expression analysis. Science 306:2255-2257.

Greengard P. 2001. The neurobiology of slow synaptic transmission. Science 294:1024-1030.

Greengard P, Valtorta F, Czernik AJ, Benfenati F. 1993. Synaptic vesicle phosphoproteins and regulation of synaptic function. Science 259:780-785.

Gruhler A, Olsen JV, Mohammed S, Mortensen P, Faergeman NJ, Mann M, Jensen ON. 2005. Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. Mol Cell Proteomics 4:310-327.

Gu Y, Hamajima N, Ihara Y. 2000. Neurofibrillary tangle-associated collapsin response mediator protein-2 (CRMP-2) is highly phosphorylated on Thr-509, Ser-518, and Ser-522. Biochemistry 39:4267-4275.

Gunasekaran K, Tsai CJ, Kumar S, Zanuy D, Nussinov R. 2003. Extended disordered proteins: targeting function with less scaffold. Trends Biochem Sci 28:81-85.

Gygi SP, Aebersold R. 2000. Mass spectrometry and proteomics. Curr Opin Chem Biol 4:489-494.

Han JD, Bertin N, Hao T, Goldberg DS, Berriz GF, Zhang LV, Dupuy D, Walhout AJ, Cusick ME, Roth FP, Vidal M. 2004. Evidence for dynamically organized modularity in the yeast protein-protein interaction network. Nature 430:88-93.

Haydon CE, Eyers PA, Aveline-Wolf LD, Resing KA, Maller JL, Ahn NG. 2003. Identification of Novel Phosphorylation Sites on Xenopus laevis Aurora A and Analysis of Phosphopeptide Enrichment by Immobilized Metal-affinity Chromatography. Mol Cell Proteomics 2:1055-1067.

Henzel WJ, Billeci TM, Stults JT, Wong SC, Grimley C, Watanabe C. 1993. Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. Proc Natl Acad Sci U S A 90:5011-5015.

Hisatsune C, Umemori H, Mishina M, Yamamoto T. 1999. Phosphorylation-dependent interaction of the N-methyl-D-aspartate receptor epsilon 2 subunit with phosphatidylinositol 3-kinase. Genes Cells 4:657-666.

Ho Y, Gruhler A, Heilbut A, Bader GD, Moore L, Adams SL, Millar A, Taylor P, Bennett K, Boutilier K, Yang L, Wolting C, Donaldson I, Schandorff S, Shewnarane J, Vo M, Taggart J, Goudreault M, Muskat B, Alfarano C, Dewar D, Lin Z, Michalickova K, Willems AR, Sassi H, Nielsen PA, Rasmussen KJ, Andersen JR, Johansen LE, Hansen LH, Jespersen H, Podtelejnikov A, Nielsen E, Crawford J, Poulsen V, Sorensen BD, Matthiesen J, Hendrickson RC, Gleeson F, Pawson T, Moran MF, Durocher D, Mann M, Hogue CW, Figeys D, Tyers M. 2002. Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. Nature 415:180-183.

Holland PM, Cooper JA. 1999. Protein modification: docking sites for kinases. Curr Biol 9:R329-331.

Houde M, Bertholet S, Gagnon E, Brunet S, Goyette G, Laplante A, Princiotta MF, Thibault P, Sacks D, Desjardins M. 2003. Phagosomes are competent organelles for antigen cross-presentation. Nature 425:402-406.

Huang Y, Yario TA, Steitz JA. 2004. A molecular link between SR protein dephosphorylation and mRNA export. Proc Natl Acad Sci U S A 101:9666-9670.

Hughes CA, Bennett V. 1995. Adducin: a physical model with implications for function in assembly of spectrin-actin complexes. J Biol Chem 270:18990-18996.

Hunter T, Sefton BM. 1980. Transforming gene product of Rous sarcoma virus phosphorylates tyrosine. Proc Natl Acad Sci U S A 77:1311-1315.

Husi H, and Grant, S. G. 2002. Construction of a Protein-Protein Interaction Database (PPID) for Synaptic Biology. In Neuroscience Databases: A practical Guide (Boston/Dordrecht/London, Kluwer Academic Publishers):pp. 51-62.

Husi H, Grant SG. 2001. Isolation of 2000-kDa complexes of N-methyl-D-aspartate receptor and postsynaptic density 95 from mouse brain. J Neurochem 77:281-291.

Husi H, Ward MA, Choudhary JS, Blackstock WP, Grant SG. 2000. Proteomic analysis of NMDA receptor-adhesion protein signaling complexes. Nat Neurosci 3:661-669.

Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z, Dunker AK. 2004. The importance of intrinsic disorder for protein phosphorylation. Nucleic Acids Res 32:1037-1049.

Impey S, McCorkle SR, Cha-Molstad H, Dwyer JM, Yochum GS, Boss JM, McWeeney S, Dunn JJ, Mandel G, Goodman RH. 2004. Defining the CREB regulon: a genome-wide analysis of transcription factor regulatory regions. Cell 119:1041-1054.

Ingebritsen TS, Cohen P. 1983. Protein phosphatases: properties and role in cellular regulation. Science 221:331-338.

Ishihama Y, Sato T, Tabata T, Miyamoto N, Sagane K, Nagasu T, Oda Y. 2005. Quantitative mouse brain proteomics using culture-derived isotope tags as internal standards. Nat Biotechnol 23:617-621.

Jaffe H, Veeranna, Pant HC. 1998. Characterization of serine and threonine phosphorylation sites in beta-elimination/ethanethiol addition-modified proteins by electrospray tandem mass spectrometry and database searching. Biochemistry 37:16211-16224.

Jeong H, Mason SP, Barabasi AL, Oltvai ZN. 2001. Lethality and centrality in protein networks. Nature 411:41-42.

Jeong H, Tombor B, Albert R, Oltvai ZN, Barabasi AL. 2000. The large-scale organization of metabolic networks. Nature 407:651-654.

Jiang Z, Tang H, Havlioglu N, Zhang X, Stamm S, Yan R, Wu JY. 2003. Mutations in tau gene exon 10 associated with FTDP-17 alter the activity of an exonic splicing enhancer to interact with Tra2 beta. J Biol Chem 278:18997-19007.

Johansen JW, Ingebritsen TS. 1987. Effects of phosphorylation of protein phosphatase 1 by pp60v-src on the interaction of the enzyme with substrates and inhibitor proteins. Biochim Biophys Acta 928:63-75.

Jordan BA, Fernholz BD, Boussac M, Xu C, Grigorean G, Ziff EB, Neubert TA. 2004. Identification and verification of novel rodent postsynaptic density proteins. Mol Cell Proteomics 3:857-871.

Kandel ER. 2001. The molecular biology of memory storage: a dialogue between genes and synapses. Science 294:1030-1038.

Kaytor MD, Orr HT. 2002. The GSK3 beta signaling cascade and neurodegenerative disease. Curr Opin Neurobiol 12:275-278.

Kim CH, Chung HJ, Lee HK, Huganir RL. 2001. Interaction of the AMPA receptor subunit GluR2/3 with PDZ domains regulates hippocampal long-term depression. Proc Natl Acad Sci U S A 98:11725-11730.

Kim E, Sheng M. 2004. PDZ domain proteins of synapses. Nat Rev Neurosci 5:771-781.

Kim M, Jiang LH, Wilson HL, North RA, Surprenant A. 2001. Proteomic and functional evidence for a P2X7 receptor signalling complex. Embo J 20:6347-6358.

Kishimoto A, Nishiyama K, Nakanishi H, Uratsuji Y, Nomura H, Takeyama Y, Nishizuka Y. 1985. Studies on the phosphorylation of myelin basic protein by protein kinase C and adenosine 3':5'-monophosphate-dependent protein kinase. J Biol Chem 260:12492-12499.

Kojima T, Zama T, Wada K, Onogi H, Hagiwara M. 2001. Cloning of human PRP4 reveals interaction with Clk1. J Biol Chem 276:32247-32256.

Kondo S, Yamamoto N, Murakami T, Okumura M, Mayeda A, Imaizumi K. 2004. Tra2 beta, SF2/ASF and SRp30c modulate the function of an exonic splicing enhancer in exon 10 of tau pre-mRNA. Genes Cells 9:121-130.

Kotecha SA, Jackson MF, Al-Mahrouki A, Roder JC, Orser BA, MacDonald JF. 2003. Co-stimulation of mGluR5 and N-methyl-D-aspartate receptors is required for potentiation of excitatory synaptic transmission in hippocampal neurons. J Biol Chem 278:27742-27749.

Kyte J, Doolittle RF. 1982. A simple method for displaying the hydropathic character of a protein. J Mol Biol 157:105-132.

Lai MC, Lin RI, Huang SY, Tsai CW, Tarn WY. 2000. A human importin-beta family protein, transportin-SR2, interacts with the phosphorylated RS domain of SR proteins. J Biol Chem 275:7950-7957.

Langen H, Berndt P, Roder D, Cairns N, Lubec G, Fountoulakis M. 1999. Two-dimensional map of human brain proteins. Electrophoresis 20:907-916. [pii].

Larsen MR, Thingholm TE, Jensen ON, Roepstorff P, Jorgensen TJ. 2005. Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. Mol Cell Proteomics.

Lee G, Neve RL, Kosik KS. 1989. The microtubule binding domain of tau protein. Neuron 2:1615-1624.

Lee HK, Barbarosie M, Kameyama K, Bear MF, Huganir RL. 2000. Regulation of distinct AMPA receptor phosphorylation sites during bidirectional synaptic plasticity. Nature 405:955-959.

Lee HK, Kameyama K, Huganir RL, Bear MF. 1998. NMDA induces long-term synaptic depression and dephosphorylation of the GluR1 subunit of AMPA receptors in hippocampus. Neuron 21:1151-1162.

Levene FALaPA. 1933. Serine phosphoric acid obtained on hydrolysis of vitellinic acid. J. Biol. Chem 98:109–114.

Li KW, Hornshaw MP, Van Der Schors RC, Watson R, Tate S, Casetta B, Jimenez CR, Gouwenberg Y, Gundelfinger ED, Smalla KH, Smit AB. 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.

Lim J, Lu KP. 2005. Pinning down phosphorylated tau and tauopathies. Biochim Biophys Acta 1739:311-322.

Lin CL, Leu S, Lu MC, Ouyang P. 2004. Over-expression of SR-cyclophilin, an interaction partner of nuclear pinin, releases SR family splicing factors from nuclear speckles. Biochem Biophys Res Commun 321:638-647.

Linding R, Jensen LJ, Diella F, Bork P, Gibson TJ, Russell RB. 2003. Protein disorder prediction: implications for structural proteomics. Structure (Camb) 11:1453-1459.

Link AJ, Eng J, Schieltz DM, Carmack E, Mize GJ, Morris DR, Garvik BM, Yates JR, 3rd. 1999. Direct analysis of protein complexes using mass spectrometry. Nat Biotechnol 17:676-682.

Liu L, Lin JJ, Chen X, Liu X, Xu P. 2003. Neural expression and regulation of NSSR1 proteins. Neuroreport 14:1847-1850.

Lizcano JM, Deak M, Morrice N, Kieloch A, Hastie CJ, Dong L, Schutkowski M, Reimer U, Alessi DR. 2002. Molecular basis for the substrate specificity of

NIMA-related kinase-6 (NEK6). Evidence that NEK6 does not phosphorylate the hydrophobic motif of ribosomal S6 protein kinase and serum- and glucocorticoid-induced protein kinase in vivo. J Biol Chem 277:27839-27849.

Lominadze G, Powell DW, Luerman GC, Link AJ, Ward RA, McLeish KR. 2005. Proteomic analysis of human neutrophil granules. Mol Cell Proteomics.

Lonart G, Schoch S, Kaeser PS, Larkin CJ, Sudhof TC, Linden DJ. 2003. Phosphorylation of RIM1alpha by PKA triggers presynaptic long-term potentiation at cerebellar parallel fiber synapses. Cell 115:49-60.

Loyet KM, Stults JT, Arnott D. 2005. Mass spectrometric contributions to the practice of phosphorylation site mapping through 2003: a literature review. Mol Cell Proteomics 4:235-245.

Lubec G, Krapfenbauer K, Fountoulakis M. 2003. Proteomics in brain research: potentials and limitations. Prog Neurobiol 69:193-211.

Luo L. 2000. Rho GTPases in neuronal morphogenesis. Nat Rev Neurosci 1:173-180.

Mandelkow E. 1999. Alzheimer's disease. The tangled tale of tau. Nature 402:588-589.

Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. 2002. The protein kinase complement of the human genome. Science 298:1912-1934.

Matsuoka Y, Li X, Bennett V. 1998. Adducin is an in vivo substrate for protein kinase C: phosphorylation in the MARCKS-related domain inhibits activity in promoting spectrin-actin complexes and occurs in many cells, including dendritic spines of neurons. J Cell Biol 142:485-497.

Mauceri D, Cattabeni F, Di Luca M, Gardoni F. 2004. Calcium/calmodulin-dependent protein kinase II phosphorylation drives synapse-associated protein 97 into spines. J Biol Chem 279:23813-23821.

Mazroui R, Huot ME, Tremblay S, Filion C, Labelle Y, Khandjian EW. 2002. Trapping of messenger RNA by Fragile X Mental Retardation protein into cytoplasmic granules induces translation repression. Hum Mol Genet 11:3007-3017.

McDonald IK, Thornton JM. 1994. Satisfying hydrogen bonding potential in proteins. J Mol Biol 238:777-793.

McGee AW, Bredt DS. 1999. Identification of an intramolecular interaction between the SH3 and guanylate kinase domains of PSD-95. J Biol Chem 274:17431-17436.

McGee AW, Dakoji SR, Olsen O, Bredt DS, Lim WA, Prehoda KE. 2001. Structure of the SH3-guanylate kinase module from PSD-95 suggests a mechanism for regulated assembly of MAGUK scaffolding proteins. Mol Cell 8:1291-1301.

Moon JI, Jung YW, Ko BH, De Pinto V, Jin I, Moon IS. 1999. Presence of a voltage-dependent anion channel 1 in the rat postsynaptic density fraction. Neuroreport 10:443-447.

Mootha VK, Bunkenborg J, Olsen JV, Hjerrild M, Wisniewski JR, Stahl E, Bolouri MS, Ray HN, Sihag S, Kamal M, Patterson N, Lander ES, Mann M. 2003. Integrated analysis of protein composition, tissue diversity, and gene regulation in mouse mitochondria. Cell 115:629-640.

Muslin AJ, Tanner JW, Allen PM, Shaw AS. 1996. Interaction of 14-3-3 with signaling proteins is mediated by the recognition of phosphoserine. Cell 84:889-897.

Nada S, Shima T, Yanai H, Husi H, Grant SG, Okada M, Akiyama T. 2003. Identification of PSD-93 as a substrate for the Src family tyrosine kinase Fyn. J Biol Chem 278:47610-47621.

Nakazawa T, Komai S, Tezuka T, Hisatsune C, Umemori H, Semba K, Mishina M, Manabe T, Yamamoto T. 2001. Characterization of Fyn-mediated tyrosine phosphorylation sites on GluR epsilon 2 (NR2B) subunit of the N-methyl-D-aspartate receptor. J Biol Chem 276:693-699.

Neubauer G, Gottschalk A, Fabrizio P, Seraphin B, Luhrmann R, Mann M. 1997. Identification of the proteins of the yeast U1 small nuclear ribonucleoprotein complex by mass spectrometry. Proc Natl Acad Sci U S A 94:385-390.

Neubauer G, King A, Rappsilber J, Calvio C, Watson M, Ajuh P, Sleeman J, Lamond A, Mann M. 1998. Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex. Nat Genet 20:46-50.

Niethammer M, Kim E, Sheng M. 1996. Interaction between the C terminus of NMDA receptor subunits and multiple members of the PSD-95 family of membrane-associated guanylate kinases. J Neurosci 16:2157-2163.

Nishimura T, Fukata Y, Kato K, Yamaguchi T, Matsuura Y, Kamiguchi H, Kaibuchi K. 2003. CRMP-2 regulates polarized Numb-mediated endocytosis for axon growth. Nat Cell Biol 5:819-826.

Nolen B, Yun CY, Wong CF, McCammon JA, Fu XD, Ghosh G. 2001. The structure of Sky1p reveals a novel mechanism for constitutive activity. Nat Struct Biol 8:176-183.

Nourry C, Grant SG, Borg JP. 2003. PDZ domain proteins: plug and play! Sci STKE 2003:RE7.

Nuhse TS, Stensballe A, Jensen ON, Peck SC. 2003. Large-scale analysis of in vivo phosphorylated membrane proteins by immobilized metal ion affinity chromatography and mass spectrometry. Mol Cell Proteomics.

Nuhse TS, Stensballe A, Jensen ON, Peck SC. 2004. Phosphoproteomics of the Arabidopsis Plasma Membrane and a New Phosphorylation Site Database. Plant Cell.

Obenauer JC, Cantley LC, Yaffe MB. 2003. Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs. Nucleic Acids Res 31:3635-3641.

Oda K, Matsuoka Y, Funahashi A, Kitano H. 2005. A comprehensive pathway map of epidermal growth factor receptor signaling. Mol Sys Biol.

Oda Y, Nagasu T, Chait BT. 2001. Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. Nat Biotechnol 19:379-382.

Oldfield CJ, Cheng Y, Cortese MS, Brown CJ, Uversky VN, Dunker AK. 2005. Comparing and combining predictors of mostly disordered proteins. Biochemistry 44:1989-2000.

Omkumar RV, Kiely MJ, Rosenstein AJ, Min KT, Kennedy MB. 1996. Identification of a phosphorylation site for calcium/calmodulindependent protein kinase II in the NR2B subunit of the N-methyl-D-aspartate receptor. J Biol Chem 271:31670-31678.

Orosz F, Kovacs GG, Lehotzky A, Olah J, Vincze O, Ovadi J. 2004. TPPP/p25: from unfolded protein to misfolding disease: prediction and experiments. Biol Cell 96:701-711.

Otzen DE, Lundvig DM, Wimmer R, Nielsen LH, Pedersen JR, Jensen PH. 2005. p25alpha is flexible but natively folded and binds tubulin with oligomeric stoichiometry. Protein Sci 14:1396-1409.

Palade GE. 1966. Structure and function at the cellular level. JAMA 198:815–825.

Pandey A, Podtelejnikov AV, Blagoev B, Bustelo XR, Mann M, Lodish HF. 2000. Analysis of receptor signaling pathways by mass spectrometry: identification of vav-2 as a substrate of the epidermal and platelet- derived growth factor receptors. Proc Natl Acad Sci U S A 97:179-184.

Pawson T, Nash P. 2003. Assembly of cell regulatory systems through protein interaction domains. Science 300:445-452.

Peng J, Kim MJ, Cheng D, Duong DM, Gygi SP, Sheng M. 2004. Semiquantitative proteomic analysis of rat forebrain postsynaptic density fractions by mass spectrometry. J Biol Chem 279:21003-21011.

Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20:3551-3567.

Pisitkun T, Shen RF, Knepper MA. 2004. Identification and proteomic profiling of exosomes in human urine. Proc Natl Acad Sci U S A 101:13368-13373.

Posewitz MC, Tempst P. 1999. Immobilized gallium(III) affinity chromatography of phosphopeptides. Anal Chem 71:2883-2892.

Puntervoll P, Linding R, Gemund C, Chabanis-Davidson S, Mattingsdal M, Cameron S, Martin DM, Ausiello G, Brannetti B, Costantini A, Ferre F, Maselli V, Via A, Cesareni G, Diella F, Superti-Furga G, Wyrwicz L, Ramu C, McGuigan C, Gudavalli R, Letunic I, Bork P, Rychlewski L, Kuster B, Helmer-Citterich M, Hunter WN, Aasland R, Gibson TJ. 2003. ELM server: A new resource for investigating short functional sites in modular eukaryotic proteins. Nucleic Acids Res 31:3625-3630.

Rall TW, Sutherland EW. 1958. Formation of a cyclic adenine ribonucleotide by tissue particles. J Biol Chem 232:1065-1076.

Ralser M, Albrecht M, Nonhoff U, Lengauer T, Lehrach H, Krobitsch S. 2005. An integrative approach to gain insights into the cellular function of human ataxin-2. J Mol Biol 346:203-214.

Ralser M, Nonhoff U, Albrecht M, Lengauer T, Wanker EE, Lehrach H, Krobitsch S. 2005. Ataxin-2 and huntingtin interact with endophilin-A complexes to function in plastin-associated pathways. Hum Mol Genet.

Ramón y Cajal S. 1894. La fine structure des centres nerveux. Proc. R. Soc. Lond. 55:444-468.

Rappsilber J, Ryder U, Lamond AI, Mann M. 2002. Large-scale proteomic analysis of the human spliceosome. Genome Res 12:1231-1245.

Ray LB, Sturgill TW. 1987. Rapid stimulation by insulin of a serine/threonine kinase in 3T3-L1 adipocytes that phosphorylates microtubule-associated protein 2 in vitro. Proc Natl Acad Sci U S A 84:1502-1506.

Ridsdale RA, Beniac DR, Tompkins TA, Moscarello MA, Harauz G. 1997. Three-dimensional structure of myelin basic protein. II. Molecular modeling and considerations of predicted structures in multiple sclerosis. J Biol Chem 272:4269-4275.

Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M, Seraphin B. 1999. A generic protein purification method for protein complex characterization and proteome exploration. Nat Biotechnol 17:1030-1032.

Roche KW, O'Brien RJ, Mammen AL, Bernhardt J, Huganir RL. 1996. Characterization of multiple phosphorylation sites on the AMPA receptor GluR1 subunit. Neuron 16:1179-1188.

Rout MP, Aitchison JD, Suprapto A, Hjertaas K, Zhao Y, Chait BT. 2000. The yeast nuclear pore complex: composition, architecture, and transport mechanism. J Cell Biol 148:635-651.

Rychlewski L, Kschischo M, Dong L, Schutkowski M, Reimer U. 2004. Target specificity analysis of the Abl kinase using peptide microarray data. J Mol Biol 336:307-311.

Salter MW, Kalia LV. 2004. Src kinases: a hub for NMDA receptor regulation. Nat Rev Neurosci 5:317-328.

Santoni V, Molloy M, Rabilloud T. 2000. Membrane proteins and proteomics: un amour impossible? Electrophoresis 21:1054-1070. [pii].

Satoh K, Takeuchi M, Oda Y, Deguchi-Tawarada M, Sakamoto Y, Matsubara K, Nagasu T, Takai Y. 2002. Identification of activity-regulated proteins in the postsynaptic density fraction. Genes Cells 7:187-197.

Schulman H, Greengard P. 1978. Stimulation of brain membrane protein phosphorylation by calcium and an endogenous heat-stable protein. Nature 271:478-479.

Schwartzberg PL, Goff SP, Robertson EJ. 1989. Germ-line transmission of a c-abl mutation produced by targeted gene disruption in ES cells. Science 246:799-803.

Scott JD, Stofko RE, McDonald JR, Comer JD, Vitalis EA, Mangili JA. 1990. Type II regulatory subunit dimerization determines the subcellular localization of the cAMP-dependent protein kinase. J Biol Chem 265:21561-21566.

Seabold GK, Burette A, Lim IA, Weinberg RJ, Hell JW. 2003. Interaction of the tyrosine kinase Pyk2 with the N-methyl-D-aspartate receptor complex via the Src homology 3 domains of PSD-95 and SAP102. J Biol Chem 278:15040-15048.

Shah K, Liu Y, Deirmengian C, Shokat KM. 1997. Engineering unnatural nucleotide specificity for Rous sarcoma virus tyrosine kinase to uniquely label its direct substrates. Proc Natl Acad Sci U S A 94:3565-3570.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498-2504.

Shapira M, Zhai RG, Dresbach T, Bresler T, Torres VI, Gundelfinger ED, Ziv NE, Garner CC. 2003. Unitary assembly of presynaptic active zones from Piccolo-Bassoon transport vesicles. Neuron 38:237-252.

Shen T, Zong C, Hamelberg D, McCammon JA, Wolynes PG. 2005. The folding energy landscape and phosphorylation: modeling the conformational switch of the NFAT regulatory domain. Faseb J 19:1389-1395.

Sheng M. 2001. Molecular organization of the postsynaptic specialization. Proc Natl Acad Sci U S A 98:7058-7061.

Sheng M, Kim MJ. 2002. Postsynaptic signaling and plasticity mechanisms. Science 298:776-780.

Shevchenko A, Chernushevich I, Ens W, Standing KG, Thomson B, Wilm M, Mann M. 1997. Rapid 'de novo' peptide sequencing by a combination of nanoelectrospray, isotopic labeling and a quadrupole/time-of-flight mass spectrometer. Rapid Commun Mass Spectrom 11:1015-1024.

Shevchenko A, Chernushevich I, Wilm M, Mann M. 2000. De Novo peptide sequencing by nanoelectrospray tandem mass spectrometry using triple quadrupole and quadrupole/time-of-flight instruments. Methods Mol Biol 146:1-16.

Shevchenko A, Loboda A, Ens W, Standing KG. 2000. MALDI quadrupole time-of-flight mass spectrometry: a powerful tool for proteomic research. Anal Chem 72:2132-2141.

Shevchenko A, Wilm M, Vorm O, Jensen ON, Podtelejnikov AV, Neubauer G, Mortensen P, Mann M. 1996. A strategy for identifying gel-separated proteins in sequence databases by MS alone. Biochem Soc Trans 24:893-896.

Shin C, Feng Y, Manley JL. 2004. Dephosphorylated SRp38 acts as a splicing repressor in response to heat shock. Nature 427:553-558.

Shin C, Manley JL. 2002. The SR protein SRp38 represses splicing in M phase cells. Cell 111:407-417.

Shoji S, Parmelee DC, Wade RD, Kumar S, Ericsson LH, Walsh KA, Neurath H, Long GL, Demaille JG, Fischer EH, Titani K. 1981. Complete amino acid sequence of the catalytic subunit of bovine cardiac muscle cyclic AMP-dependent protein kinase. Proc Natl Acad Sci U S A 78:848-851.

Shu H, Chen S, Bi Q, Mumby M, Brekken DL. 2004. Identification of phosphoproteins and their phosphorylation sites in the WEHI-231 B lymphoma cell line. Mol Cell Proteomics 3:279-286.

Siomi MC, Higashijima K, Ishizuka A, Siomi H. 2002. Casein kinase II phosphorylates the fragile X mental retardation protein and modulates its biological properties. Mol Cell Biol 22:8438-8447.

Steen H, Kuster B, Mann M. 2001. Quadrupole time-of-flight versus triple-quadrupole mass spectrometry for the determination of phosphopeptides by precursor ion scanning. J Mass Spectrom 36:782-790.

Steen H, Pandey A, Andersen JS, Mann M. 2002. Analysis of tyrosine phosphorylation sites in signaling molecules by a phosphotyrosine-specific immonium ion scanning method. Sci STKE 2002:PL16.

Stewart AA, Ingebritsen TS, Manalan A, Klee CB, Cohen P. 1982. Discovery of a Ca2+- and calmodulin-dependent protein phosphatase: probable identity with calcineurin (CaM-BP80). FEBS Lett 137:80-84.

Stoilov P, Daoud R, Nayler O, Stamm S. 2004. Human tra2-beta1 autoregulates its protein concentration by influencing alternative splicing of its pre-mRNA. Hum Mol Genet 13:509-524.

Sudhof TC. 2004. The synaptic vesicle cycle. Annu Rev Neurosci 27:509-547.

Sweatt JD, Kandel ER. 1989. Persistent and transcriptionally-dependent increase in protein phosphorylation in long-term facilitation of Aplysia sensory neurons. Nature 339:51-54.

Szafron D, Lu P, Greiner R, Wishart DS, Poulin B, Eisner R, Lu Z, Anvik J, Macdonell C, Fyshe A, Meeuwis D. 2004. Proteome Analyst: custom predictions with explanations in a web-based tool for high-throughput proteome annotations. Nucleic Acids Res 32:W365-371.

Takahashi M, Tomizawa K, Ishiguro K, Sato K, Omori A, Sato S, Shiratsuchi A, Uchida T, Imahori K. 1991. A novel brain-specific 25 kDa protein (p25) is phosphorylated by a Ser/Thr-Pro kinase (TPK II) from tau protein kinase fractions. FEBS Lett 289:37-43.

Taoka M, Wakamiya A, Nakayama H, Isobe T. 2000. Protein profiling of rat cerebella during development. Electrophoresis 21:1872-1879. [pii].

Tavares GA, Panepucci EH, Brunger AT. 2001. Structural characterization of the intramolecular interaction between the SH3 and guanylate kinase domains of PSD-95. Mol Cell 8:1313-1325.

Taylor SS, Buechler JA, Yonemoto W. 1990. cAMP-dependent protein kinase: framework for a diverse family of regulatory enzymes. Annu Rev Biochem 59:971-1005.

Taylor SW, Fahy E, Ghosh SS. 2003. Global organellar proteomics. Trends Biotechnol 21:82-88.

Tezuka T, Umemori H, Akiyama T, Nakanishi S, Yamamoto T. 1999. PSD-95 promotes Fyn-mediated tyrosine phosphorylation of the N-methyl-D-aspartate receptor subunit NR2A. Proc Natl Acad Sci U S A 96:435-440.

Tilleman K, Stevens I, Spittaels K, Haute CV, Clerens S, Van Den Bergh G, Geerts H, Van Leuven F, Vandesande F, Moens L. 2002. Differential expression of brain proteins in glycogen synthase kinase-3 transgenic mice: a proteomics point of view. Proteomics 2:94-104.

Tilleman K, Van den Haute C, Geerts H, van Leuven F, Esmans EL, Moens L. 2002. Proteomics analysis of the neurodegeneration in the brain of tau transgenic mice. Proteomics 2:656-665.

Tingley WG, Ehlers MD, Kameyama K, Doherty C, Ptak JB, Riley CT, Huganir RL. 1997. Characterization of protein kinase A and protein kinase C phosphorylation of the N-methyl-D-aspartate receptor NR1 subunit using phosphorylation site-specific antibodies. J Biol Chem 272:5157-5166.

Tirian L, Hlavanda E, Olah J, Horvath I, Orosz F, Szabo B, Kovacs J, Szabad J, Ovadi J. 2003. TPPP/p25 promotes tubulin assemblies and blocks mitotic spindle formation. Proc Natl Acad Sci U S A 100:13976-13981.

Tokuda M, Hatase O. 1998. Regulation of neuronal plasticity in the central nervous system by phosphorylation and dephosphorylation. Mol Neurobiol 17:137-156.

Tompa P. 2003. Intrinsically unstructured proteins evolve by repeat expansion. Bioessays 25:847-855.

Tonks NK, Diltz CD, Fischer EH. 1988. Purification of the major protein-tyrosine-phosphatases of human placenta. J Biol Chem 263:6722-6730.

Topinka JR, Bredt DS. 1998. N-terminal palmitoylation of PSD-95 regulates association with cell membranes and interaction with K+ channel Kv1.4. Neuron 20:125-134.

Trinidad JC, Thalhammer A, Specht CG, Schoepfer R, Burlingame AL. 2005. Phosphorylation state of postsynaptic density proteins. J Neurochem 92:1306-1316.

Tu JC, Xiao B, Yuan JP, Lanahan AA, Leoffert K, Li M, Linden DJ, Worley PF. 1998. Homer binds a novel proline-rich motif and links group 1 metabotropic glutamate receptors with IP3 receptors. Neuron 21:717-726.

Turner KM, Burgoyne RD, Morgan A. 1999. Protein phosphorylation and the regulation of synaptic membrane traffic. Trends Neurosci 22:459-464.

Uchida Y, Ohshima T, Sasaki Y, Suzuki H, Yanai S, Yamashita N, Nakamura F, Takei K, Ihara Y, Mikoshiba K, Kolattukudy P, Honnorat J, Goshima Y. 2005. Semaphorin3A signalling is mediated via sequential Cdk5 and GSK3beta phosphorylation of CRMP2: implication of common phosphorylating mechanism underlying axon guidance and Alzheimer's disease. Genes Cells 10:165-179.

Ueda T, Greengard P. 1977. Adenosine 3':5'-monophosphate-regulated phosphoprotein system of neuronal membranes. I. Solubilization, purification, and some properties of an endogenous phosphoprotein. J Biol Chem 252:5155-5163.

Ueda T, Maeno H, Greengard P. 1973. Regulation of endogenous phosphorylation of specific proteins in synaptic membrane fractions from rat brain by adenosine 3':5'-monophosphate. J Biol Chem 248:8295-8305.

Valcarcel J, Gaur RK, Singh R, Green MR. 1996. Interaction of U2AF65 RS region with pre-mRNA branch point and promotion of base pairing with U2 snRNA [corrected]. Science 273:1706-1709.

Vale RD. 2003. The molecular motor toolbox for intracellular transport. Cell 112:467-480.

Vieth M, Sutherland JJ, Robertson DH, Campbell RM. 2005. Kinomics: characterizing the therapeutically validated kinase space. Drug Discov Today 10:839-846.

Volknandt W. 1995. The synaptic vesicle and its targets. Neuroscience 64:277-300.

Wagner S, Chiosea S, Ivshina M, Nickerson JA. 2004. In vitro FRAP reveals the ATP-dependent nuclear mobilization of the exon junction complex protein SRm160. J Cell Biol 164:843-850.

Walikonis RS, Jensen ON, Mann M, Provance DW, Jr., Mercer JA, Kennedy MB. 2000. Identification of proteins in the postsynaptic density fraction by mass spectrometry. J Neurosci 20:4069-4080.

Walsh DA, Perkins JP, Krebs EG. 1968. An adenosine 3',5'-monophosphate-dependant protein kinase from rabbit skeletal muscle. J Biol Chem 243:3763-3765.

Washburn MP, Wolters D, Yates JR, 3rd. 2001. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol 19:242-247.

Weiler IJ, Irwin SA, Klintsova AY, Spencer CM, Brazelton AD, Miyashiro K, Comery TA, Patel B, Eberwine J, Greenough WT. 1997. Fragile X mental retardation protein is translated near synapses in response to neurotransmitter activation. Proc Natl Acad Sci U S A 94:5395-5400.

Wenschuh H, Volkmer-Engert R, Schmidt M, Schulz M, Schneider-Mergener J, Reineke U. 2000. Coherent membrane supports for parallel microsynthesis and screening of bioactive peptides. Biopolymers 55:188-206.

Westphal RS, Tavalin SJ, Lin JW, Alto NM, Fraser ID, Langeberg LK, Sheng M, Scott JD. 1999. Regulation of NMDA receptors by an associated phosphatase-kinase signaling complex. Science 285:93-96.

Whitman M, Downes CP, Keeler M, Keller T, Cantley L. 1988. Type I phosphatidylinositol kinase makes a novel inositol phospholipid, phosphatidylinositol-3-phosphate. Nature 332:644-646.

Wilks AF, Harpur AG, Kurban RR, Ralph SJ, Zurcher G, Ziemiecki A. 1991. Two novel protein-tyrosine kinases, each with a second phosphotransferase-related

catalytic domain, define a new class of protein kinase. Mol Cell Biol 11:2057-2065.

Wilm M, Mann M. 1996. Analytical properties of the nanoelectrospray ion source. Anal Chem 68:1-8.

Witzmann FA, Arnold RJ, Bai F, Hrncirova P, Kimpel MW, Mechref YS, McBride WJ, Novotny MV, Pedrick NM, Ringham HN, Simon JR. 2005. A proteomic survey of rat cerebral cortical synaptosomes. Proteomics 5:2177-2201.

Wright PE, Dyson HJ. 1999. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. J Mol Biol 293:321-331.

Wu CC, MacCoss MJ, Howell KE, Yates JR. 2003. A method for the comprehensive proteomic analysis of membrane proteins. Nat Biotechnol 21:532-538.

Wu CC, Taylor RS, Lane DR, Ladinsky MS, Weisz JA, Howell KE. 2000. GMx33: a novel family of trans-Golgi proteins identified by proteomics. Traffic 1:963-975.

Wu CC, Yates JR, 3rd, Neville MC, Howell KE. 2000. Proteomic analysis of two functional states of the Golgi complex in mammary epithelial cells. Traffic 1:769-782.

Xiao SH, Manley JL. 1997. Phosphorylation of the ASF/SF2 RS domain affects both protein-protein and protein-RNA interactions and is necessary for splicing. Genes Dev 11:334-344.

Yamauchi T. 2002. Molecular constituents and phosphorylation-dependent regulation of the post-synaptic density. Mass Spectrom Rev 21:266-286.

Yang M, Leonard JP. 2001. Identification of mouse NMDA receptor subunit NR2A C-terminal tyrosine sites phosphorylated by coexpression with v-Src. J Neurochem 77:580-588.

Yao WD, Gainetdinov RR, Arbuckle MI, Sotnikova TD, Cyr M, Beaulieu JM, Torres GE, Grant SG, Caron MG. 2004. Identification of PSD-95 as a regulator of dopamine-mediated synaptic and behavioral plasticity. Neuron 41:625-638.

Yin G, Zheng Q, Yan C, Berk BC. 2005. GIT1 is a scaffold for ERK1/2 activation in focal adhesions. J Biol Chem 280:27705-27712.

Yoshimura Y, Sogawa Y, Yamauchi T. 1999. Protein phosphatase 1 is involved in the dissociation of Ca2+/calmodulin-dependent protein kinase II from postsynaptic densities. FEBS Lett 446:239-242.

Yoshimura Y, Yamauchi Y, Shinkawa T, Taoka M, Donai H, Takahashi N, Isobe T, 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.

Zahler AM, Lane WS, Stolk JA, Roth MB. 1992. SR proteins: a conserved family of pre-mRNA splicing factors. Genes Dev 6:837-847.

Zhang C, Williams EH, Guo Y, Lum L, Beachy PA. 2004. Extensive phosphorylation of Smoothened in Hedgehog pathway activation. Proc Natl Acad Sci U S A 101:17900-17907.

Zhang H, Webb DJ, Asmussen H, Niu S, Horwitz AF. 2005. A GIT1/PIX/Rac/PAK signaling module regulates spine morphogenesis and synapse formation through MLC. J Neurosci 25:3379-3388.

Zhang H, Zha X, Tan Y, Hornbeck PV, Mastrangelo AJ, Alessi DR, Polakiewicz RD, Comb MJ. 2002. Phosphoprotein analysis using antibodies broadly reactive against phosphorylated motifs. J Biol Chem 277:39379-39387.

Zheng F, Gingrich MB, Traynelis SF, Conn PJ. 1998. Tyrosine kinase potentiates NMDA receptor currents by reducing tonic zinc inhibition. Nat Neurosci 1:185-191.

Zhou H, Watts JD, Aebersold R. 2001. A systematic approach to the analysis of protein phosphorylation. Nat Biotechnol 19:375-378.

Zimowska G, Shi J, Munguba G, Jackson MR, Alpatov R, Simmons MN, Shi Y, Sugrue SP. 2003. Pinin/DRS/memA interacts with SRp75, SRm300 and SRrp130 in corneal epithelial cells. Invest Ophthalmol Vis Sci 44:4715-4723.