MCP Papers in Press. Published on October 15, 2007 as Manuscript M700241-MCP200

Motif decomposition of the phosphotyrosine proteome reveals a new N-terminal binding motif for SHIP2

Martin Lee Miller[1]

Stefan Hanke[2]

Anders Mørkeberg Hinsby[1]

Carsten Friis[1]

Søren Brunak[1]

Matthias Mann[2]

Nikolaj Blom[1,3]

1) Center for Biological Sequence Analysis. Technical University of Denmark. Kemitorvet, Bldg. 208, DK-2800 Lyngby, Denmark. Fax:+45 4593 1585, Tel:+45 4525 2477

2) Department of Proteomics and Signal Transduction. Max-Planck-Institute for Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany. Fax: +49 (89) 8578 3209, Tel: +49 (89) 8578 2557

3) Correspondence: 3, nikob@cbs.dtu.dk, current email: nblm@novozymes.com

Downloaded from www.mcponline.org by on January 9, 2009

Copyright 2007 by The American Society for Biochemistry and Molecular Biology, Inc.

Keywords: phosphoproteomics, sequence motif decomposition, phosphotyrosine interaction motifs, peptide pull-down, SILAC, SHIP2, signal transduction, mass spectrometry

Running title: Motif decomposition of the pY proteome

Abbreviations: Phosphotyrosine (pY), Partitioning Around Medoids (PAM), Human Protein Reference Database (HPRD), Immunoreceptor tyrosine-based inhibition motif (ITIM), Phosphotyrosine binding (PTB), Src-homology 2 (SH2), Hidden Markov Models (HMMs), Gene Ontology (GO), Fetal bovine serum (FBS), International Protein Index (IPI) , Tandem mass spectrometry (MS/MS), Protein tyrosine kinase (PTK), Cys_2His_2 zinc finger protein (zf-C2H2), Neural Wiskott-Aldrich syndrome protein (N-WASP)

Downloaded from www.mcponline.org by on January 9, 2009

ASBMB Molecular & Cellular Proteomics

Summary

Advances in mass spectrometry-based proteomics have yielded a substantial mapping of the tyrosine phosphoproteome and thus provided an important step towards a systematic analysis of intracellular signaling networks in higher eukaryotes. In this study we decompose an uncharacterized proteomic data set of 481 unique phosphotyrosine (pY) peptides by sequence similarity to known ligands of the Src Homology 2 (SH2) and the phosphotyrosine binding (PTB) domains. From 20 clusters we extract 16 known and four new interaction motifs. Using quantitative mass spectrometry we pull down pY-specific binding partners for peptides corresponding to the extracted motifs. We confirm numerous of previously known interaction motifs and find 15 new interactions mediated by phosphosites not previously known to bind SH2 or PTB domains. Remarkably, a novel hydrophobic N-terminal motif ([LVI]-[LVI]-pY) is identified and validated as a binding motif for the SH2 domain containing inositol phosphatase SHIP2. Our decomposition of the *in vivo* pY proteome furthermore suggests that two thirds of the pY sites mediate interaction while the remaining third governs processes such as enzyme activation and nucleic acid binding.

Downloaded from www.mcponline.org by on January 9, 2009

Introduction

Phosphorylation-dependent protein-protein interaction is one of the key organizing principles in intracellular signaling events. The phosphotyrosine binding domain (PTB) and the Src homology domain 2 (SH2) are modular domains that typically bind phosphotyrosine (pY) containing peptides (Pawson, 2004; Yaffe, 2002). "Linear motifs" (unstructured sequence recognition patches with conserved residues at specific positions (Bork and Koonin, 1996)) that direct pY-dependent interaction have traditionally been studied using degenerate oriented peptide libraries. Such studies revealed that PTB and SH2 domains have preference for specific amino acids N- and C-terminal to the pY residue, respectively (Songyang et al., 1995; Songyang et al., 1994).

Recent methodological developments in mass spectrometry (MS)-based proteomics have made it possible to identify hundreds to thousands of protein phosphorylation sites in a single project (Ballif et al., 2004; Beausoleil et al., 2004; Brill et al., 2004; Ficarro et al., 2003; Ficarro et al., 2005; Olsen et al., 2006; Rush et al., 2005; Salomon et al., 2003; Zheng et al., 2005). Extensive mapping of the phosphoproteome is an important step towards analyzing the regulatory components of the cell. Since the majority of newly identified phosphopeptides are uncharacterized with respect to signaling context, there is now a unique opportunity to mine the phosphoproteome for novel phosphorylation motifs. Methods have been developed that successfully mine for over-represented motifs from large protein data sets in general (Jonassen et al., 1995; Nevill-Manning et al., 1998; Rigoutsos and Floratos, 1998) and more recently also from phosphoproteomic data sets (Schwartz and Gygi, 2005). However, these methods do not

Molecular & Cellular Proteomics ASBMB

Downloaded from www.mcponline.org by on January 9, 2009

partition the data set into smaller subsets with high sequence similarity prior to motif extraction. Sequence patches flanking the motif also govern phosphorylation-dependent recognition (Forman-Kay and Pawson, 1999); consequently, there is a risk of extracting false positive motifs from functionally unrelated peptides. Furthermore, the above mentioned methods are *in silico* approaches and do not combine prediction with experimental validation.

To overcome such limitations in the area of pY motif discovery and classification, one may partition the data set into smaller subsets e.g. by sequence similarity with known kinase or binding substrates prior to motif extraction. Thus, the risk of retrieving false positive motifs is minimized, since overrepresented motifs are extracted from peptides closely related in sequence and function. Besides pY recognition motifs for kinases and interaction domains, there may also potentially exist pY motifs that mediate other processes than binding such as e.g. enzyme activation and nucleic acid binding. Thus, it is essential to validate the extracted motifs both by experimental and bioinformatical means in order to obtain a functional classification.

With this in mind we have developed a motif extraction and validation methodology and classified pY motifs on a proteome level. Operating in sequence space, we stretched the MS-mapped pY peptides over a backbone of ligands already known to be involved in pY-dependent interaction. Using experimentally verified pY ligands of the PTB and SH2 domains as both a clustering backbone and as a control for the partitioning, we split a literature-extracted data set of mammalian pY peptides into 20 different clusters. We obtained a meaningful clustering since the controls partitioned correctly into separate clusters.

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

From the 20 clusters we extracted both known and unknown phosphorylation motifs, and peptides matching these motifs were synthesized and assayed for phosphorylation-specific interaction partners using a peptide pull-down assay based on quantitative proteomics (Schulze and Mann, 2004). In contrast to the oriented peptide library approach that uses artificial degenerate peptides, we used naturally occurring peptides as baits to pull down binding partners from the cell lysate. Moreover, since the interaction partners are in competition for binding, mimicking the *in vivo* binding situation, the risk of finding kinetically unfavourable interaction motifs is minimized. Finally, this technique can potentially identify new types of domains with modification-specific binding capability.

Using the pull-down assay we identified the expected binding partners for numerous known C-terminally directed SH2 domain motifs. We also found 15 new phosphorylation-dependent interactions, mediated by phosphosites not previously shown to direct interaction. Surprisingly, we identified a new N-terminal hydrophobic motif ([LVI]-[LVI]-pY), for the SH2 domain containing inositol phosphatase SHIP2. The specificity of the motif was confirmed by mutational analysis. Surprisingly, this motif is N-terminally directed which is in contrast to previous observations showing that binding of prototypical SH2 domains are directed by C-terminal recognition (Songyang and Cantley, 1995). Until now the only other known SH2-domain binding motif which is partly directed by N-terminal recognition is the immunoreceptor tyrosine-based inhibition motif (ITIM) [ILV]-x-pY-x-x-[ILV] (Burshtyn et al., 1997; Vely and Vivier, 1997).

On a proteome level we analyzed which Gene Ontology (GO) categories were overrepresented in proteins matching the extracted motifs. We found that motifs that

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

mediate interaction in the pull-down assay are typically found in proteins involved in signal transduction, whereas non-binding motifs are found in enzymes and ion- and nucleic acid binding proteins. Thus, we estimate that one third of the *in vivo* pY sites are not directly involved in interaction via domains such as SH2 and PTB, but rather are sites that could alter the catalytic activity of enzymes or modulate the DNA binding affinity of e.g. transcription factors.

Downloaded from www.mcponline.org by on January 9, 2009

Experimental procedures

Data set preparation

Large-scale data sets of tyrosine phosphorylation sites mapped in tandem mass spectrometry (MS/MS) experiments with mammalian cell lines were collected from the literature (Brill et al., 2004; Ficarro et al., 2005; Hinsby et al., 2003; Hinsby et al., 2004; Rush et al., 2005; Salomon et al., 2003; Zheng et al., 2005) yielding a total of 847 tyrosine phosphorylation sites. To filter out phosphopeptides from closely related homologs and orthologs only unique 13-mer peptide sequences with the pY centrally positioned were considered. This reduced the MS based data set to 481 phosphopeptides distributed in 380 proteins. Furthermore, 162 experimentally verified pY peptide ligands of one phosphotyrosine binding (PTB) domain and 10 different Src-homology 2 (SH2) domains were extracted from the Phospho.ELM database (Diella et al., 2004). The 162 peptides were included in the data set as positive controls, resulting in a data set of 643 pY peptides (see Supplementary Table 3). The criteria for selecting the positive controls were the existence of a consensus binding motif and that a suitable amount of examples could be obtained.

Generation of weight matrices

13-mers of the 162 phosphopeptide ligands of the 11 respective PTB and SH2 domains (see Table I) were used to create 11 weight matrices using the weight matrix mode of EasyGibbs 1.0 (http://www.cbs.dtu.dk/biotools/EasyGibbs/) (Nielsen et al., 2004). Default settings were used except motif length was set to 13 fixed around the central pY

Downloaded from www.mcponline.org by on January 9, 2009

ASBMB Molecular & Cellular Proteomics

residue. Subsequently, all phosphopeptides in the MS based data set (481) and the positive control data set (162) were scored by each of the 11 weight matrices, and thus each phosphopeptide could be represented as a vector of the 11 weight matrix scores.

Clustering using Partitioning Around Medoids (PAM)

A matrix consisting of the 11 weight matrix scores and the 643 phosphopeptides was generated and subsequently clustered by the PAM (Partitioning Around Medoids) method (Kaufman and Rousseeuw, 1990) using the *cluster* package in R (http://www.r-project.org/). The PAM algorithm is a robust version of k-means, and it searches for a specified number of medoids (representatives), *k*, around which clusters are constructed. The clusters are generated by minimizing the sum of the dissimilarities of all observations and assigning them to their closest medoid. Using a hypergeometrical test (see statistics) the optimal number of clusters ($k = 20$) was inferred since this resulted in the best partitioning of the positive controls. We use z-scores, i.e. multiples of standard deviations from mean, to account for the different numeric ranges of the measured parameters.

The choice of an appropriate clustering algorithm is a complex one, since no given algorithm is universally superior (Fraley and Raftery, 1998; Jain et al., 1999). Rather, the best choice will depend on the data set and in particular on what constitutes a good distance measure for it. Another relevant concern is the desired outcome and whether a hierarchical or partitional result is preferable. Many sophisticated methods exist capable of automatically determining the number of 'natural clusters' in the data like the popular density-based clustering algorithms which can describe very complex non-circular relations in the data (Ester et al., 1996). It is, however, not clear whether the

Downloaded from www.mcponline.org by on January 9, 2009

ability to recognize non-circular structures in the data is beneficial in this case. Proteins which share the same features are likely to be related, and will form a circular relation in feature space. On the other hand, an elongated cluster in feature space will contain proteins which share only some features but not others, and the biological implications thereof can be quite diverse. Other than being computationally effective and easy to implement, the PAM algorithm was selected because it satisfies the need for a robust clustering algorithm and because its reliance on an Euclidean distance measure ensures that the result can be easily interpreted. The primary weakness of PAM is the need to arbitrarily select a number of clusters for the data, which in this case is overcome by the mentioned application of the hypergeometric test.

Dendrogram and Sequence Logo Plots

Weight matrices of the peptides in the 20 clusters were made using positional weighting of the 3 residues flanking the central pY residue (Nielsen et al., 2004) and used to calculate distance matrices as described (Lund et al., 2004). The distance matrices were used as input to the program *neighbor* from version 3.5 of the PHYLIP package (http://evolution.genetics.washington.edu/phylip.html). To estimate the significance of the neighbor-joining clustering we employed the bootstrap method and estimated the consensus tree by bootstrapping for 1000 repetitions as described earlier (Lund et al., 2004).

The frequencies of amino acids at particular positions in each cluster were calculated and subsequently sequence logo plots were used for graphic visualization (Schneider and Stephens, 1990). Each position in the aligned sequences corresponds to a

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

column in the logo plot. The height of the column represents the degree of conservation at that position, while the height of the individual letters is proportional to the relative frequency of this amino acid residue. The maximal height of the column for the 20 amino acid alphabet is $\log_2 20 = 4.32$ bits.

Extraction of motifs and selection of peptides to synthesize

The identified phosphomotifs in each of the 20 clusters were found using the publicly available TEIRISIAS pattern discovery tool from IBM bioinformatics (Rigoutsos and Floratos, 1998). Parameters were set so the extracted motifs were within a window of 13 residues centered on the phosphoresidue. The minimal number of literals in the motif was set to 4, and the amino acids were grouped according to their chemical nature ([AG] [DE] [FY] [KR] [ILMV] [QN] [ST] P W H C (Rigoutsos and Floratos, 1998)). For each of the 20 clusters the most abundant motif was selected and subsequently one peptide matching the motif was chosen from the respective cluster. Since multiple peptides in each cluster matched the extracted motif, peptides from mouse and peptides not previously known to be involved in phosphorylation-dependent interaction were preferred. In the few cases (3) where mouse sequences could not be obtained, peptides from humans with high homology in mouse were chosen.

Gene Ontology analysis

Gene Ontology categories were obtained from GOA Mouse database version 29.0. The extracted motifs were matched to proteins in the IPI mouse proteome version 3.20. Using

Downloaded from www.mcponline.org by on January 9, 2009

ASBMB Molecular & Cellular Proteomics

a hypergeometrical test (see next) with the total proteome as background we found the ten most overrepresented GO terms in retrieved proteins. The hits were inspected manually and the consensus GO term was assessed for each motif. For the purpose of the hypergeometrical test, each annotated GO category was taken to include all of its ancestral terms, to avoid problems with diverging levels of annotation.

Statistics

To determine if the positive controls were significantly overrepresented in specific clusters compared to the whole data set, hypergeometric sampling without replacement (Johnson et al., 1992) was performed. The hypergeometric test is a statistical test used to describe the arbitrariness of a sampling without replacement from a background of true or false examples. The probability (P) to observe a given or more extreme situation by a pure coincidence is given by the hypergeometric distribution:

$$P(X = x|N, M, K) = \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}}$$

where N is the total number of peptides, M is the number of peptides in the given set, K is the number of peptides in a particular cluster, and x is the number of K that belongs to M. A Bonferroni correction was performed in order to correct for multiple comparisons. In the case of GO analysis, we performed the test once for each GO category present in the data, and evaluated the probability of sampling the set of retrieved proteins from the background of the total proteome by mere chance, considering a protein 'true' or 'false' depending on whether it had been assigned the category in question. The end result of

Downloaded from www.mcponline.org by on January 9, 2009

this test was one p-value for each GO category, describing the degree of over-representation of that particular assignment in the retrieved set against the background of the entire proteome.

Cell culture

Mouse C2C12 muscle cells were grown in arginine- and lysine deficient Dulbecco's modified Eagle's medium (DMEM) with 10% dialyzed fetal bovine serum (FBS) for at least 5 passages and then switched to 2% dialyzed FBS to differentiate the cells for 8 days. In accordance with the stable isotope labeling of amino acids in cell culture (SILAC) procedure, one cell population was supplemented with normal isotopic abundance L-12 arginine (Sigma) and L-lysine, and the other with >99% isotopic abundance $^{13}C_6$ $^{15}N_4$-arginine and $^{13}C_8$-lysine (Aldrich), as described (Ong et al., 2003). Thereby full labeling of all proteins was achieved.

Peptide synthesis and pull-downs

Desthiobiotinylated peptides were synthesized on a solid-phase peptide synthesizer using amide resin (Intavis, Germany). All peptides were designed as 15-mers with the pY residue placed centrally at position 7 or 8, expect one peptide from cluster 1 (see Table I) which was 20 amino acids long. The peptides were synthesized with an N-terminal biotin and an SG dipeptide linker. Peptides were synthesized as pairs in phosphorylated and a non-phosphorylated 'control' form. The identity and purity of the synthesized peptides was confirmed by mass spectrometric analysis. For pull-downs, 1.5 nmol of immobilized

Downloaded from www.mcponline.org by on January 9, 2009

peptide was added to an average of 1.5 mg of cell lysate. Dynabeads MyOne Streptavidin were saturated with biotinylated peptide prior to incubation with cell lysates. Cells were lysed as described in (Schulze et al., 2005) and equal amounts of protein were incubated over night with the respective immobilized peptides at 4°C. After three rounds of washing with lysis buffer, beads of pull-down pairs with 'phosphorylated' and 'control' were combined (Schulze and Mann, 2004) and bound proteins were eluted using 16mM biotin. Eluted proteins were precipitated and subsequently digested with trypsin for LC–MS/MS analysis.

LC/MS/MS, database searching and quantitation

After reduction in 1 µg DTT and alkylation with 5 µg iodoacetamide the eluted proteins were in-solution digested with 1 µg endoproteinase Lys-C (Wako) for 3h at room temperature. Subsequently, samples were diluted with 4 volumes of 50mM NH_4HCO_3 and further digested with 1 µg trypsin (Promega) overnight at room temperature. Peptide mixtures were desalted on STAGE tips (Rappsilber et al., 2003) and loaded onto reversed phase analytical columns for liquid chromatography (Ishihama et al., 2002). Peptides were eluted from the analytical column by a multistep linear gradient running from 2 to 40% acetonitrile in 100 min and sprayed directly into the orifice of an LTQ-FT or an LTQ-Orbitrap mass spectrometer (Thermo Electron, Germany). Proteins were identified by tandem mass spectrometry (MS/MS) by information-dependent acquisition of fragmentation spectra of multiply charged peptides. The peaklist was generated using in house software, raw2msm version 1.2, with default settings. The identified proteins were then searched against the mouse International Protein Index (IPI) Database using the

Downloaded from www.mcponline.org by on January 9, 2009

Mascot (version 2.1.0) algorithm (Perkins et al., 1999). The MS/MS ion search parameters were set with the following parameters: Enzyme specificity for trypsin: trypsin/P + DP; Maximum number of miss cleavages: 2; Fixed modifications: Carbamidomethyl-cysteine; Variable modifications: Oxidation (M), N-Acetyl (Protein), Deamidation (NQ), $^{13}C_6$ $^{15}N_4$-arginine, $^{13}C_8$-lysine, Pyro (N-term QC); Mass tolerance for precursor ions: 5 ppm; Fragment mass tolerance: 0.5 Da. Database version: IPI_mouse mouse_v314 with 68655 entries. Common contaminants like human keratins were manually added. No species specific restrictions were used. MSQuant (http://msquant.sourceforge.net) was used for quantitation and spectra validation. MSQuant uses Peak area and XIC for quantitation.

Determination of significant binding partners

Intensity ratios of labeled to unlabeled forms of each validated tryptic peptide, and the associated average ratio for the whole protein were obtained by MSQuant. We used 'cross-over' experiments, in which the specific interaction partners were required to have inverse ratios compared to the 'normal' experiment (Schulze and Mann, 2004). A significant binder was defined as a protein with a log ratio at least 3 standard deviations over the log average ratio (>3 log(σ) + log(μ)) of all the proteins identified in a pull-down experiment. Furthermore, the binder had to be confirmed in at least two pull-down experiments ('normal' and 'cross-over' experiment) and we only report specific binders for the phosphopeptide and not the non-phosphorylated peptides. Finally, at least one peptide had to have a score above 30, corresponding to $p<0.05$. In the 64 pull-downs performed we identified a few sequence unspecific binders with high affinity to either the

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

phosphorylated or non-phosphorylated peptides (Staphylococcal nuclease domain-containing protein, Eukaryotic translation initiation factor, peptidylprolyl isomerase B, RNA-binding protein SiahBP, and RIKEN cDNA 2410104I19). These proteins were excluded since we consider them as false positive binders, i.e. they bind in a sequence unspecific manner and occasionally strongest to the non-phosphorylated peptide. In all the pull-down experiments an average of 140 ±41 proteins was quantified with an average ratio of 1.217 ±0.529.

Downloaded from www.mcponline.org by on January 9, 2009

ASBMB
Molecular & Cellular Proteomics

Results

A new clustering approach for motif extraction and classification

In order to discover new interaction motifs in large aligned peptide data sets, we have developed a method that partitions the sequences based on similarity with proteins previously known to be involved in interaction. The method is generally applicable for peptide-interaction datasets. In this work we have clustered and classified motifs in the pY proteome. A flow chart of the method is presented in Figure 1. In short, we used verified pY ligands of the PTB and SH2 domains as a backbone to cluster a data set of uncharacterized phosphopeptides mapped in MS-based proteomic experiments with mammalian cell lines. From each clusters the most conserved motif was extracted. In order to classify the function of each motifs in an unbiased and systematic manner we conducted both experimental and bioinformatical investigations. Peptides matching the motifs were assayed for binding partners using a peptide-protein interaction screen based on quantitative proteomics (Schulze et al., 2005; Schulze and Mann, 2004). On a proteome level we analyzed which GO categories were overrepresented in proteins matching the extracted pY motifs. Thus, we have established a workflow that extracts, verifies and classifies motifs in phosphoproteome.

Clustering (Fig. 1A-C)

In detail, we extracted 162 known 13-mer phosphopeptide ligands of ten different SH2 domains and a general PTB domain from the phospho-ELM database (Diella et al., 2004). In this database, which is the key repository for high-quality annotated

Downloaded from www.mcponline.org by on January 9, 2009

phosphorylation sites, these 11 domains had the highest number of annotated ligands (see Table 1 for details). Thus, the 162 substrates of the 11 domains represent the broadest available in vivo dataset of pY ligands. These ligands were aligned with the pY centrally positioned, and used to generate a position specific weight matrix for each of the 11 domains. The 481 phosphopeptides in the MS-based data set and the 162 ligands of the PTB and SH2 domains ('positive controls') were then scored by each of the 11 weight matrices. Consequently, each phosphopeptide was represented as a vector of the 11 weight matrix scores (Fig. 1A). Based on this profile of vectors we clustered the total of 643 peptides using the k-means derived Partitioning around Medoids (PAM) algorithm (Bateman et al., 2002). The PAM algorithm searches for a predefined number of medoids (representatives) around which clusters are constructed. We tried various cluster sizes and by using a hypergeometrical distribution test the number of clusters was set to 20, since this gave the statistically best partition of the positive controls into different clusters (Fig. 1B and C).

We were able to obtain a convenient fit of the model since the positive controls were distributed such that all were statistically overrepresented in separate clusters ($p<0.05$) (see Table 1). For example, 10 out of a total of 10 ligands of the PTB domains were grouped in cluster 1 ($p = 7.34E-12$). Only the 31 ligands for the Grb2 SH2 domain were split up in two significant groups (clusters 6 and 7). Furthermore, eight out of the 20 clusters did not contain a significant overrepresentation of the known ligands used.

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

Motif extraction (Fig. 1D)

From each of the 20 clusters the most conserved motif was extracted with the TEIRISIAS pattern discovery tool from IBM bioinformatics (Fig. 1D)(Rigoutsos and Floratos, 1998) as described in the methods section. The 20 identified motifs are presented in Table I followed by the number matches to peptides in the particular cluster and in the total data set. For example, a unique motif, N-P-x-pY-x-[ST], was extracted from cluster 1, since all peptides (7) that matched this motif were in this cluster. To compare the 20 identified motifs with already characterized interaction motifs, we matched each motif to a comprehensive library of pY interaction motifs in the Human Protein Reference Database (HPRD) (Peri et al., 2003). Using the same example again, the motif extracted from cluster 1 could be matched to the N-P-x-pY motif described for the SHC and IRS-1 PTB domains (Yaffe, 2002), although our extended motif contains a Ser or Thr residue at position +2 from the pY residue. Considering that the ligands (the positive controls) of the PTB domain were grouped in cluster 1, it is not surprising that we extracted the N-P-x-pY-x-[ST] motif in this cluster, however, only three of the seven matching peptides in cluster 1 were from these PTB ligands (data not shown). Since this was a general trend it shows the ability of the clustering method to gather previously uncharacterized peptides with high sequence similarity to known ligands of the different pY binding domains. In total, 16 of the 20 identified motifs could be matched to the motif library, showing an overall consistency between the positive controls and the extracted and matched motifs. In four clusters new pY motifs were identified.

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

Gene Ontology analysis (Fig. 1E)

We employed a Gene Ontology (GO) analysis of the 20 extracted motifs on a proteome level to investigate which type of proteins the extracted motifs generally are present in (Fig 1.E). We retrieved all proteins in the mouse proteome that matched the motifs and used a hypergeometrical test with the total proteome as background and hereby found the ten most overrepresented GO terms in retrieved proteins (see Methods for details). Using the same example again, the N-P-x-Y-x-[ST] motif from cluster 1 was overrepresented in proteins involved in processes like 'receptor activity', 'intrinsic to membrane' with the consensus parent GO term assessed to be 'transmembrane receptor activity' (see Supplementary Table II). Thus, analyzing the motif on a proteome level indicates that proteins containing the motif are involved in early signaling transduction. This makes sense in the light that this is a motif for the PTB domain, which are found in proteins that function as molecular scaffolds and adaptors in signaling pathways (Yaffe, 2002).

pY-specific interaction partners (Fig. 1F)

In order to experimentally verify the 20 extracted motifs we used a phosphorylation-specific peptide-protein interaction screen (Fig. 1F) (Schulze et al., 2005; Schulze and Mann, 2004). This assay is based on differential labeling of proteins using stable isotope labeling of amino acids in cell culture (SILAC) allowing to distinguish specific binders from background binders by their isotope ratios determined by quantitative mass spectrometry (Ong et al., 2002; Ong et al., 2003). The peptides are synthesized in

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

phosphorylated and a non-phosphorylated form and used as baits to pull down competing binding partners from cell lysate, thus mimicking the *in vivo* binding situation.

We synthesized peptide-pairs matching the 20 extracted motifs. If there were multiple matches we chose peptides with pY sites not previously known to mediate interaction. Using this experimental approach we could test our clustering and motif extraction method and investigate the relevance of known motifs in a near *in vivo* situation, i.e. endogenous proteins competing for binding, and potentially discover binding partners for novel motifs. Again using cluster 1 as example, we synthesized a peptide pair from the Engulfment and cell motility protein 2, where the Tyr-48 residue was either phosphorylated or non-phosphorylated. This is an uncharacterized phosphosite identified in a large-scale phosphoproteomic study (Rush et al., 2005), and has not previously been shown to direct pY-dependent interaction. Using this peptide pair as bait we retrieved one specific binder with a ratio more than three standard deviations over the log mean out of a total of 162 background binders. This protein, SHC-transforming protein 1 (SHC), which contains both a PTB and a SH2 domain, had a total Mascot score 1654 with 15 identified peptides of which nine were quantifiable (see Supplementary Table I). Theoretically, it could be either the SH2 or the PTB domain that binds the bait phosphopeptide, however, since the peptide matches the consensus N-P-x-pY motif known to direct PTB domain binding, it is most likely that SHC binds to the peptide through its PTB domain.

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

Assaying the pY sequence space for interaction partners

The identified phosphospecific binding partners of the representative peptides from each cluster can be seen in Table I. Besides the aforementioned SHC protein, these proteins all contain SH2 domains, making it very likely that this domain governed the phospho-specific interaction. The majority (13 of 20) of the peptides retrieved one or more interaction proteins. Of these proteins seven were unique, since some proteins were identified several times. This is not surprising since some of the clusters were close to each other in sequence space resulting in extraction of similar motifs and ultimately retrieving the same interaction partners.

To get an overview of the results of the pull-down experiments, the GO analysis, and the sequence similarity between the different clusters, we generated weight matrices of the peptides in the individual clusters and constructed a dendrogram based on an alignment of these matrices (Lund et al., 2004). The tree can be seen in Figure 2 together with sequence logo plots, where the height of each position represents the degree of conservation (Schneider and Stephens, 1990). The logo plots visually illustrate a successful partitioning, since each cluster has a distinct pattern, where particular amino acids are highly abundant at specific positions flanking the central pY residue. The tree is colored according to the identified interaction partners retrieved by peptides matching the motifs in the different clusters. For example, clusters 8 and 11 are close in sequence space with an overrepresentation of hydrophobic residues, especially methionine, at position +3 from the central tyrosine residue. Rather than being distinct clusters, these are more likely to be subsets of the same motif. Thus, from these two clusters the motifs pY-[ILMV]-x-M-x-P and pY-[ILMV]-P-M-x-P were extracted and matched to the consensus pY-x-x-M

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

in the library of motifs. Accordingly, the two peptide-pairs synthesized from these clusters both retrieve the PI3K-p85α protein. In the same manner the majority of the expected interaction partners were identified using peptides matching to the well characterized C-terminally directed pY motifs, such as pY-x-N, pY-x-x-P, and pY-[IVL]-x-[IVL], that retrieved Grb2, RasGAP, and SHP2, respectively. This illustrates the clear consistency between the sequence similarity of the clusters, the conserved residues in the motifs, and the interaction partners identified.

There were four clusters (3, 9, 10 and 18) where we did not identify the partners (Vav, Fes/Fps and STAT) as expected from the signature of motifs alone (see Table I). For instance, we did not pull down the SH2 domain protein Fps/Fes when using peptides matching a pY-[ED]-x-[IVL] motif (cluster 9 and 10), which has previously been shown to direct binding (Songyang et al., 1994). This motif was defined using *in vitro* oriented peptide library experiments, which have the inherent risk of defining motifs that are not relevant *in vivo*. Whether this is the case, the Fps/Fes protein was not present in the cell lysate or because of technical limitations remains unclear.

In total, we here report 15 phosphorylation-dependent interactions mediated by phosphosites not previously known to direct protein interaction (see Supplementary Table 1).

Non-binding pY motifs

We observed that motifs that do not mediate interaction in the pull-down assay are typically found in proteins involved in other processes than signal transduction (see Fig. 2). Particularly interesting, we extracted a highly conserved H-[ST]-G-x-K-P-pY-x-C-x-

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

x-x-C-G motif from a number of closely related peptides from Cys_2His_2 zinc finger proteins (zf-C2H2) concentrated in cluster 19, that did not pull down any phosphorylation-specific interaction partner. The phospho-site of the first tyrosine residue in the zf-C2H2 domain was also described recently in a study that mined for novel motifs in the phosphoproteome (Schwartz and Gygi, 2005), although the identified motif in this work (E-x-x-pY) was different from our top-scoring motif (H-[ST]-G-x-K-P-Y-x-C-x-x-x-C-G), which does not contain an acidic residue in position -3. However, our second best motif in cluster 19 (H-x-G-E-x-x-pY) closely resembles that reported by Schwartz and Gygi, 2005.

The H-[ST]-G-x-K-P-Y-x-C-x-x-x-C-G motif is extremely specific for proteins containing the zf-C2H2 domain: Out of 33758 proteins we retrieved 656 matches, all of which had the GO term 'nucleic acid biding' (GO:0003676) (see Supplementary table 2), while 647 of the 656 matches had the term 'zinc ion binding' (GO:0008270) ($p<1E-100$).

Recently, a role for phosphorylation of zf-C2H2 domains in inhibition of transcription has been suggested (Dovat et al., 2002; Saydam et al., 2002), supposedly as a consequence of the negatively charged phosphomoity that reduces DNA affinity (Jantz and Berg, 2004). Indirectly, our results support this; since we did not retrieve any interaction partner for the synthesized phosphopeptide matching the zinc finger motif, it is unlikely that this motif directs protein-protein interaction, but rather phosphorylation of this motif modulates nucleic acid binding.

Similarly, the novel motifs from cluster 16 and 20 could mediate mechanisms other than protein-protein interactions, for example, a kinase motif that mediates enzyme activation, nucleic acid binding, protein folding etc. Supporting this, we found these

Downloaded from www.mcponline.org by on January 9, 2009

ASBMB Molecular & Cellular Proteomics

motifs to be present in proteins overrepresented in proteins with GO terms 'ion binding' and 'kinase activity', respectively. Likewise, the motifs from cluster where we did not find the expected partners (3, 9, 10 and 18) are all except motif 3 overrepresented in proteins involved in enzymatic processes and ion binding (see Fig. 2 or Supplementary table 2).

This indicates that the motifs not mediating protein binding could govern processes such as phosphorylation-dependent enzyme activation and nucleic acid binding. Taken together with the results from the pull-down experiment, where 13 of 20 motifs mediated interaction, we estimate that one-third of the pY motifs in the proteome mediate other processes than interaction through prototypic SH2 and PTB domains.

Identification of a new N-terminal hydrophobic pY binding motif for SHIP2

From cluster 4 we extracted a new N-terminal hydrophobic motif [DE]-x-x-x-[ILV]-[ILV]-pY. We synthesized a peptide pair from the Neural Wiskott-Aldrich syndrome protein (N-WASP) matching the motif where Tyr-253 was either phosphorylated or non-phosphorylated. This phospho-Tyr-253 was identified in a large-scale phosphoproteomic experiment (Rush et al., 2005). It has previously been shown that phosphorylation of Tyr-253 modulates localization of N-WASP from nucleus to cytoplasm and thereby possibly stimulating cell migration (Wu et al., 2004). Using the N-WASP peptide-pair in our quantitative proteomics experiment, we find the SH2-domain bearing inositol 5-phosphatase 2 (SHIP2) as a significant binding partner with 12 peptides (ratio 9.5±1.5) out of 131 proteins background partners (ratio 1.2±0.9) (see Supplementary Table I). To address the specificity of this N-terminal hydrophobic phospho-motif we repeated the

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

experiment with the only difference being that the two hydrophobic residues were mutated to alanines (-A-A-pY-) and paired this peptide with the wild type N-WASP phosphopeptide (-V-I-pY-) (Figure 3). The SHIP2 protein bound specifically to the wild type phospho-peptide (ratio 10.2±3) but not the mutated phospho-peptide, confirming the specificity of the hydrophobic motif. In a third interaction experiment we obtained similar results by using a phospho-peptide where all three residues in the extracted motif were mutated (see Supplementary Table I). This supports the notion that recognition is based on two hydrophobic amino acids adjacent to the pY and we could confine the motif to [ILV]-[ILV]-pY.

We next extracted three additional peptides from the same cluster matching the [ILV]-[ILV]-pY motif. Two out of three of these phospho-peptide-pairs also specifically pulled down SHIP2 (see Supplementary Table I). Together, these experiments show that this indeed is a generic binding motif for SHIP2.

Similarly to previous studies using this phospho-specific pull-down method (Schulze et al., 2005; Schulze and Mann, 2004) in this work we only retrieved specific interaction partners containing either an SH2 or a PTB domain. Since SHIP2 contains an SH2 domain it is highly unlikely that interaction between SHIP2 and the phospho-peptides could be mediated by other domains than through the SH2 domain.

Since the observed interaction seems to be one of the first examples where N-terminal residues of the ligands guide SH2 domain binding, we wanted to exclude the possibility that the peptide was bound in a reverse fashion. We synthesized the N-WASP peptide-pair with the reverse sequence (GTKEIFDpYIVKSTER) and found that SHIP2 was not retrieved using this scrambled peptide pair as bait (see Supplementary Table I).

Molecular & Cellular Proteomics ASBMB

Downloaded from www.mcponline.org by on January 9, 2009

Thus, the assay has directional specificity and only the two hydrophobic residues N-terminal and not C-terminal from the pY direct SH2 domain mediated SHIP2 binding.

To investigate the significance of the motif on a proteome level, we used the aforementioned GO analysis, and found that proteins containing the motif indeed are involved in signal transduction. The [ILV]-[ILV]-Y is not particularly specific in itself, since it matches 3618 times, which is about 10% of all proteins in the mouse proteome. However, out of the 3618 proteins, 2495 can be backtracked to the term 'cell surface receptor linked signal transduction' (GO:0007166) ($p<1E-100$). In summary, the N-terminal hydrophobic motif we have characterized mediates SHIP2 interaction and is generally overrepresented in proteins involved in signal transduction.

.

ASBMB Molecular & Cellular Proteomics

Downloaded from www.mcponline.org by on January 9, 2009

Discussion

Currently, there is a unique opportunity to mine and classify pY motifs in the proteome: 1) a significant portion of the pY sites have been mapped by MS-based proteomics experiments; 2) although not exhaustively, classical biochemical studies have pinpointed interaction between SH2 and PTB domain containing proteins and specific pY ligands; 3) advances of high through-put validation methods, like the peptide pull-down assay used in this study, make it possible to validate dozens of interaction motifs. Together, these developments constitute the basis for discovering and validating pY motifs on a global level.

Methodological considerations

Compared to previously published methods that mine for motifs in large-scale phosphoproteomic data sets we partition by similarity with functionally characterized peptides prior to motif extraction. By this clustering approach we obtain high resolution and can extract meaningful patterns from functionally and physically related groups of peptides. We used the binding ligands of PTB and SH2 domains both as a clustering backbone and as positive controls in the clustering and could consequently obtain a satisfactory fit of the model. Ideally, as more interaction data becomes available one would use independent controls; however this was not possible due to data limitations. For all the 20 extracted motifs we could either detect a match to an existing pY motif, retrieve a binding partner, or obtain a meaningful GO term for all proteins containing the

Molecular & Cellular Proteomics ASBMB

Downloaded from www.mcponline.org by on January 9, 2009

motif, and thus, we estimate that we have an insignificantly low false positive motif extraction rate in our method.

In the few cases where we did not retrieve the expected binding partner from the pY motif alone, i.e. Fps/Fes, STAT and Vav, it can be speculated that 1) the motifs, originally defined *in vitro*, could be low affinity motifs *in vivo*; 2) the proteins were not expressed in the C2C12 muscle cell line used in this study or 3) it could be due to a technical limitation in our assay.

Implications for a new motif for SHIP2

One of our findings is that the SH2 domain containing 5' inositol phosphatase SHIP2 binds to a novel N-terminal hydrophobic motif, [ILV]-[ILV]-pY . The specificity of this motif was confirmed by mutational analysis. A scrambled peptide pair with the reverse sequence and thus with the motif C-terminal of the pY does not retrieve SHIP2, confirming that the recognition lies on the N-terminal side of the pY. SHIP2 is also retrieved by two other peptide pairs matching the motifs, indicating this is a generic motif for SHIP2 binding.

In general, the C-terminal residues pY +1 and +3 are considered as the most important for the binding specificity of prototypic SH2 domains (Pawson, 2004; Yaffe, 2002). Interestingly, SHIP2's SH2 binding motif we describe is N-terminal, indicating that the peptide binding groove of some SH2 domains also accommodates residues N-terminal of the pY. In agreement with this, the binding properties of the tandem SH2 domains of the protein tyrosine phosphatase SHP-2 are governed by residues pY -2 to +5 (Eck et al., 1996). SHP-1 and SHP-2 both bind the [ILV]-x-pY-x-x-[ILV] ITIM motif in

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

the cytoplasmic part of Fc receptors and the pY -2 hydrophobic residues has specifically been shown to mediate binding (Beebe et al., 2000; Burshtyn et al., 1997; Huyer and Ramachandran, 1998). In contrast to the prototypic SH2 domains that have an Arg in the αA2 binding-pocket groove, the tandem SH2 domains of the SHP-1 and SHP-2 phosphatases instead have Gly (Liu et al., 2006). Supposedly, this creates a gap which is filled by the side chain of the pY -2 residue of the bound peptide. Supporting this hypothesis, it has been shown that a single point-mutation in αA2 Gly → Arg disrupts the pY -2 mediated binding specificity of SHP-2 (Huyer and Ramachandran, 1998). Furthermore, it is known that SAP and Eat2 SH2 domains in part are directed by N-terminal binding to ITIM motifs (Poy et al., 1999).

The binding motif of the SH2 domain of SHIP2 has not previously been investigated using degenerate peptide binding experiments, however, the ITIM motif has also been reported as docking point for SHIP2's SH2 domain (Bruhns et al., 2000; Muraille et al., 2000). Combined with our observations, this indicates that the binding specificity of SHIP2 may be conferred by hydrophobic residues immediately upstream of the pY with contributions from downstream hydrophobic patches. To our knowledge, other that that of SHP, this is the only other reported N-terminally directed pY binding motif.

It could be speculated that the SHIP inositol phosphatases could have a binding mechanism similar to that of SHP protein tyrosine phosphatases. However, in contrast to SHP the SH2 domains of the SHIP phosphatases resemble the prototypic SH2 domain since they have the highly conserved αA2 Arg (Liu et al., 2006), indicating that the N-terminal directed binding mechanisms differ between SHIP and SHP phosphatases. Since

Downloaded from www.mcponline.org by on January 9, 2009

ASBMB Molecular & Cellular Proteomics

the crystal structure of the SHIP2 phosphatase with a bound ligand has not been resolved, the specific binding mechanisms are yet to be described.

SHIP2 is involved in membrane signaling by dephosphorylating the 5'-phosphate group of the key secondary messenger phosphatidylinositol 3,4,5-trisphosphate (PIP3). Through this action SHIP2 inhibits PI3K mediated receptor tyrosine kinase (RTK) signaling, since PIP3 is generated by PI3K (Damen et al., 1996). The new motif we describe for SHIP2's SH2 domain fits into this overall function, since a GO term analysis of all proteins containing the [ILV]-[ILV]-pY motif revealed that these proteins are involved in 'cell surface receptor linked signal transduction'. Presumably, SHIP2 could bind to a number of yet unidentified membrane signaling proteins through its SH2 domain and in this way be translocated to the membrane or act cooperatively with these proteins. Negative regulators of PI3K and PIP3 are attractive as anti-obesity and diabetes drug targets, since PI3K is the main effector of insulin signaling. Recently, a role for SHIP2 as a candidate for such therapeutic intervention has been proposed by studies of SHIP2 knockout mice (Marion et al., 2002; Sleeman et al., 2005). Thus, the new motif described in this report may not only be involved in regulation of SHIP2 mediated signal transduction but could also be relevant for medical use.

Conclusions

This work presents the first system-wide approach to mine the proteome for pY interaction motifs using both bioinformatics methods and experimental validation. Strikingly, 16 of the 20 motifs extracted could be matched to previously described interaction motifs. Our experimental validation shows that the majority of the pY

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

interaction motifs that previously have been defined *in vitro* are also able to pull down interaction partners from complex lysate mixtures. The GO analysis revealed that motifs that mediate interaction in the pull-down assay are found in proteins involved in signal transduction, whereas remaining non-binding motifs are found in enzymes and ion- and nucleic acid binding proteins. This raises the intriguing possibility that about one third of the *in vivo* pY sites are not directly involved in interaction via domains such as SH2 and PTB, but rather are sites that could alter the catalytic activity of enzymes or modulate the DNA binding affinity of e.g. transcription factors.

Perspectives

The developed clustering method is applicable to other types of complex large-scale data sets of post-translational modification where a substantial amount of peptide-protein interactions have been identified. As MS-based methods map more modifications such as acetylation and methylation sites, such interactomes could be mined for conserved motifs and assayed for binding partners in a similar manner. Combining proteomics and bioinformatics enables one to do large-scale screens in an unbiased way and thus allows one to reconfirm previous knowledge and discover new mechanisms at the same time. As mass spectrometric streamlining and automation advances, entire post-translational modification proteomes can be mapped, binding motifs identified, and thus ultimately unravel the specificity of signaling networks.

Downloaded from www.mcponline.org by on January 9, 2009

ASBMB Molecular & Cellular Proteomics

Acknowledgements

We thank members of the Department of Proteomics and Signal Transduction and the Center for Biological Sequence Analysis for their comments to the manuscript. This work was supported by Interaction Proteome (IP), FP6, contract no.: LSHG-CT-2003-505520, a grant from the European Commission in the 6th framework program, and by the Danish Platform for Integrative Biology (DPIB), a grant from the Danish National Research Foundation (Danmarks Grundforskningsfond). Finally, we thank the Danish Research Agency (Forskningsstyrelsen) for funding.

Downloaded from www.mcponline.org by on January 9, 2009

Figure legends

Table I

Table I. *Clustering and motif extraction of the pY proteome*. The Partitioning around Medoids (PAM) algorithm is used to partition the data set by sequence similarity with the known ligands of the pY binding domains ('positive controls'). The first and second columns show the cluster number and the size of the clusters, respectively. The algorithm's ability to significantly partition ($p<0.05$) the positive controls into different clusters is shown in the third column. For example, eight out of a total of 10 ligands of the Crk SH2 domain are grouped in cluster 2, corresponding to a significant overrepresentation in a Bonferonni corrected hypergeometric sampling test (p=3.67E-08). The most conserved motif in each cluster is extracted and stated in column four. The pY residue is indicated in red and an 'x' represents any amino acid. Also indicated is the number of occurrences of the motif in the respective cluster and in the total data set. The identified motifs are matched to a library of known pY binding motifs and the expected binding partner is indicated (column 5). 15-20 mer peptides matching the motifs were synthesized in pairs – one with a pY as indicated in red. Furthermore, the parent protein of the peptide is given by SwissProt entry name and the position of the pY is stated. The pY-dependent interaction partner(s) identified in a quantitative proteomic peptide-protein screen can be seen in the last column (see also Supplementary Table 1).

Downloaded from www.mcponline.org by on January 9, 2009

Figures

Figure 1

Figure 1. *Overview of the motif extraction and validation strategy*. A. Verified pY-peptide ligands (162) of different PTB and SH2 domains are used to create 11 specific weight matrices that subsequently score both a large-scale data set of 481 pY peptides identified in MS experiments and the 162 ligands themselves. All peptides are 13-mers with the pY (red) centrally positioned. Consequently, each phosphopeptide can be represented as a vector of the 11 weight matrix scores. B. This is used as input for the k-means derived Partitioning around Medoids (PAM) algorithm (Bateman et al., 2002) and the data set is split up into clusters (20, see text for details) depending on sequence similarity of the ligands used in the weight matrix generation. We use z-scores, i.e. multiples of standard deviations from mean, to account for the different numeric ranges of the measured parameters. C. Sequence logo plots (Schneider and Stephens, 1990) of the peptides in the different clusters are used to visually inspect the result of the partitioning (see Figure 2). D. After partitioning the most conserved motif from each cluster is extracted (Rigoutsos and Floratos, 1998). E. A GO analysis of the motifs on a proteome level is conducted in order to analyze which functional class of proteins the motifs generally are present in compared to the whole proteome as background. F. A peptide pull-down assay based on quantitative proteomics (Schulze and Mann, 2004) is used to evaluate the pY-specific binding capability of the motifs. Peptides matching the motifs are synthesized in phosphorylated and non-phosphorylated pairs and used to pull down pY-specific interaction partners from the lysate of C2C12 muscle cells.

Downloaded from www.mcponline.org by on January 9, 2009

ASBMB Molecular & Cellular Proteomics

ASBMB
Molecular & Cellular Proteomics

Figure 2

Figure 2. *Dendrogram representing the in vivo pY sequence space, motifs, and binding partners.* Peptides in the 20 different clusters are used to generate weight matrices that subsequently are used as input in a phylogenetic alignment. The tree is a consensus tree of 1000 bootstrap trees (Lund et al., 2004). The tree represents the distance between the clusters in sequence space, which is also visually illustrated by the sequence logo plots (Schneider and Stephens, 1990) of the peptides in each cluster. The color of the branches is based on the pY-dependent interaction partners that were experimentally identified using a peptide-protein interaction screen. Both novel and previously known consensus motifs (see Table 1, column 5) that govern these specific interactions are indicated in the same color. Branches and extracted motifs are grey if no interaction partners were retrieved via the motifs. Furthermore, overrepresented GO terms from proteins in the whole mouse proteome containing the motifs are stated (see text for details). Note that the motifs that do not retrieve a specific binding partner (grey) are typically found in proteins mediating processes such as 'ion binding', 'catalytic activity' and 'nucleic acid binding'.

Downloaded from www.mcponline.org by on January 9, 2009

Figure 3

Figure 3. *Identification of a new hydrophobic N-terminal motif for SHIP2.*

A) Mass spectrum of a doubly charged $^{13}C_6$ $^{15}N_4$-Arg-labeled tryptic peptide from Myosin Light Chain 1 pulled down with the RETSKVIYDFIEKTG peptide from N-WASP. The Myosin Light Chain 1 protein is an unspecific binding protein, since the ratio of the ion

intensities of the two differentially labeled peptides is 1:1. B) Mass spectrum of a doubly charged $^{13}C_6$ $^{15}N_4$-Arg-labeled tryptic peptide from SHIP2. SHIP2 specifically binds to the phosphorylated peptide from N-WASP as indicated by the larger peak intensity of the labeled peptides. C) When mutating the VIpY motif to AApY and comparing these two 15-mer phosphopeptides directly in a pull-down, SHIP2 is still retrieved as a specific binding partner to the VIpY motif, confirming the specificity of the hydrophobic motif immediately N-terminal to the pY.

Downloaded from www.mcponline.org by on January 9, 2009

Literature

Ballif, B.A., Villen, J., Beausoleil, S.A., Schwartz, D. and Gygi, S.P. (2004) Phosphoproteomic analysis of the developing mouse brain. *Mol Cell Proteomics*, **3**, 1093-1101.

Bateman, A., Birney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S.R., Griffiths-Jones, S., Howe, K.L., Marshall, M. and Sonnhammer, E.L. (2002) The Pfam protein families database. *Nucleic Acids Res*, **30**, 276-280.

Beausoleil, S.A., Jedrychowski, M., Schwartz, D., Elias, J.E., Villen, J., Li, J., Cohn, M.A., Cantley, L.C. and Gygi, S.P. (2004) Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc Natl Acad Sci U S A*, **101**, 12130-12135.

Beebe, K.D., Wang, P., Arabaci, G. and Pei, D. (2000) Determination of the binding specificity of the SH2 domains of protein tyrosine phosphatase SHP-1 through the screening of a combinatorial phosphotyrosyl peptide library. *Biochemistry*, **39**, 13251-13260.

Bork, P. and Koonin, E.V. (1996) Protein sequence motifs. *Curr Opin Struct Biol*, **6**, 366-376.

Brill, L.M., Salomon, A.R., Ficarro, S.B., Mukherji, M., Stettler-Gill, M. and Peters, E.C. (2004) Robust phosphoproteomic profiling of tyrosine phosphorylation sites from human T cells using immobilized metal affinity chromatography and tandem mass spectrometry. *Anal Chem*, **76**, 2763-2772.

Bruhns, P., Vely, F., Malbec, O., Fridman, W.H., Vivier, E. and Daeron, M. (2000) Molecular basis of the recruitment of the SH2 domain-containing inositol 5-phosphatases SHIP1 and SHIP2 by fcgamma RIIB. *J Biol Chem*, **275**, 37357-37364.

Burshtyn, D.N., Yang, W., Yi, T. and Long, E.O. (1997) A novel phosphotyrosine motif with a critical amino acid at position -2 for the SH2 domain-mediated activation of the tyrosine phosphatase SHP-1. *J Biol Chem*, **272**, 13066-13072.

Damen, J.E., Liu, L., Rosten, P., Humphries, R.K., Jefferson, A.B., Majerus, P.W. and Krystal, G. (1996) The 145-kDa protein induced to associate with Shc by multiple

Downloaded from www.mcponline.org by on January 9, 2009

cytokines is an inositol tetraphosphate and phosphatidylinositol 3,4,5-triphosphate 5-phosphatase. *Proc Natl Acad Sci U S A*, **93**, 1689-1693.

Diella, F., Cameron, S., Gemund, C., Linding, R., Via, A., Kuster, B., Sicheritz-Ponten, T., Blom, N. and Gibson, T.J. (2004) Phospho.ELM: a database of experimentally verified phosphorylation sites in eukaryotic proteins. *BMC Bioinformatics*, **5**, 79.

Dovat, S., Ronni, T., Russell, D., Ferrini, R., Cobb, B.S. and Smale, S.T. (2002) A common mechanism for mitotic inactivation of C2H2 zinc finger DNA-binding domains. *Genes Dev*, **16**, 2985-2990.

Eck, M.J., Pluskey, S., Trub, T., Harrison, S.C. and Shoelson, S.E. (1996) Spatial constraints on the recognition of phosphoproteins by the tandem SH2 domains of the phosphatase SH-PTP2. *Nature*, **379**, 277-280.

Ester, M., Kriegel, H.P., Sander, S. and Xu, X. (1996) A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *Proc. 2nd Int. Conf. on Knowledge Discovery and Data Mining*, **1**, 226-231.

Ficarro, S., Chertihin, O., Westbrook, V.A., White, F., Jayes, F., Kalab, P., Marto, J.A., Shabanowitz, J., Herr, J.C., Hunt, D.F. and Visconti, P.E. (2003) Phosphoproteome analysis of capacitated human sperm. Evidence of tyrosine phosphorylation of a kinase-anchoring protein 3 and valosin-containing protein/p97 during capacitation. *J Biol Chem*, **278**, 11579-11589.

Ficarro, S.B., Salomon, A.R., Brill, L.M., Mason, D.E., Stettler-Gill, M., Brock, A. and Peters, E.C. (2005) Automated immobilized metal affinity chromatography/nano-liquid chromatography/electrospray ionization mass spectrometry platform for profiling protein phosphorylation sites. *Rapid Commun Mass Spectrom*, **19**, 57-71.

Forman-Kay, J.D. and Pawson, T. (1999) Diversity in protein recognition by PTB domains. *Curr Opin Struct Biol*, **9**, 690-695.

Fraley, C. and Raftery, A.E. (1998) How Many Clusters? Which Clustering Method? Answers Via Model-Based Cluster Analysis. *The Computer Journal*, **41**, 578-588.

Hinsby, A.M., Olsen, J.V., Bennett, K.L. and Mann, M. (2003) Signaling initiated by overexpression of the fibroblast growth factor receptor-1 investigated by mass spectrometry. *Mol Cell Proteomics*, **2**, 29-36.

Hinsby, A.M., Olsen, J.V. and Mann, M. (2004) Tyrosine phosphoproteomics of fibroblast growth factor signaling: a role for insulin receptor substrate-4. *J Biol Chem*, **279**, 46438-46447.

Downloaded from www.mcponline.org by on January 9, 2009

Huyer, G. and Ramachandran, C. (1998) The specificity of the N-terminal SH2 domain of SHP-2 is modified by a single point mutation. *Biochemistry*, **37**, 2741-2747.

Ishihama, Y., Rappsilber, J., Andersen, J.S. and Mann, M. (2002) Microcolumns with self-assembled particle frits for proteomics. *J Chromatogr A*, **979**, 233-239.

Jain, A.K., Murth, M.N. and Flynn, P.J. (1999) Data clustering: a review. *ACM Computing Surveys*, **31**, 264-323.

Jantz, D. and Berg, J.M. (2004) Reduction in DNA-binding affinity of Cys2His2 zinc finger proteins by linker phosphorylation. *Proc Natl Acad Sci U S A*, **101**, 7589-7593.

Johnson, L.N., Kotz, S. and Kemp, A.W. (1992) Univariate Discrete Distributions. **Second Edition**.

Jonassen, I., Collins, J.F. and Higgins, D.G. (1995) Finding flexible patterns in unaligned protein sequences. *Protein Sci*, **4**, 1587-1595.

Kaufman, L. and Rousseeuw, P.J. (1990) Finding Groups in Data: an Introduction to Cluster Analysis.

Liu, B.A., Jablonowski, K., Raina, M., Arce, M., Pawson, T. and Nash, P.D. (2006) The human and mouse complement of SH2 domain proteins-establishing the boundaries of phosphotyrosine signaling. *Mol Cell*, **22**, 851-868.

Lund, O., Nielsen, M., Kesmir, C., Petersen, A.G., Lundegaard, C., Worning, P., Sylvester-Hvid, C., Lamberth, K., Roder, G., Justesen, S., Buus, S. and Brunak, S. (2004) Definition of supertypes for HLA molecules using clustering of specificity matrices. *Immunogenetics*, **55**, 797-810.

Marion, E., Kaisaki, P.J., Pouillon, V., Gueydan, C., Levy, J.C., Bodson, A., Krzentowski, G., Daubresse, J.C., Mockel, J., Behrends, J., Servais, G., Szpirer, C., Kruys, V., Gauguier, D. and Schurmans, S. (2002) The gene INPPL1, encoding the lipid phosphatase SHIP2, is a candidate for type 2 diabetes in rat and man. *Diabetes*, **51**, 2012-2017.

Muraille, E., Bruhns, P., Pesesse, X., Daeron, M. and Erneux, C. (2000) The SH2 domain containing inositol 5-phosphatase SHIP2 associates to the immunoreceptor tyrosine-based inhibition motif of Fc gammaRIIB in B cells under negative signaling. *Immunol Lett*, **72**, 7-15.

Nevill-Manning, C.G., Wu, T.D. and Brutlag, D.L. (1998) Highly specific protein sequence motifs for genome analysis. *Proc Natl Acad Sci U S A*, **95**, 5865-5871.

Downloaded from www.mcponline.org by on January 9, 2009

ASBMB Molecular & Cellular Proteomics

Nielsen, M., Lundegaard, C., Worning, P., Hvid, C.S., Lamberth, K., Buus, S., Brunak, S. and Lund, O. (2004) Improved prediction of MHC class I and class II epitopes using a novel Gibbs sampling approach. *Bioinformatics*, **20**, 1388-1397.

Olsen, J.V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P. and Mann, M. (2006) Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. *Cell*, **127**, 635-648.

Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A. and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics*, **1**, 376-386.

Ong, S.E., Kratchmarova, I. and Mann, M. (2003) Properties of 13C-substituted arginine in stable isotope labeling by amino acids in cell culture (SILAC). *J Proteome Res*, **2**, 173-181.

Pawson, T. (2004) Specificity in signal transduction: from phosphotyrosine-SH2 domain interactions to complex cellular systems. *Cell*, **116**, 191-203.

Peri, S., Navarro, J.D., Amanchy, R., Kristiansen, T.Z., Jonnalagadda, C.K., Surendranath, V., Niranjan, V., Muthusamy, B., Gandhi, T.K., Gronborg, M., Ibarrola, N., Deshpande, N., Shanker, K., Shivashankar, H.N., Rashmi, B.P., Ramya, M.A., Zhao, Z., Chandrika, K.N., Padma, N., Harsha, H.C., Yatish, A.J., Kavitha, M.P., Menezes, M., Choudhury, D.R., Suresh, S., Ghosh, N., Saravana, R., Chandran, S., Krishna, S., Joy, M., Anand, S.K., Madavan, V., Joseph, A., Wong, G.W., Schiemann, W.P., Constantinescu, S.N., Huang, L., Khosravi-Far, R., Steen, H., Tewari, M., Ghaffari, S., Blobe, G.C., Dang, C.V., Garcia, J.G., Pevsner, J., Jensen, O.N., Roepstorff, P., Deshpande, K.S., Chinnaiyan, A.M., Hamosh, A., Chakravarti, A. and Pandey, A. (2003) Development of human protein reference database as an initial platform for approaching systems biology in humans. *Genome Res*, **13**, 2363-2371.

Perkins, D.N., Pappin, D.J., Creasy, D.M. and Cottrell, J.S. (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*, **20**, 3551-3567.

Poy, F., Yaffe, M.B., Sayos, J., Saxena, K., Morra, M., Sumegi, J., Cantley, L.C., Terhorst, C. and Eck, M.J. (1999) Crystal structures of the XLP protein SAP reveal a class of SH2 domains with extended, phosphotyrosine-independent sequence recognition. *Mol Cell*, **4**, 555-561.

Rappsilber, J., Ishihama, Y. and Mann, M. (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal Chem*, **75**, 663-670.

Downloaded from www.mcponline.org by on January 9, 2009

ASBMB Molecular & Cellular Proteomics

Rigoutsos, I. and Floratos, A. (1998) Combinatorial pattern discovery in biological sequences: The TEIRESIAS algorithm. *Bioinformatics*, **14**, 55-67.

Rush, J., Moritz, A., Lee, K.A., Guo, A., Goss, V.L., Spek, E.J., Zhang, H., Zha, X.M., Polakiewicz, R.D. and Comb, M.J. (2005) Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nat Biotechnol*, **23**, 94-101.

Salomon, A.R., Ficarro, S.B., Brill, L.M., Brinker, A., Phung, Q.T., Ericson, C., Sauer, K., Brock, A., Horn, D.M., Schultz, P.G. and Peters, E.C. (2003) Profiling of tyrosine phosphorylation pathways in human cells using mass spectrometry. *Proc Natl Acad Sci U S A*, **100**, 443-448.

Saydam, N., Adams, T.K., Steiner, F., Schaffner, W. and Freedman, J.H. (2002) Regulation of metallothionein transcription by the metal-responsive transcription factor MTF-1: identification of signal transduction cascades that control metal-inducible transcription. *J Biol Chem*, **277**, 20438-20445.

Schneider, T.D. and Stephens, R.M. (1990) Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res*, **18**, 6097-6100.

Schulze, W.X., Deng, L. and Mann, M. (2005) Phosphotyrosine interactome of the ErbB-receptor kinase family. *Mol Syst Biol*, **1**, 2005 0008.

Schulze, W.X. and Mann, M. (2004) A novel proteomic screen for peptide-protein interactions. *J Biol Chem*, **279**, 10756-10764.

Schwartz, D. and Gygi, S.P. (2005) An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. *Nat Biotechnol*, **23**, 1391-1398.

Sleeman, M.W., Wortley, K.E., Lai, K.M., Gowen, L.C., Kintner, J., Kline, W.O., Garcia, K., Stitt, T.N., Yancopoulos, G.D., Wiegand, S.J. and Glass, D.J. (2005) Absence of the lipid phosphatase SHIP2 confers resistance to dietary obesity. *Nat Med*, **11**, 199-205.

Songyang, Z. and Cantley, L.C. (1995) Recognition and specificity in protein tyrosine kinase-mediated signalling. *Trends Biochem Sci*, **20**, 470-475.

Songyang, Z., Margolis, B., Chaudhuri, M., Shoelson, S.E. and Cantley, L.C. (1995) The phosphotyrosine interaction domain of SHC recognizes tyrosine-phosphorylated NPXY motif. *J Biol Chem*, **270**, 14863-14866.

Songyang, Z., Shoelson, S.E., McGlade, J., Olivier, P., Pawson, T., Bustelo, X.R., Barbacid, M., Sabe, H., Hanafusa, H., Yi, T. and et al. (1994) Specific motifs recognized

Downloaded from www.mcponline.org by on January 9, 2009

ASBMB Molecular & Cellular Proteomics

by the SH2 domains of Csk, 3BP2, fps/fes, GRB-2, HCP, SHC, Syk, and Vav. *Mol Cell Biol*, **14**, 2777-2785.

Vely, F. and Vivier, E. (1997) Conservation of structural features reveals the existence of a large family of inhibitory cell surface receptors and noninhibitory/activatory counterparts. *J Immunol*, **159**, 2075-2077.

Wu, X., Suetsugu, S., Cooper, L.A., Takenawa, T. and Guan, J.L. (2004) Focal adhesion kinase regulation of N-WASP subcellular localization and function. *J Biol Chem*, **279**, 9565-9576.

Yaffe, M.B. (2002) Phosphotyrosine-binding domains in signal transduction. *Nat Rev Mol Cell Biol*, **3**, 177-186.

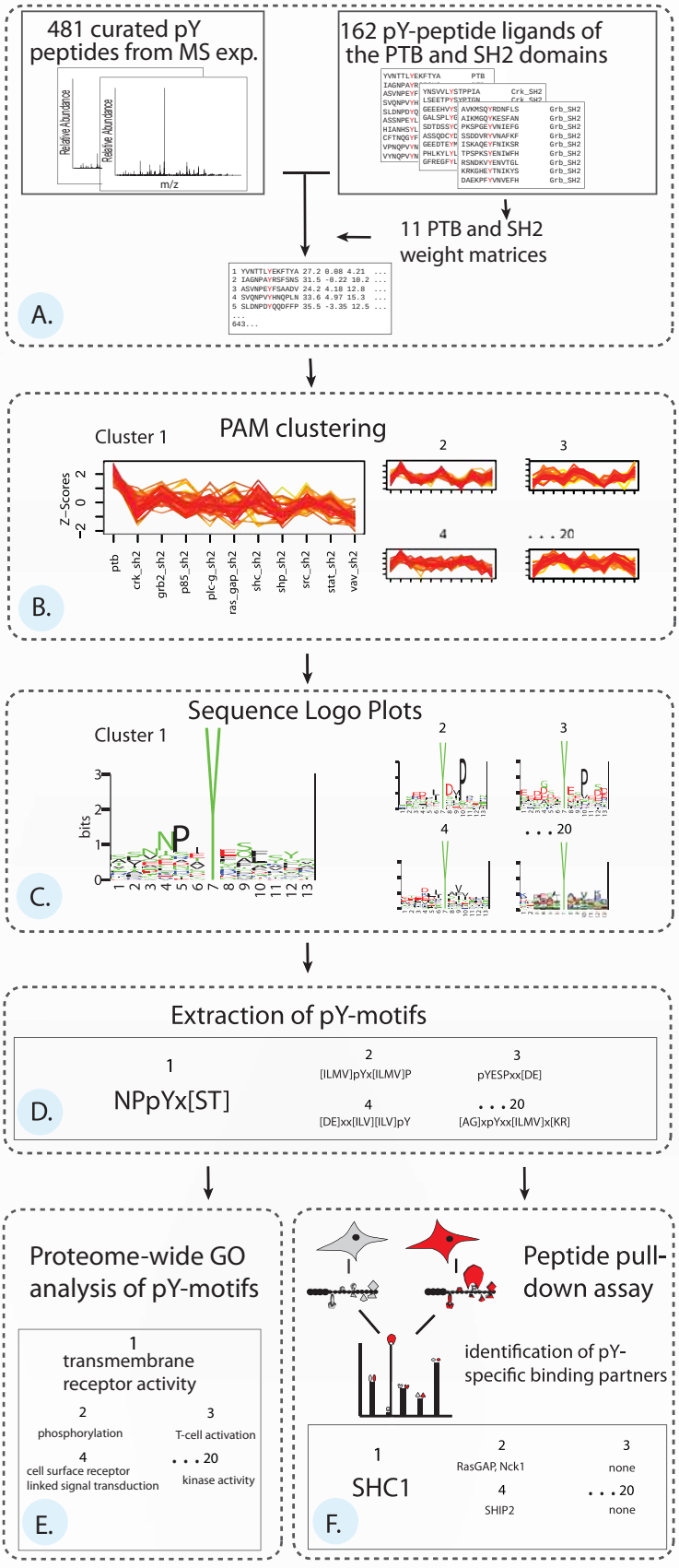
Zheng, H., Hu, P., Quinn, D.F. and Wang, Y.K. (2005) Phosphotyrosine proteomic study of interferon alpha signaling pathway using a combination of immunoprecipitation and immobilized metal affinity chromatography. *Mol Cell Proteomics*, **4**, 721-730.

Downloaded from www.mcponline.org by on January 9, 2009

LUS-TER	SI-ZE	POSITIVE CONTROLS	EXTRACTED MOTIFS	MATCHED MOTIFS, EXPECTED PARTNER	PUBMED ID	PEPTIDES SYNTHESIZED	IDENTIFIED PARTNERS
1	41	PTB 10 of 10	NPxpYx[ST] 7 of 7	SHC PTB (NPxpY)	7542744 7541030	KEVCDGWSLPNPEpYYTLRYA ELMO2_MOUSE, 48	SHC
2	36	Crk SH2 8 of 10	[ILMV]pYx[ILMV]P 8 of 14	Crk/RasGAP SH2 (pYxxP)	9233798 11607838	KPSTDPLpYDTPDTRG RIN1_MOUSE, 35	RasGAP Nck1
3	29	Vav SH2 4 of 4	pYESPxx[DE] 5 of 5	Vav SH2 (pYESP)	9151714	TETKTITpYESPQIDG E41L2_MOUSE, 889	none
4	28		[DE]xxx[ILV][ILV]pY 4 of 6	new motif		RETSKVIpYDFIEKTG WASL_MOUSE, 253	SHIP2
5	20		[DE][DE]xxxpYxN 4 of 6	Grb2 SH2 (pYxN)	11994738	VYDEDSPpYQNIKILH SPSY_MOUSE, 147	Grb2 RasGAP
6	41	Grb2 SH2 14 of 31	pYxN[ILMV]xxL 5 of 7	Grb2 SH2 (pYxN)	11994738	ELFDDPSpYVNIQNLD SHC1_MOUSE, 423	Grb2
7	21	Grb2 SH2 6 of 31	[DE]pYxN[ILMV] 4 of 11	Grb2 SH2 (pYxN)	11994738	QPASVTDpYQNVSFSN ITSN2_HUMAN, 858	Grb2
8	45		pY[ILMV]xMxP 4 of 10	p85 PI3K SH2 (pYxxM)	7511210 11994738	PQRVDPNGpYMMMSPS IRS1_MOUSE, 658	p85-PI3Kα p85-PI3Kβ
9	40		pY[DE]x[ILMV]x[ILMV] 5 of 22	FpsFes SH2 (pY[ED]x[IV])	7511210	AGKQKLQpYEGIFIKD SF3A1_MOUSE, 757	none
10	32		[DE]xxpY[DE]x[ILMV] 7 of 27	FpsFes SH2 (pY[ED]x[IV])	7511210	DGGSDQNpYDIVTIGA INP4A_HUMAN, 355	none
11	35	PI3K SH2 16 of 24	pY[ILMV]PMxP 6 of 7	p85 PI3K SH2 (pYxxM)	7511210 11994738	NLHTDDGpYMPMSPGV IRS1_MOUSE, 608	p85-PI3Kα
12	23		DpY[ILMV]x[ILMV] 7 of 18	SHP2 SH2 (pY[IVL]x[IVL])	7680959	DLINRMDpYVEINIDH VIGLN_MOUSE, 437	SHP2
13	28	SHP2 SH2 7 of 12	[ILMV]xpY[ILMV]x[ILMV]D 6 of 7	SHP2 SH2 (pY[IVL]x[IVL])	7680959	DIKEKLCpYVALDFEQ ACTB_MOUSE, 218	SHP2
14	32	PLCg SH2 5 of 16	[ILMV]pYxx[ILMV][ILMV] 5 of 11	general/SHP2 SH2 (pY[IVL]x[IVL])	7511210 7680959	GKSKQPLpYSSIVTVE O88185_MOUSE, 948	SHP2 SHIP2
15	29	RasGAP SH2 7 of 8	[AG][ILMV]pYxxP 6 of 10	Crk/RasGAP SH2 (pYxxP)	9233798 11607838	GVVDSGVpYAVPPPAE BCAR1_HUMAN, 410	RasGAP
16	37	Shc SH2 9 of 13	PxEpYxxxxx[ILMV] 3 of 3	new motif		TTEAPGEYFFSDGVR IMDH1_MOUSE, 400	none
17	25	Src SH2 7 of 14	pY[DE]x[ILMV]H 4 of 6	FpsFes SH2 (pY[ED]x[IV])	7511210	ELTAEFLpYDEVHPKQ TWF2_MOUSE, 309	RasGAP
18	32	Stat SH2 15 of 19	pY[ILMV]PQ 4 of 4	STAT SH2 (pYxxQ)	14966128	SGENFVPpYMPQFQTC LEPR_MOUSE, 1138	none
19	33		H[ST]GxKPpYxCxxCG 10 of 10	new motif		RIHTGEKPpYECVQCGK ZNF24_MOUSE, 335	none
20	36		[AG]xpYxx[ILMV]x[KR] 8 of 15	new motif		KKNRIAIpYELLFKEG RS10_MOUSE, 12	none

Downloaded from www.mcponline.org by on January 9, 2009

Figure 1



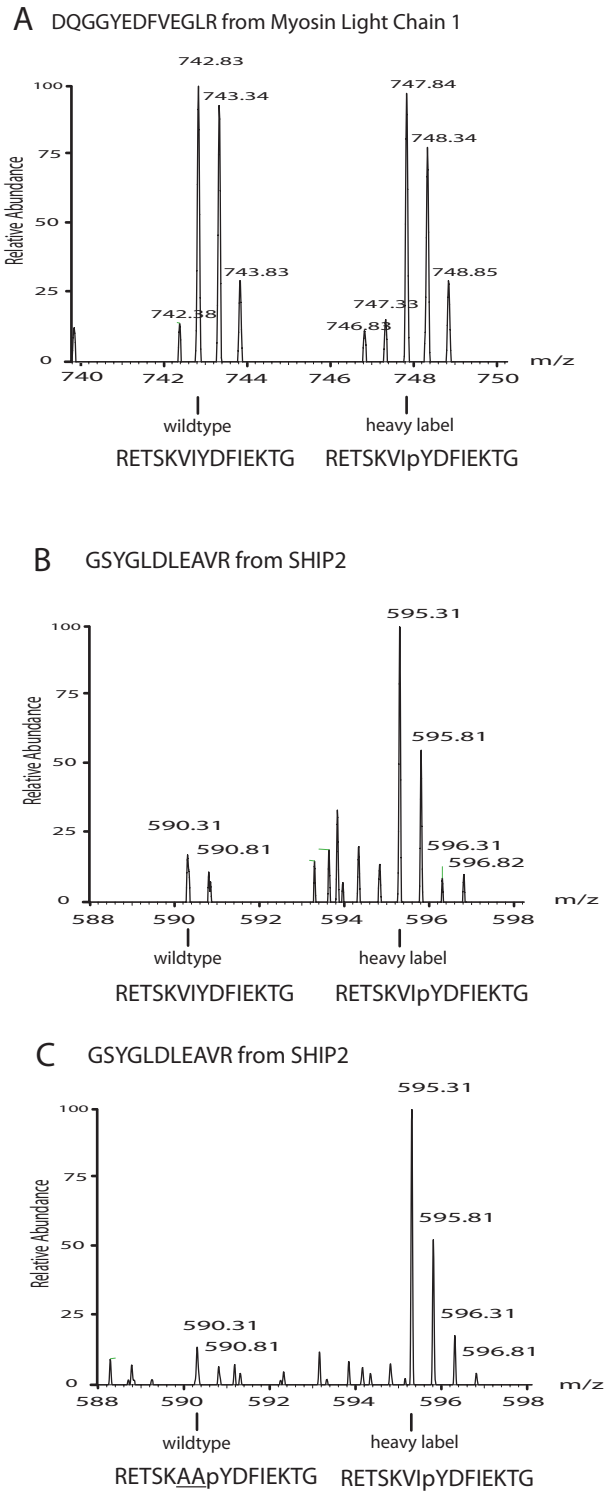
Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

Figure 2



Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

Figure 3



Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

This supplementary table 1 contains the results from all of the 32 peptide pull-down experiments that were performed in this report. Column A states the cluster number that the peptides were extracted from. The date of the performed experiment can be seen in column B. The peptide-pairs used in the experiments are found in column C and D with the phosphotyrosine is indicated in red. The SwissProt entry name of the parent protein of the bait peptides and the residue number of the phosphorylation site are stated in column E and F, respectively, followed the instrument used for analysis. The Mascot hit number of the significant binder is in column H ('none' is stated if none such was retrieved). The following columns are: protein name, gene name, description, IPI and Uniprot accession numbers. The log average of the intensity ratios of labeled to unlabeled tryptic peptides of the significant binder is in column O. The log mean of the intensity ratios of labeled to unlabeled form of all proteins identified in the experiment and the log standard deviation are stated in column P and Q respectively. The binder's number of standard deviations over the log mean of all the proteins is stated in column N. Similarly, the non-log values are in column R-U. The Mascot score of the significant binder is found in column V followed by the number of identified (W) and quantified (X) peptides. The $p<0.05$-value, e.g. the score at which a peptide´s likelihood to be wrongly identified is 5%, was between 28 and 32 in all experiments. Finally, the number of MS3 fragmented peptides is found in column Y giving the corrected mascot score in the last column.

Cluster	Date	Bait peptide1	Bait peptide2	SwissProt	pY site	Instrument	Mascot hit	Protein	Gene	Description	IPI	Uniprot	No. of stdev (log)	Ave i/w (log)	Mean (log)	Stdev (log)	Ave i/w	Stdev i/w		Stdev	Mascot score	No.of peps	No.of quant.peps	No.of MS3 peps	MS3 mascot score
1	20060607	IEVCDGWSLPNPEYYTLRY	SLPNPE	ELMO2_MOUSE	48	LTQ-FT	39	SHC	Shc1	SHC trans	IPI0012P98083		6.84	1.45	0.03	0.21	27.89	22.05	1.35	2.4	832	15	9	11	1654.38
1	20060607	KEVCDGWSLPNPEYYTLRY	SLPNPE	ELMO2_MOUSE	48	LTQ-FT	35	SHC	Shc1	SHC trans	IPI0012P98083		-5.34	-1.19	-0.08	0.21	0.07	0.05	0.93	0.41	1048	19	9	15	1774.39
2	20060324	KPSTDPLYDTPDTRG	PLYDTI	RIN1_MOUSE	35	LTQ-FT	47	RasGAP	Rasa1	RAS p21 p	IPI0013Q3V090		4.85	1.18	0	0.24	15.03	2.83	1.33	2.05	576	13	5	11	1539.54
2	20060324	KPSTDPLYDTPDTRG	PLYDTI	RIN1_MOUSE	35	LTQ-FT	37	RasGAP	Rasa1	RAS p21 p	IPI0013Q3V090		-3.71	-1.04	-0.23	0.22	0.09	0.07	0.65	0.3	729	16	4	12	1828.56
2	20060324	KPSTDPLYDTPDTRG	PLYDTI	RIN1_MOUSE	35	LTQ-FT	123	Nck1	Nck1	SH2/SH3 a	IPI0030Q8BQ28		3.64	0.89	0	0.24	7.67	2.32	1.33	2.05	169	4	3	3	371.7
2	20060324	KPSTDPLYDTPDTRG	PLYDTI	RIN1_MOUSE	35	LTQ-FT	131	Nck1	Nck1	SH2/SH3 a	IPI0030Q8BQ28		-4.48	-1.21	-0.23	0.22	0.06	0.05	0.85		161	4	2	4	403.93
3	20060324	TETKTITYESPQIDG	ITYESI	E41L2_MOUSE	889	LTQ-FT	none																		
3	20060324	TETKTITYESPQIDG	ITYESI	E41L2_MOUSE	889	LTQ-FT	none																		
4	20060324	RETSKVIYDFIEKTG	VIYDF	WASL_MOUSE	253	LTQ-FT	59	SHIP2	011 (Shi	Inositol	IPI0031Q6P549		6.52	0.98	0.04	0.14	9.51	1.55	1.2	0.86	449	12	3	6	724.68
4	20060324	RETSKVIYDFIEKTG	VIYDF	WASL_MOUSE	253	LTQ-FT	68	SHIP2	011 (Shi	Inositol	IPI0031Q6P549		-3.48	-0.86	-0.19	0.19	0.14	0.03	0.73	0.47	462	11	5	9	874.4
4	20061102	RETSKVIYDFIEKTG	VIYDF	WASL_MOUSE	253	Q-Orbitr	911	Nck1	Nck1	SH2/SH3 a	IPI0030O55033		4.61	0.7	-0.21	0.2	5.03	0	0.69	0.49	31	1	1	0	31
4	20061102	RETSKVIYDFIEKTG	VIYDF	WASL_MOUSE	253	Q-Orbitr	389	Nck1	Nck1	SH2/SH3 a	IPI0030055033		-3.5	-0.68	0.07	0.21	0.21	0.05	1.26	0.36	121	4	3	0	121
4	20061102	RETSKVIYDFIEKTG	VIYDF	WASL_MOUSE	253	Q-Orbitr	361	SHIP2	011 (Shi	Inositol	IPI0031008611		5.33	0.84	-0.21	0.2	6.97	0	0.69	0.49	92	2	1	0	92
4	20061102	RETSKVIYDFIEKTG	VIYDF	WASL_MOUSE	253	Q-Orbitr	398	SHIP2	011 (Shi	Inositol	IPI0031008611		-5.86	-1.18	0.07	0.21	0.07	0.01	1.26	0.36	67	2	2	0	67
4	20061102	GTKEIFDYIVKSTER	FDYIV	Reverse WASL_MOUS	253	Q-Orbitr	none																		
4	20061102	GTKEIFDYIVKSTER	FDYIV	Reverse WASL_MOUS	253	Q-Orbitr	none																		
4	20060607	RETSKAAYDFIEKTG	VIYDF	WASL_MOUSE_AA	253	LTQ-FT	73	SHIP2	011 (Shi	Inositol	IPI0031Q6P549		3.94	1.01	0.03	0.25	10.22	3.04	1.25	0.94	453	10	6	7	718.8
4	20060607	RETSKVIYDFIEKTG	AAYDF	WASL_MOUSE_AA	253	LTQ-FT	113	SHIP2	011 (Shi	Inositol	IPI0031Q6P549		-4.26	-1.19	-0.03	0.27	0.06	0.02	1.21	1.51	307	8	4	6	523.82
4	20060607	RATSKAAYDFIEKTG	VIYDF	WASL_MOUSE_AAA	253	LTQ-FT	210	SHIP2	011 (Shi	Inositol	IPI0031Q6P549		6.9	1.41	0.03	0.2	25.51	0	1.32	2.17	46	1	1	0	46
4	20060607	RETSKVIYDFIEKTG	AAYDF	WASL_MOUSE_AAA	253	LTQ-FT	122	SHIP2	011 (Shi	Inositol	IPI0031Q6P549		-4.96	-0.96	-0.05	0.18	0.11	0.06	0.98	0.56	263	6	3	5	431.58
4	20061102	SEDRDLLYSIDAYRS	LLYSII	ANLN_MOUSE	666	Q-Orbitr	none																		
4	20061102	SEDRDLLYSIDAYRS	LLYSII	ANLN_MOUSE	666	Q-Orbitr	none																		
4	20061102	QDHNHLLYSTIPRMQ	LLYST	CTND1_MOUSE	96	Q-Orbitr	211	SHIP2	011 (Shi	Inositol	IPI0031008611		7.76	0.93	-0.18	0.14	8.6	1.48	0.71	0.47	290	7	5	0	290
4	20061102	QDHNHLLYSTIPRMQ	LLYST	CTND1_MOUSE	96	Q-Orbitr	235	SHIP2	011 (Shi	Inositol	IPI0031008611		-8.61	-1.13	0	0.13	0.07	0.03	1.06	0.42	314	8	6	0	314
4	20061102	AETSKLIYDFIEDQG	LIYDF	WASP_MOUSE	293	Q-Orbitr	461	SHIP2	011 (Shi	Inositol	IPI0031008611		5.55	0.46	-0.25	0.13	2.87	0	0.6	0.25	34	1	1	0	34
4	20061102	AETSKLIYDFIEDQG	LIYDF	WASP_MOUSE	293	Q-Orbitr	386	SHIP2	011 (Shi	Inositol	IPI0031008611		-5.48	-1.1	0.01	0.2	0.08	0.02	1.11	0.58	92	3	3	0	92
4	20061102	AETSKLIYDFIEDQG	LIYDF	WASP_MOUSE	293	Q-Orbitr	470	Nck1	Nck1	SH2/SH3 a	IPI0030055033		4.92	0.38	-0.25	0.13	2.39	0	0.6	0.25	55	2	1	0	55

Downloaded from www.mcponline.org on January 9, 2009

4	20061102_	AETSKLIYDFIEDQG	KLIYDF]	WASP_MOUSE	293	Q-Orbitr	511	Nck1	Nck1	SH2/SH3 a	IPI0030055033	-3.73	-0.74	0.01	0.2	0.18	0.01	1.11	0.58	51	2	2	0	51			
5	20060324!	VYDEDSPYQNIKILH	)SPYQN]	SPSY_MOUSE	147	LTQ-FT	61	Grb2	Grb2	Splice Is	IPI0011Q60631	4.61	1.22	0.08	0.25	16.41	4.71	1.57	2.43	463	12	8	8	971.22			
5	20060324!	VYDEDSPYQNIKILH	)SPYQN]	SPSY_MOUSE	147	LTQ-FT	46	Grb2	Grb2	Splice Is	IPI0011Q60631	-5.71	-1.37	-0.11	0.22	0.04	0.02	0.85	0.35	521	13	9	9	1030.15			
5	20060324!	VYDEDSPYQNIKILH	)SPYQN]	SPSY_MOUSE	147	LTQ-FT	64	RasGAP	Rasa1	RAS p21 p	IPI0013Q3V090	4.95	1.3	0.08	0.25	19.9	3.45	1.57	2.43	405	11	4	7	878.68			
5	20060324!	VYDEDSPYQNIKILH	)SPYQN]	SPSY_MOUSE	147	LTQ-FT	48	RasGAP	Rasa1	RAS p21 p	IPI0013Q3V090	-3.41	-0.86	-0.11	0.22	0.14	0.1	0.85	0.35	555	13	6	9	1315.26			
6	20060405!	ELFDDPSYVNIQNLD	)PSYVN]	SHC1_MOUSE	423	LTQ-FT	46	Grb2	Grb2	Splice Is	IPI0011Q60631	5.53	1.32	-0.16	0.27	20.68	6.6	0.95	1.98	610	15	8	12	1339.32			
6	20060405!	ELFDDPSYVNIQNLD	)PSYVN]	SHC1_MOUSE	423	LTQ-FT	50	Grb2	Grb2	Splice Is	IPI0011Q60631	-6.89	-1.75	0.2	0.28	0.02	0.01	1.85	1.04	480	12	6	9	1008.42			
7	20060324!	QPASVTDYQNVSFSN	/TDYQN\	ITSN2_HUMAN	858	LTQ-FT	172	Grb2	Grb2	Splice Is	IPI0011Q60631	3.82	1.1	-0.36	0.38	12.46	3.15	0.64	1.04	263	9	5	4	504.75			
7	20060324!	QPASVTDYQNVSFSN	/TDYQN\	ITSN2_HUMAN	858	LTQ-FT	73	Grb2	Grb2	Splice Is	IPI0011Q60631	-3.29	-1.3	0.07	0.42	0.05	0.01	1.91	2.16	438	11	5	5	678.06			
8	20051002!	PQRVDPNGYMMMSPS	)PNGYM]	IRS1_MOUSE	658	LTQ-FT	32	p85-PI3Kα	(p85 α	Phosphati	IPI0026P26450	4.81	1.34	0.05	0.27	21.78	9.36	1.61	2.99	369	7	4	4	752.85			
8	20051002!	PQRVDPNGYMMMSPS	)PNGYM]	IRS1_MOUSE	658	LTQ-FT	40	p85-PI3Kα	(p85 α	Phosphati	IPI0026P26450	4.28	1.11	-0.02	0.26	12.88	2.08	1.36	2.12	254	5	3	4	696.81			
8	20051002!	PQRVDPNGYMMMSPS	)PNGYM]	IRS1_MOUSE	658	LTQ-FT	70	p85-PI3Kβ2	(p85	Phosphati	IPI0011008908	4.56	1.27	0.05	0.27	18.65	0	1.61	2.99	35	1	1	0	35			
8	20051002!	PQRVDPNGYMMMSPS	)PNGYM]	IRS1_MOUSE	658	LTQ-FT	56	p85-PI3Kβ2	(p85	Phosphati	IPI0011008908	3.82	0.99	-0.02	0.26	9.7	0	1.36	2.12	31	1	1	0	31			
9	20060324!	AGKQKLQYEGIFIKD	KLQYEG]	SF3A1_MOUSE	757	LTQ-FT	none																				
9	20060324!	AGKQKLQYEGIFIKD	KLQYEG]	SF3A1_MOUSE	757	LTQ-FT	none																				
10	20060324!	DGGSDQNYDIVTIGA	)QNYDI\	INP4A_HUMAN	355	LTQ-FT	none																				
10	20060324!	DGGSDQNYDIVTIGA	)QNYDI\	INP4A_HUMAN	355	LTQ-FT	none																				
11	20051002!	NLHTDDGYMPMSPGV	)DGYMPM	IRS1_MOUSE	608	LTQ-FT	51	p85-PI3Kα	(p85 α	Phosphati	IPI0026P26450	3.86	1.05	-0.02	0.28	11.14	8.64	1.18	1.27	255	5	3	0	255			
11	20051002!	NLHTDDGYMPMSPGV	)DGYMPM	IRS1_MOUSE	608	LTQ-FT	54	p85-PI3Kα	(p85 α	Phosphati	IPI0026P26450	4.05	0.98	-0.09	0.27	9.59	3.24	0.99	1.03	278	6	4	4	561.48			
12	20060324!	DLINRMDYVEINIDH	RMDYVE]	VIGLN_MOUSE	437	LTQ-FT	35	SHP2	n11 (Sh	Splice Is	IPI0011P35235	3.91	1.35	-0.29	0.42	22.14	5.74	0.93	2.03	571	12	5	7	964.73			
12	20060324!	DLINRMDYVEINIDH	RMDYVE]	VIGLN_MOUSE	437	LTQ-FT	78	SHP2	n11 (Sh	Splice Is	IPI0011P35235	-4.29	-1.4	0.43	0.43	0.04	0.01	4.17	4.51	452	13	2	9	767.81			
13	20060405!	DIKEKLCYVALDFEQ	KLCYVA]	ACTB_MOUSE	218	LTQ-FT	24	SHP2	n11 (Sh	Splice Is	IPI0011P35235	3.99	1.18	-0.02	0.3	15.17	5.32	1.3	1.75	1045	25	13	16	1757.82			
13	20060405!	DIKEKLCYVALDFEQ	KLCYVA]	ACTB_MOUSE	218	LTQ-FT	23	SHP2	n11 (Sh	Splice Is	IPI0011P35235	-5.74	-1.36	0.06	0.25	0.04	0.02	1.3	0.66	1116	27	11	19	2222.67			
14	20060405!	GKSKQPLYSSIVTVE	)PLYSS]	O88185_MOUSE	948	LTQ-FT	48	SHIP2	)l1 (Shi	Inositol	IPI0031Q6P549	4.1	0.96	-0.03	0.24	9.1	3.57	1.58	1.45	512	12	3	6	682.61			
14	20060405!	GKSKQPLYSSIVTVE	)PLYSS]	O88185_MOUSE	948	LTQ-FT	50	SHIP2	)l1 (Shi	Inositol	IPI0031Q6P549	-3.02	-0.6	0.07	0.22	0.25	0	1.3	0.47	403	8	1	8	751.58			
14	20060405!	GKSKQPLYSSIVTVE	)PLYSS]	O88185_MOUSE	948	LTQ-FT	41	SHP2	n11 (Sh	Splice Is	IPI0011P35235	4.56	1.07	-0.03	0.24	11.79	0	1.18	1.45	581	14	2	9	870.88			
14	20060405!	GKSKQPLYSSIVTVE	)PLYSS]	O88185_MOUSE	948	LTQ-FT	56	SHP2	n11 (Sh	Splice Is	IPI0011P35235	-5.87	-1.22	0.07	0.22	0.06	0.02	1.3	0.47	483	15	2	10	1109.35			
15	20060324!	GVVDSGVYAVPPPAE	GVYAV]	BCAR1_HUMAN	410	LTQ-FT	72	RasGAP	Rasa1	RAS p21 p	IPI0013Q3V090	5.65	1.28	0.14	0.2	18.88	0	1.61	1.75	320	8	2	7	833.62			
15	20060324!	GVVDSGVYAVPPPAE	GVYAV]	BCAR1_HUMAN	410	LTQ-FT	77	RasGAP	Rasa1	RAS p21 p	IPI0013Q3V090	-3.85	-1.19	-0.24	0.25	0.06	0	0.67	0.38	225	5	1	5	790.96			
16	20060324!	TTEAPGEYFFSDGVR	PGEYFF]	IMDH1_MOUSE	400	LTQ-FT	none																				
16	20060324!	TTEAPGEYFFSDGVR	PGEYFF]	IMDH1_MOUSE	400	LTQ-FT	none																				
17	20060405!	ELTAEFLYDEVHPKQ	EFLYDE\	TWF2_MOUSE	309	LTQ-FT	34	RasGAP	Rasa1	RAS p21 p	IPI0013Q3V090	6.87	1.44	0.03	0.2	27.23	13.96	1.42	2.94	671	16	9	13	1769.66			
17	20060405!	ELTAEFLYDEVHPKQ	EFLYDE\	TWF2_MOUSE	309	LTQ-FT	31	RasGAP	Rasa1	RAS p21 p	IPI0013Q3V090	-4.98	-0.9	0.01	0.18	0.13	0	1.08	0.28	489	12	1	8	989.63			
18	20060324!	SGENFVPYMPQFQTC	VPYMP(	LEPR_MOUSE	###	LTQ-FT	none																				
18	20060324!	SGENFVPYMPQFQTC	VPYMP(	LEPR_MOUSE	###	LTQ-FT	none																				
19	20060105!	RIHTGEKPYECVQCGK	EKPYEC	ZNF24_MOUSE	335	LTQ-FT	none																				
19	20060105!	RIHTGEKPYECVQCGK	EKPYEC	ZNF24_MOUSE	335	LTQ-FT	none																				
20	20060324!	KKNRIAIYELLFKEG	AIYELI	RS10_MOUSE	12	LTQ-FT	none																				
20	20060324!	KKNRIAIYELLFKEG	AIYELI	RS10_MOUSE	12	LTQ-FT	none																				

Downloaded from www.mcponline.org by on January 9, 2009

Molecular & Cellular Proteomics

This supplementary table 2 lists overrepresented GO terms in proteins matching the extracted motifs using a hypergeometrical test with the whole proteome as background. In the first two columns the cluster number and the extracted motifs can be found. In column C the overrepresented GO ID can be seen followed by the p-value. Column E states the number of proteins with the GO ID that also matches the extracted motif. Column D is the number of proteins in the proteome with the GO ID. Column G is the number of proteins in the proteome that matches the extracted motif followed by the total number of protein in the proteome. The GO name can be seen in Column I, and finally the assessed consensus GO term can be found in column J.

Cluster	Extracted motif	GO ID	p-value	Motif matches with GO ID	GO ID matches in proteome	Motif matches in proteome	Proteome size	GO Name	Consensus GO term of top 10
1	NP.Y.[ST]	GO:0044464	4.79E-71	6532	21607	9129	33758	cell part	
1	NP.Y.[ST]	GO:0005623	4.79E-71	6532	21607	9129	33758	cell	
1	NP.Y.[ST]	GO:0004888	7.59E-62	1326	3355	9129	33758	transmembrane receptor activity	transmembrane receptor activity
1	NP.Y.[ST]	GO:0044425	1.92E-61	2672	7749	9129	33758	membrane part	
1	NP.Y.[ST]	GO:0016020	1.28E-60	3380	10201	9129	33758	membrane	
1	NP.Y.[ST]	GO:0016021	3.03E-59	2464	7083	9129	33758	integral to membrane	
1	NP.Y.[ST]	GO:0031224	9.65E-59	2466	7098	9129	33758	intrinsic to membrane	
1	NP.Y.[ST]	GO:0004872	3.46E-52	1596	4329	9129	33758	receptor activity	
1	NP.Y.[ST]	GO:0043169	6.60E-50	2090	5985	9129	33758	cation binding	
1	NP.Y.[ST]	GO:0005575	1.28E-49	6877	23401	9129	33758	cellular_component	
1	NP.Y.[ST]	GO:0006096	1.49E-31	82	218	2968	33758	glycolysis	
2	[ILMV]Y.[ILMV]P	GO:0004365	2.21E-31	61	125	2968	33758	glyceraldehyde-3-phosphate dehydrogenase (phosphory	
2	[ILMV]Y.[ILMV]P	GO:0008943	2.21E-31	61	125	2968	33758	glyceraldehyde-3-phosphate dehydrogenase activity	
2	[ILMV]Y.[ILMV]P	GO:0006468	6.75E-29	232	1242	2968	33758	protein amino acid phosphorylation	phosphorylation
2	[ILMV]Y.[ILMV]P	GO:0006007	1.33E-28	82	237	2968	33758	glucose catabolism	
2	[ILMV]Y.[ILMV]P	GO:0019320	1.85E-28	82	238	2968	33758	hexose catabolism	
2	[ILMV]Y.[ILMV]P	GO:0046365	1.85E-28	82	238	2968	33758	monosaccharide catabolism	
2	[ILMV]Y.[ILMV]P	GO:0016903	5.28E-28	70	181	2968	33758	oxidoreductase activity	
2	[ILMV]Y.[ILMV]P	GO:0046164	6.88E-28	82	242	2968	33758	alcohol catabolism	
2	[ILMV]Y.[ILMV]P	GO:0016620	3.46E-27	65	162	2968	33758	oxidoreductase activity	

Downloaded from www.mcponline.org by on January 9, 2009

3	YESP..[DE]	GO:0046641	5.09E-10	3	4	18	33758	positive regulation of alpha-beta T cell proliferation	
3	YESP..[DE]	GO:0046640	5.09E-10	3	4	18	33758	regulation of alpha-beta T cell proliferation	
3	YESP..[DE]	GO:0004715	1.27E-09	3	5	18	33758	non-membrane spanning protein tyrosine kinase activity	
3	YESP..[DE]	GO:0046638	1.27E-09	3	5	18	33758	positive regulation of alpha-beta T cell differentiation	
3	YESP..[DE]	GO:0046635	1.27E-09	3	5	18	33758	positive regulation of alpha-beta T cell activation	T cell activation
3	YESP..[DE]	GO:0046637	1.27E-09	3	5	18	33758	regulation of alpha-beta T cell differentiation	
3	YESP..[DE]	GO:0046632	1.27E-09	3	5	18	33758	alpha-beta T cell differentiation	
3	YESP..[DE]	GO:0046634	1.27E-09	3	5	18	33758	regulation of alpha-beta T cell activation	
3	YESP..[DE]	GO:0046633	2.54E-09	3	6	18	33758	alpha-beta T cell proliferation	
3	YESP..[DE]	GO:0046631	4.45E-09	3	7	18	33758	alpha-beta T cell activation	
4	[ILV][ILV]Y	GO:0007166	1.00E-100	2495	3618	12941	33758	cell surface receptor linked signal transduction	cell surface receptor linked signal transduction
4	[ILV][ILV]Y	GO:0016021	1.00E-100	4269	7083	12941	33758	integral to membrane	
4	[ILV][ILV]Y	GO:0031224	1.00E-100	4275	7098	12941	33758	intrinsic to membrane	
4	[ILV][ILV]Y	GO:0044425	1.00E-100	4485	7749	12941	33758	membrane part	
4	[ILV][ILV]Y	GO:0016020	1.00E-100	5724	10201	12941	33758	membrane	
4	[ILV][ILV]Y	GO:0004930	1.00E-100	2109	2630	12941	33758	G-protein coupled receptor activity	
4	[ILV][ILV]Y	GO:0004888	1.00E-100	2384	3355	12941	33758	transmembrane receptor activity	
4	[ILV][ILV]Y	GO:0007186	1.00E-100	2152	2833	12941	33758	G-protein coupled receptor protein signaling pathway	
4	[ILV][ILV]Y	GO:0001584	1.00E-100	1712	2104	12941	33758	rhodopsin-like receptor activity	
4	[ILV][ILV]Y	GO:0004872	2.03E-306	2794	4329	12941	33758	receptor activity	
5	[DE][DE]...Y.N	GO:0001518	1.67E-37	20	25	294	33758	voltage-gated sodium channel complex	
5	[DE][DE]...Y.N	GO:0005248	2.75E-36	20	27	294	33758	voltage-gated sodium channel activity	
5	[DE][DE]...Y.N	GO:0005272	1.50E-29	20	46	294	33758	sodium channel activity	
5	[DE][DE]...Y.N	GO:0005083	2.42E-18	26	301	294	33758	small GTPase regulator activity	
5	[DE][DE]...Y.N	GO:0006814	5.65E-18	20	152	294	33758	sodium ion transport	
5	[DE][DE]...Y.N	GO:0005099	4.85E-15	8	11	294	33758	Ras GTPase activator activity	Ras GTPase activator activity
5	[DE][DE]...Y.N	GO:0046580	1.45E-14	8	12	294	33758	negative regulation of Ras protein signal transduction	
5	[DE][DE]...Y.N	GO:0030695	4.53E-14	29	578	294	33758	GTPase regulator activity	
5	[DE][DE]...Y.N	GO:0051058	8.64E-14	8	14	294	33758	negative regulation of small GTPase mediated signal tra	
5	[DE][DE]...Y.N	GO:0005244	1.34E-12	21	324	294	33758	voltage-gated ion channel activity	
6	Y.N[ILMV]..L	GO:0050875	3.38E-13	315	17780	455	33758	cellular physiological process	
6	Y.N[ILMV]..L	GO:0016311	4.26E-13	24	270	455	33758	dephosphorylation	
6	Y.N[ILMV]..L	GO:0046839	1.48E-12	11	37	455	33758	phospholipid dephosphorylation	
6	Y.N[ILMV]..L	GO:0030258	1.22E-11	11	44	455	33758	lipid modification	
6	Y.N[ILMV]..L	GO:0050808	1.06E-10	10	40	455	33758	synapse organization and biogenesis	
6	Y.N[ILMV]..L	GO:0051258	1.69E-10	13	87	455	33758	protein polymerization	

Downloaded from www.mcponline.org by on January 9, 2009

6	Y.N[ILMV]..L	GO:0030198	2.08E-10	14	107	455	33758	extracellular matrix organization and biogenesis	
6	Y.N[ILMV]..L	GO:0043062	2.08E-10	14	107	455	33758	extracellular structure organization and biogenesis	
6	Y.N[ILMV]..L	GO:0006355	4.47E-10	91	3476	455	33758	regulation of transcription	regulation of transcription
6	Y.N[ILMV]..L	GO:0006351	4.80E-10	92	3536	455	33758	transcription	
7	[DE]Y.N[ILMV]	GO:0004985	2.06E-14	10	17	548	33758	opioid receptor activity	
7	[DE]Y.N[ILMV]	GO:0016668	7.25E-13	8	11	548	33758	oxidoreductase activity	
7	[DE]Y.N[ILMV]	GO:0050808	1.10E-12	12	40	548	33758	synapse organization and biogenesis	
7	[DE]Y.N[ILMV]	GO:0004731	1.78E-11	6	6	548	33758	purine-nucleoside phosphorylase activity	
7	[DE]Y.N[ILMV]	GO:0004791	1.23E-10	6	7	548	33758	thioredoxin-disulfide reductase activity	
7	[DE]Y.N[ILMV]	GO:0016654	1.23E-10	6	7	548	33758	oxidoreductase activity	
7	[DE]Y.N[ILMV]	GO:0005262	4.38E-10	15	112	548	33758	calcium channel activity	
7	[DE]Y.N[ILMV]	GO:0005158	4.86E-10	8	20	548	33758	insulin receptor binding	insulin receptor binding
7	[DE]Y.N[ILMV]	GO:0016667	7.75E-10	8	21	548	33758	oxidoreductase activity	
7	[DE]Y.N[ILMV]	GO:0007191	1.22E-08	7	19	548	33758	dopamine receptor	
8	Y[ILMV].M.P	GO:0004984	1.24E-52	184	1362	1277	33758	olfactory receptor activity	
8	Y[ILMV].M.P	GO:0001584	1.13E-34	201	2104	1277	33758	rhodopsin-like receptor activity	
8	Y[ILMV].M.P	GO:0004888	8.31E-33	267	3355	1277	33758	transmembrane receptor activity	
8	Y[ILMV].M.P	GO:0004930	8.91E-32	225	2630	1277	33758	G-protein coupled receptor activity	
8	Y[ILMV].M.P	GO:0004872	2.60E-27	302	4329	1277	33758	receptor activity	
8	Y[ILMV].M.P	GO:0007186	1.15E-25	221	2833	1277	33758	G-protein coupled receptor protein signaling pathway	
8	Y[ILMV].M.P	GO:0007166	2.07E-25	261	3618	1277	33758	cell surface receptor linked signal transduction	
8	Y[ILMV].M.P	GO:0004871	9.81E-22	338	5459	1277	33758	signal transducer activity	signal transducer activity
8	Y[ILMV].M.P	GO:0016021	6.67E-21	408	7083	1277	33758	integral to membrane	
9	Y[DE].[ILMV].[ILMV]	GO:0016020	8.95E-21	540	10201	1277	33758	membrane	
9	Y[DE].[ILMV].[ILMV]	GO:0008066	2.39E-33	106	418	2280	33758	glutamate receptor activity	
9	Y[DE].[ILMV].[ILMV]	GO:0007215	2.07E-26	81	313	2280	33758	glutamate signaling pathway	
9	Y[DE].[ILMV].[ILMV]	GO:0007216	1.28E-25	79	307	2280	33758	metabotropic glutamate receptor signaling pathway	
9	Y[DE].[ILMV].[ILMV]	GO:0019028	1.38E-25	54	150	2280	33758	viral capsid	
9	Y[DE].[ILMV].[ILMV]	GO:0008067	1.97E-25	89	381	2280	33758	metabotropic glutamate	
9	Y[DE].[ILMV].[ILMV]	GO:0044423	2.92E-25	54	152	2280	33758	virion part	
9	Y[DE].[ILMV].[ILMV]	GO:0019012	2.92E-25	54	152	2280	33758	virion	
9	Y[DE].[ILMV].[ILMV]	GO:0003824	4.04E-21	880	10036	2280	33758	catalytic activity	catalytic activity
9	Y[DE].[ILMV].[ILMV]	GO:0003674	5.20E-21	2138	29740	2280	33758	molecular_function	
9	Y[DE].[ILMV].[ILMV]	GO:0030554	3.62E-18	311	2845	2280	33758	adenyl nucleotide binding	
10	[DE]..Y[DE].[ILMV]	GO:0016787	2.97E-24	327	3774	1687	33758	hydrolase activity	
10	[DE]..Y[DE].[ILMV]	GO:0003824	5.99E-21	677	10036	1687	33758	catalytic activity	catalytic activity

Downloaded from www.mcponline.org by on January 9, 2009

ASBMB Molecular & Cellular Proteomics

10	[DE]..Y[DE].[ILMV]	GO:0017076	2.41E-18	285	3435	1687	33758	purine nucleotide binding	
10	[DE]..Y[DE].[ILMV]	GO:0005509	9.61E-18	151	1454	1687	33758	calcium ion binding	
10	[DE]..Y[DE].[ILMV]	GO:0007018	1.14E-17	51	253	1687	33758	microtubule-based movement	
10	[DE]..Y[DE].[ILMV]	GO:0030705	2.18E-17	52	266	1687	33758	cytoskeleton-dependent intracellular transport	
10	[DE]..Y[DE].[ILMV]	GO:0007017	6.41E-17	63	382	1687	33758	microtubule-based process	
10	[DE]..Y[DE].[ILMV]	GO:0007010	1.99E-15	90	725	1687	33758	cytoskeleton organization and biogenesis	
10	[DE]..Y[DE].[ILMV]	GO:0007156	6.02E-15	48	263	1687	33758	homophilic cell adhesion	
10	[DE]..Y[DE].[ILMV]	GO:0000166	2.97E-14	314	4177	1687	33758	nucleotide binding	
11	Y[ILMV]PM.P	GO:0005158	1.34E-20	7	20	13	33758	insulin receptor binding	insulin receptor binding
11	Y[ILMV]PM.P	GO:0005102	7.47E-09	7	822	13	33758	receptor binding	
11	Y[ILMV]PM.P	GO:0005515	1.55E-06	11	6998	13	33758	protein binding	
11	Y[ILMV]PM.P	GO:0019901	1.58E-06	3	61	13	33758	protein kinase binding	
11	Y[ILMV]PM.P	GO:0019900	2.41E-06	3	70	13	33758	kinase binding	
11	Y[ILMV]PM.P	GO:0030225	1.07E-05	2	13	13	33758	macrophage differentiation	
11	Y[ILMV]PM.P	GO:0004871	2.87E-05	9	5459	13	33758	signal transducer activity	
11	Y[ILMV]PM.P	GO:0019899	6.12E-05	3	206	13	33758	enzyme binding	
11	Y[ILMV]PM.P	GO:0030224	0.0001168	2	42	13	33758	monocyte differentiation	
11	Y[ILMV]PM.P	GO:0008286	0.0001403	2	46	13	33758	insulin receptor signaling pathway	
12	DY[ILMV].[ILMV]	GO:0030106	1.79E-150	131	172	1281	33758	MHC class I receptor activity	
12	DY[ILMV].[ILMV]	GO:0042612	7.12E-150	131	173	1281	33758	MHC class I protein complex	
12	DY[ILMV].[ILMV]	GO:0042611	1.07E-141	131	188	1281	33758	MHC protein complex	
12	DY[ILMV].[ILMV]	GO:0001772	3.18E-132	132	213	1281	33758	immunological synapse	
12	DY[ILMV].[ILMV]	GO:0019882	2.70E-115	131	258	1281	33758	antigen presentation	
12	DY[ILMV].[ILMV]	GO:0006955	1.54E-51	156	1016	1281	33758	immune response	immune response
12	DY[ILMV].[ILMV]	GO:0006952	1.70E-46	162	1189	1281	33758	defense response	
12	DY[ILMV].[ILMV]	GO:0044459	2.69E-46	175	1371	1281	33758	plasma membrane part	
12	DY[ILMV].[ILMV]	GO:0009607	2.91E-45	164	1242	1281	33758	response to biotic stimulus	
12	DY[ILMV].[ILMV]	GO:0005886	6.48E-38	189	1798	1281	33758	plasma membrane	
13	[ILMV].Y[ILMV].[ILMV]D	GO:0005884	1.85E-25	17	46	236	33758	actin filament	
13	[ILMV].Y[ILMV].[ILMV]D	GO:0005200	9.41E-22	17	71	236	33758	structural constituent of cytoskeleton	
13	[ILMV].Y[ILMV].[ILMV]D	GO:0005524	7.57E-12	53	2735	236	33758	ATP binding	
13	[ILMV].Y[ILMV].[ILMV]D	GO:0006476	2.23E-11	7	18	236	33758	protein amino acid deacetylation	
13	[ILMV].Y[ILMV].[ILMV]D	GO:0030554	3.31E-11	53	2845	236	33758	adenyl nucleotide binding	
13	[ILMV].Y[ILMV].[ILMV]D	GO:0005158	5.36E-11	7	20	236	33758	insulin receptor binding	insulin receptor binding
13	[ILMV].Y[ILMV].[ILMV]D	GO:0048005	2.33E-10	7	24	236	33758	antigen presentation	
13	[ILMV].Y[ILMV].[ILMV]D	GO:0042613	5.20E-10	6	15	236	33758	MHC class II protein complex	

Downloaded from www.mcponline.org by on January 9, 2009

13	[ILMV].Y[ILMV].[ILMV]D	GO:0016575	5.20E-10	6	15	236	33758	histone deacetylation	
14	[ILMV]Y..[ILMV][ILMV]	GO:0015629	8.06E-10	17	359	236	33758	actin cytoskeleton	
14	[ILMV]Y..[ILMV][ILMV]	GO:0016021	1.88E-122	1829	7083	5481	33758	integral to membrane	
14	[ILMV]Y..[ILMV][ILMV]	GO:0031224	8.30E-122	1830	7098	5481	33758	intrinsic to membrane	
14	[ILMV]Y..[ILMV][ILMV]	GO:0044425	8.14E-108	1911	7749	5481	33758	membrane part	
14	[ILMV]Y..[ILMV][ILMV]	GO:0004930	3.03E-99	846	2630	5481	33758	G-protein coupled receptor activity	G-protein coupled receptor activity
14	[ILMV]Y..[ILMV][ILMV]	GO:0016020	4.59E-97	2322	10201	5481	33758	membrane	
14	[ILMV]Y..[ILMV][ILMV]	GO:0007186	8.51E-97	886	2833	5481	33758	G-protein coupled receptor protein signaling pathway	
14	[ILMV]Y..[ILMV][ILMV]	GO:0001584	7.12E-88	699	2104	5481	33758	rhodopsin-like receptor activity	
14	[ILMV]Y..[ILMV][ILMV]	GO:0007166	1.11E-81	1016	3618	5481	33758	cell surface receptor linked signal transduction	
14	[ILMV]Y..[ILMV][ILMV]	GO:0004888	1.84E-81	960	3355	5481	33758	transmembrane receptor activity	
14	[ILMV]Y..[ILMV][ILMV]	GO:0044464	1.23E-73	4085	21607	5481	33758	cell part	
15	[AG][ILMV]Y..P	GO:0004197	7.31E-15	31	242	762	33758	cysteine-type endopeptidase activity	
15	[AG][ILMV]Y..P	GO:0008234	5.68E-14	33	296	762	33758	cysteine-type peptidase activity	
15	[AG][ILMV]Y..P	GO:0004198	2.05E-13	12	27	762	33758	calpain activity	
15	[AG][ILMV]Y..P	GO:0004984	6.70E-11	71	1362	762	33758	olfactory receptor activity	
15	[AG][ILMV]Y..P	GO:0004930	2.43E-10	110	2630	762	33758	G-protein coupled receptor activity	
15	[AG][ILMV]Y..P	GO:0030286	1.99E-09	13	65	762	33758	dynein complex	
15	[AG][ILMV]Y..P	GO:0046839	6.50E-09	10	37	762	33758	phospholipid dephosphorylation	
15	[AG][ILMV]Y..P	GO:0007166	1.03E-08	133	3618	762	33758	cell surface receptor linked signal transduction	cell surface receptor linked signal transduction
15	[AG][ILMV]Y..P	GO:0043235	1.49E-08	15	105	762	33758	receptor complex	
15	[AG][ILMV]Y..P	GO:0050853	1.60E-08	7	15	762	33758	B cell receptor signaling pathway	
16	P.EY.....[ILMV]	GO:0005891	3.84E-10	8	41	235	33758	voltage-gated calcium channel complex	
16	P.EY.....[ILMV]	GO:0003938	2.29E-09	4	4	235	33758	IMP dehydrogenase activity	
16	P.EY.....[ILMV]	GO:0005245	1.20E-08	8	62	235	33758	voltage-gated calcium channel activity	
16	P.EY.....[ILMV]	GO:0005262	9.71E-08	9	112	235	33758	calcium channel activity	
16	P.EY.....[ILMV]	GO:0006177	1.57E-07	4	8	235	33758	GMP biosynthesis	
16	P.EY.....[ILMV]	GO:0046037	1.57E-07	4	8	235	33758	GMP metabolism	
16	P.EY.....[ILMV]	GO:0051018	1.57E-07	4	8	235	33758	protein kinase A binding	
16	P.EY.....[ILMV]	GO:0043167	1.60E-07	77	6299	235	33758	ion binding	ion binding
16	P.EY.....[ILMV]	GO:0043169	2.00E-07	74	5985	235	33758	cation binding	
16	P.EY.....[ILMV]	GO:0035019	3.33E-07	3	3	235	33758	somatic stem cell maintenance	
17	Y[DE].[ILMV]H	GO:0051020	9.24E-12	15	106	440	33758	GTPase binding	
17	Y[DE].[ILMV]H	GO:0019899	3.60E-11	19	206	440	33758	enzyme binding	
17	Y[DE].[ILMV]H	GO:0050794	5.18E-11	127	5572	440	33758	regulation of cellular process	
17	Y[DE].[ILMV]H	GO:0005085	7.46E-11	23	326	440	33758	guanyl-nucleotide exchange factor activity	

Downloaded from www.mcponline.org by on January 9, 2009

ASBMB Molecular & Cellular Proteomics

17	Y[DE].[ILMV]H	GO:0050789	1.65E-10	134	6107	440	33758	regulation of biological process	
17	Y[DE].[ILMV]H	GO:0030695	1.96E-10	30	578	440	33758	GTPase regulator activity	
17	Y[DE].[ILMV]H	GO:0005622	2.95E-10	242	13615	440	33758	intracellular	
17	Y[DE].[ILMV]H	GO:0031323	1.74E-09	96	3990	440	33758	regulation of cellular metabolism	
17	Y[DE].[ILMV]H	GO:0019222	4.67E-09	97	4127	440	33758	regulation of metabolism	regulation of metabolism
17	Y[DE].[ILMV]H	GO:0051244	4.91E-09	113	5101	440	33758	regulation of cellular physiological process	
18	Y[ILMV]PQ	GO:0000145	5.83E-22	14	30	277	33758	exocyst	
18	Y[ILMV]PQ	GO:0006904	1.69E-18	14	48	277	33758	vesicle docking during exocytosis	
18	Y[ILMV]PQ	GO:0048278	6.04E-18	14	52	277	33758	vesicle docking	
18	Y[ILMV]PQ	GO:0044448	4.31E-17	14	59	277	33758	cell cortex part	
18	Y[ILMV]PQ	GO:0005938	1.62E-15	14	75	277	33758	cell cortex	
18	Y[ILMV]PQ	GO:0004629	1.74E-14	11	40	277	33758	phospholipase C activity	
18	Y[ILMV]PQ	GO:0004435	1.30E-13	10	34	277	33758	phosphoinositide phospholipase C activity	
18	Y[ILMV]PQ	GO:0004434	1.30E-13	10	34	277	33758	inositol or phosphatidylinositol phosphodiesterase activity	
18	Y[ILMV]PQ	GO:0006887	3.06E-12	14	127	277	33758	exocytosis	
18	Y[ILMV]PQ	GO:0005509	8.52E-11	39	1454	277	33758	calcium ion binding	ion binding
19	H[ST]G.KPY.C..CG	GO:0003676	1.00E-100	656	6624	656	33758	nucleic acid binding	nucleic acid binding
19	H[ST]G.KPY.C..CG	GO:0008270	1.00E-100	647	3698	656	33758	zinc ion binding	
19	H[ST]G.KPY.C..CG	GO:0046914	1.00E-100	647	4442	656	33758	transition metal ion binding	
19	H[ST]G.KPY.C..CG	GO:0043169	1.00E-100	647	5985	656	33758	cation binding	
19	H[ST]G.KPY.C..CG	GO:0046872	1.00E-100	647	6200	656	33758	metal ion binding	
19	H[ST]G.KPY.C..CG	GO:0043167	1.00E-100	647	6299	656	33758	ion binding	
19	H[ST]G.KPY.C..CG	GO:0005622	1.31E-263	656	13615	656	33758	intracellular	
19	H[ST]G.KPY.C..CG	GO:0006355	1.11E-238	410	3476	656	33758	regulation of transcription	
19	H[ST]G.KPY.C..CG	GO:0006351	7.41E-237	411	3536	656	33758	transcription	
19	H[ST]G.KPY.C..CG	GO:0045449	4.75E-227	411	3738	656	33758	regulation of transcription	
20	[AG].Y..[ILMV].[KR]	GO:0006468	3.70E-44	194	1242	1753	33758	protein amino acid phosphorylation	
20	[AG].Y..[ILMV].[KR]	GO:0004365	5.56E-42	59	125	1753	33758	glyceraldehyde-3-phosphate dehydrogenase (phosphory	
20	[AG].Y..[ILMV].[KR]	GO:0008943	5.56E-42	59	125	1753	33758	glyceraldehyde-3-phosphate dehydrogenase activity	
20	[AG].Y..[ILMV].[KR]	GO:0004672	3.68E-41	188	1235	1753	33758	protein kinase activity	kinase activity
20	[AG].Y..[ILMV].[KR]	GO:0016310	5.90E-40	200	1394	1753	33758	phosphorylation	
20	[AG].Y..[ILMV].[KR]	GO:0016903	3.11E-39	67	181	1753	33758	oxidoreductase activity	
20	[AG].Y..[ILMV].[KR]	GO:0006096	3.44E-39	73	218	1753	33758	glycolysis	
20	[AG].Y..[ILMV].[KR]	GO:0016773	5.62E-39	201	1427	1753	33758	phosphotransferase activity	
20	[AG].Y..[ILMV].[KR]	GO:0006796	3.19E-38	220	1668	1753	33758	phosphate metabolism	
20	[AG].Y..[ILMV].[KR]	GO:0006793	3.19E-38	220	1668	1753	33758	phosphorus metabolism	

Downloaded from www.mcponline.org by on January 9, 2009

ASBMB Molecular & Cellular Proteomics

ID (SwissProt/GenBank)	Phospho Residue	Number In Group	Peptide Sequence	Domain Family	Domain	PubMed ID
TIE2_HUMAN	1108	10	YVNTTLYEKFTYA	PTB	DOK-R_PTB	12665569
IL4RA_HUMAN	497	10	IAGNPAYRSFSNS	PTB	FRIP_PTB	9697832
IGF1R_HUMAN	980	10	ASVNPEYFSAADV	PTB	IRS-1_PTB;Shc_PTB	7559478
EGFR_HUMAN	1110	10	SVQNPVYHNQPLN	PTB	PTB_EGFR	8845374
EGFR_HUMAN	1172	10	SLDNPDYQQDFFP	PTB	PTB_EGFR	8845374
INSR_HUMAN	999	10	ASSNPEYLSASDV	PTB	IRS-2_PTB	8662806
TPOR_HUMAN	626	10	HIANHSYLPLSYW	PTB	Shc_PTB	9122198
IL2RB_HUMAN	364	10	CFTNQGYFFFHLP	PTB	Shc_PTB	8643566
PLS1_HUMAN	69	10	VPNQPVYNQPVYN	PTB	Shc_PTB	12871937
PLS1_HUMAN	74	10	VYNQPVYNQPVGA	PTB	Shc_PTB	12871937
VGFR1_HUMAN	1333	10	YNSVVLYSTPPIX	Crk_SH2	Crk_SH2	9600074
PAXI_HUMAN	31	10	LSEETPYSYPTGN	Crk_SH2	Crk_SH2	16254239
PAXI_HUMAN	118	10	GEEEHVYSFPNKQ	Crk_SH2	Crk_SH2	16254239
PAXI_HUMAN	181	10	GALSPLYGVPETN	Crk_SH2	Crk_SH2	16254239
GAB1_HUMAN	373	10	SDTDSSYCIPTAG	Crk_SH2	Crk_SH2	10734310
GAB1_HUMAN	406	10	ASSQDCYDIPRAF	Crk_SH2	Crk_SH2	10734310
CBL_HUMAN	700	10	GEEDTEYMTPSSR	Crk_SH2	Crk_SH2	12522270
EPOR_HUMAN	456	10	PHLKYLYLVVSDS	Crk_SH2	CrkL_SH2	12027890
CBL_HUMAN	337	10	GFREGFYLFPDGR	Crk_SH2	CrkL_SH2	10608804
CBL_HUMAN	774	10	ENEDDGYDVPKPP	Crk_SH2	CrkL_SH2	12522270
INSR_HUMAN	1189	31	DIYETDYYRKGGK	Grb_SH2	Grb10_SH2	9506989
INSR_HUMAN	1190	31	IYETDYYRKGGKG	Grb_SH2	Grb10_SH2	9506989
EPHB1_CHICK	928	31	AVKMSQYRDNFLS	Grb_SH2	Grb10_SH2	7731717
EPHB2_MOUSE	938	31	AIKMGQYKESFAN	Grb_SH2	Grb10_SH2	8798570
IRS1_RAT	895	31	PKSPGEYVNIEFG	Grb_SH2	Grb2_SH2	8530377
VGFR1_HUMAN	1213	31	SSDDVRYVNAFKF	Grb_SH2	Grb2_SH2	9600074
ACHD_TORCA	393	31	ISKAQEYFNIKSR	Grb_SH2	Grb2_SH2	9185541
CSF3R_MOUSE	788	31	TPSPKSYENIWFH	Grb_SH2	Grb2_SH2	8608222
FAK1_HUMAN	925	31	RSNDKVYENVTGL	Grb_SH2	Grb2_SH2	7997267
PTN11_HUMAN	546	31	KRKGHEYTNIKYS	Grb_SH2	Grb2_SH2	8041791
BCR_HUMAN	177	31	DAEKPFYVNVEFH	Grb_SH2	Grb2_SH2	12522270
SHC1_HUMAN	349	31	EPPDHQYYNDFPG	Grb_SH2	Grb2_SH2	9927207
SHC1_HUMAN	427	31	LFDDPSYVNVQNL	Grb_SH2	Grb2_SH2	9927207
CD28_HUMAN	191	31	RLLHSDYMNMTPR	Grb_SH2	Grb2_SH2	8992971
ITB3_HUMAN	773	31	TANNPLYKEATST	Grb_SH2	Grb2_SH2	8631894
ADRB2_MESAU	350	31	RSSSKAYGNGYSS	Grb_SH2	Grb2_SH2	8557631
LAT_MOUSE	175	31	VESCEDYVNVPES	Grb_SH2	Grb2_SH2	11368773
LAT_MOUSE	195	31	LDGSREYVNVSPE	Grb_SH2	Grb2_SH2	11368773
LAT_MOUSE	235	31	GEEAPDYENLQEL	Grb_SH2	Grb2_SH2	11368773
FAK2_HUMAN	881	31	RTDDLVYLNVMEL	Grb_SH2	Grb2_SH2	10880513
PGFRB_HUMAN	716	31	PPSAELYSNALPV	Grb_SH2	Grb2_SH2	7935391
ERBB2_HUMAN	1139	31	CSPQPEYVNQPDV	Grb_SH2	Grb2_SH2	11500516
MET_HUMAN	1356	31	VHVNATYVNVKCV	Grb_SH2	Grb2_SH2	8662889
CD19_HUMAN	348	31	SGPQNQYGNVLSL	Grb_SH2	Grb2_SH2	10706702
CD19_HUMAN	378	31	GGTAPSYGNPSSD	Grb_SH2	Grb2_SH2	10706702
TIE2_HUMAN	1102	31	LEERKTYVNTTLY	Grb_SH2	Grb2_SH2	12082108
ERBB3_HUMAN	1199	31	EDEEYEYMNRRRR	Grb_SH2	Grb7_SH2	9516479
ERBB3_HUMAN	1262	31	PDEDYEYMNRQRD	Grb_SH2	Grb7_SH2	9516479
PTN11_HUMAN	584	31	EDSARVYENVGLM	Grb_SH2	Grb7_SH2	8622870
CAV1_MOUSE	14	31	DSEGHLYTVPIRE	Grb_SH2	Grb7_SH2	15592455
PGFRB_HUMAN	775	31	SNYMAPYDNYVPS	Grb_SH2	Grb7_SH2	8940081
IRS1_RAT	608	24	LHTDDGYMPMSPG	p85_PI3K_SH2	p85_PI3K_SH2	11572799
IRS1_RAT	628	24	RKGNGDYMPMSPK	p85_PI3K_SH2	p85_PI3K_SH2	11572799

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

IRS1_RAT	939	24	ETGSEEYMNMDLG	p85_PI3K_SH2	p85_PI3K_SH2	11572799
IGF1R_HUMAN	1346	24	FDERQPYAHMNGG	p85_PI3K_SH2	p85_PI3K_SH2	8550552
CBL_HUMAN	371	24	QEQYELYCEMGST	p85_PI3K_SH2	p85_PI3K_SH2	10635327
CBL_HUMAN	731	24	QIDSCTYEAMYNI	p85_PI3K_SH2	p85_PI3K_SH2	10934191
NTRK2_MOUSE	782	24	TCPQEVYELMLGC	p85_PI3K_SH2	p85_PI3K_SH2	8226808
ERBB3_HUMAN	1276	24	GGPGGDYAAMGAC	p85_PI3K_SH2	p85_PI3K_SH2	11546794
ERBB3_HUMAN	1289	24	PASEQGYEEMRAF	p85_PI3K_SH2	p85_PI3K_SH2	11546794
GAB1_HUMAN	447	24	EELDENYVPMNPN	p85_PI3K_SH2	p85_PI3K_SH2	10734310
GAB1_HUMAN	472	24	PIQEANYVPMTPG	p85_PI3K_SH2	p85_PI3K_SH2	10734310
GAB1_HUMAN	589	24	HDSEENYVPMNPN	p85_PI3K_SH2	p85_PI3K_SH2	10734310
GAB2_MOUSE	441	24	ASSDDNYVPMNPG	p85_PI3K_SH2	p85_PI3K_SH2	11334882
GAB2_MOUSE	465	24	DNSQSVYIPMSPG	p85_PI3K_SH2	p85_PI3K_SH2	11334882
GAB2_MOUSE	573	24	GDSEENYVPMQNP	p85_PI3K_SH2	p85_PI3K_SH2	11334882
P85A_HUMAN	688	24	AEPYNLYSSLKEL	p85_PI3K_SH2	p85_PI3K_SH2	11337495
IL7RA_HUMAN	449	24	SNQEEAYVTMSSF	p85_PI3K_SH2	p85_PI3K_SH2	7719938
CTLA4_HUMAN	201	24	PLTTGVYVKMPPT	p85_PI3K_SH2	p85_PI3K_SH2	9813138
EPOR_HUMAN	504	24	EPLPPSYVACSXX	p85_PI3K_SH2	p85_PI3K_SH2	9573010
PGFRB_HUMAN	740	24	GESDGGYMDMSKD	p85_PI3K_SH2	p85_PI3K_SH2	8195171
MET_HUMAN	1313	24	YCPDPLYEVMLKC	p85_PI3K_SH2	p85_PI3K_SH2	10069803
MET_HUMAN	1349	24	TFIGEHYVHVNAT	p85_PI3K_SH2	p85_PI3K_SH2	7731718
CD19_HUMAN	500	24	SLGSQSYEDMRGI	p85_PI3K_SH2	p85_PI3K_SH2	10933394
CD19_HUMAN	531	24	EEDADSYENMDNP	p85_PI3K_SH2	p85_PI3K_SH2	10933394
VGFR1_HUMAN	794	16	SEIKTDYLSIIMD	PLCgamma_SH2	PLCgamma_SH2	9398617
VGFR1_HUMAN	1169	16	QQDGKDYIPINAI	PLCgamma_SH2	PLCgamma_SH2	9299537
VGFR2_HUMAN	801	16	GELKTGYLSIVMD	PLCgamma_SH2	PLCgamma_SH2	11741094
VGFR2_HUMAN	1175	16	QQDGKDYIVLPIS	PLCgamma_SH2	PLCgamma_SH2	10749680
ZAP70_HUMAN	315	16	PMDTSVYESPYSD	PLCgamma_SH2	PLCgamma_SH2	15592455
ZAP70_HUMAN	319	16	SVYESPYSDPEEL	PLCgamma_SH2	PLCgamma_SH2	15592455
3BP2_HUMAN	183	16	DDEDDSYLEPDSP	PLCgamma_SH2	PLCgamma_SH2	11390470
KSYK_MOUSE	342	16	PMDTEVYESPYAD	PLCgamma_SH2	PLCgamma_SH2	9820500
KSYK_MOUSE	346	16	EVYESPYADPEEI	PLCgamma_SH2	PLCgamma_SH2	9820500
LAT_MOUSE	136	16	DDYPNGYLVVLPD	PLCgamma_SH2	PLCgamma_SH2	11368773
NTRK2_MOUSE	816	16	AKASPVYLDILGX	PLCgamma_SH2	PLCgamma_SH2	10533983
PGFRB_HUMAN	1021	16	NEGDNDYIIPLPD	PLCgamma_SH2	PLCgamma1_SH2	12062403
NTRK1_HUMAN	496	16	IIENPQYFSDACV	PLCgamma_SH2	PLCgamma1_SH2	8155326
NTRK1_HUMAN	791	16	AQAPPVYLDVLGX	PLCgamma_SH2	PLCgamma1_SH2	8155326
FGFR1_HUMAN	766	16	LTSNQEYLDLSMP	PLCgamma_SH2	PLCgamma1_SH2	8622701
FGFR3_HUMAN	760	16	VTSTDEYLDLSAP	PLCgamma_SH2	PLCgamma1_SH2	11294897
EPHB3_HUMAN	614	9	YIDPFTYEDPNEA	RasGap_SH2	RasGap_SH2	9674711
DOK1_HUMAN	296	9	DSPPALYAEPLDS	RasGap_SH2	RasGap_SH2	10688886
DOK1_HUMAN	315	9	PSQDSLYSDPLDS	RasGap_SH2	RasGap_SH2	10688886
DOK1_HUMAN	362	9	PKEDPIYDEPEGL	RasGap_SH2	RasGap_SH2	15592455
DOK1_HUMAN	377	9	VPPQGLYDLPREP	RasGap_SH2	RasGap_SH2	15592455
EPHB2_MOUSE	604	9	TPGMKIYIDPFTY	RasGap_SH2	RasGap_SH2	9233798
PGFRB_HUMAN	771	9	DIESSNYMAPYDN	RasGap_SH2	RasGAP_SH2	8195171
DOK1_HUMAN	398	9	RVKEEGYELPYNP	RasGap_SH2	RasGAP_SH2	15592455
DOK1_HUMAN	409	9	NPATDDYAVPPPR	RasGap_SH2	RasGAP_SH2	15592455
ZAP70_HUMAN	474	13	LLVNRHYAKISDF	Shc_SH2	Shc_SH2	9685404
NTRK2_MOUSE	515	13	VIENPQYFGITNS	Shc_SH2	Shc_SH2	10533983
BCR_HUMAN	276	13	PWPPLEYQPYQSI	Shc_SH2	Shc_SH2	8955135
BCR_HUMAN	279	13	PLEYQPYQSIYVG	Shc_SH2	Shc_SH2	8955135
BCR_HUMAN	283	13	QPYQSIYVGGMME	Shc_SH2	Shc_SH2	8955135
GAB2_MOUSE	193	13	TSAPQEYLYLHQC	Shc_SH2	Shc_SH2	12819203
IL3RB_HUMAN	593	13	FDFNGPYLGPPHS	Shc_SH2	Shc_SH2	10362354
IL3RB_HUMAN	766	13	KSGFEGYVELPPI	Shc_SH2	Shc_SH2	9257712
PGFRB_HUMAN	579	13	SSDGHEYIYVDPM	Shc_SH2	Shc_SH2	8195171
PGFRB_HUMAN	751	13	KDESVDYVPMLDM	Shc_SH2	Shc_SH2	8195171
ERBB3_HUMAN	1328	13	AFDNPDYWHSRLF	Shc_SH2	Shc_SH2	9694850
ERBB2_HUMAN	1221	13	PAFDNLYYWDQDP	Shc_SH2	Shc_SH2	11500516
ERBB2_HUMAN	1222	13	AFDNLYYWDQDPP	Shc_SH2	Shc_SH2	11500516
PGFRB_HUMAN	1009	12	DTSSVLYTAVQPN	SHP-2_SH2	SHP-2_SH2	8195171
CSF3R_MOUSE	728	12	PALVQAYVLQGDP	SHP-2_SH2	SHP-2_SH2	10329697
GAB1_HUMAN	627	12	GDKQVEYLDLDLD	SHP-2_SH2	SHP-2_SH2	10734310

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

GAB1_HUMAN	659	12	ADERVDYVVVDQQ	SHP-2_SH2	SHP-2_SH2	10734310
GAB2_MOUSE	603	12	STGSVDYLALDFQ	SHP-2_SH2	SHP-2_SH2	9885561
GAB2_MOUSE	632	12	SDEKVDYVQVDKE	SHP-2_SH2	SHP-2_SH2	11640929
IRS1_RAT	1222	12	SEDLSTYASINFQ	SHP-2_SH2	SHP-2_SH2	12220227
IL6RB_HUMAN	759	12	TSSTVQYSTVVHS	SHP-2_SH2	SHP-2_SH2	9505191
EPOR_HUMAN	426	12	SAASFEYTILDPS	SHP-2_SH2	SHP-2_SH2	10660611
LEPR_MOUSE	985	12	RQPSVKYATLVSN	SHP-2_SH2	SHP-2_SH2	10799542
TIE2_HUMAN	1113	12	LYEKFTYAGIDCS	SHP-2_SH2	SHP-2_SH2	12665569
IRS1_RAT	1172	12	LEKSLNYIDLDLV	SHP-2_SH2	SHP-2_SH2	11572799
VGFR2_HUMAN	1214	14	CDPKFHYDNTAGI	Src_SH2	Src_SH2	11387210
Q90738_CHICK	93	14	SSLPEGYYEEAVP	Src_SH2	Src_SH2	9655255
Q90738_CHICK	94	14	SLPEGYYEEAVPV	Src_SH2	Src_SH2	9655255
Q90738_CHICK	451	14	DPGALHYDYIDVE	Src_SH2	Src_SH2	9655255
Q90738_CHICK	453	14	GALHYDYIDVEMT	Src_SH2	Src_SH2	9655255
SRC_HUMAN	529	14	TSTEPQYQPGENL	Src_SH2	Src_SH2	12387813
BCA1_HUMAN	653	14	DSPDGQYENSEGG	Src_SH2	Src_SH2	12480533
BCA1_HUMAN	664	14	GGWMEDYDYVHLQ	Src_SH2	Src_SH2	11604500
PAXI_HUMAN	40	14	PTGNHTYQEIAVP	Src_SH2	Src_SH2	16254239
EGFR_HUMAN	1016	14	VVDADEYLIPQQG	Src_SH2	Src_SH2	7680558
EGFR_HUMAN	1125	14	PSRDPHYQDPHST	Src_SH2	Src_SH2	10075741
FYN_HUMAN	530	14	TATEPQYQPGENL	Src_SH2	Src_SH2	1985196
FAK2_HUMAN	402	14	SIESDIYAEIPDE	Src_SH2	Src_SH2	15105428
PGFRB_HUMAN	581	14	DGHEYIYVDPMQL	Src_SH2	Src_SH2	8617789
IL6RB_HUMAN	905	19	EGMPKSYLPQTVR	Stat_SH2	Stat1_SH2	10661409
IL6RB_HUMAN	915	19	TVRQGGYMPQXXX	Stat_SH2	Stat1_SH2	10661409
INGR1_HUMAN	457	19	APTSFGYDKPHVL	Stat_SH2	Stat1_SH2	8156998
LEPR_MOUSE	1138	19	GENFVPYMPQFQT	Stat_SH2	Stat3_SH2	10799542
CSF3R_MOUSE	768	19	SPGVMQYIRSDST	Stat_SH2	Stat3_SH2	8608222
IL6RB_HUMAN	767	19	TVVHSGYRHQVPS	Stat_SH2	Stat3_SH2	10661409
IL6RB_HUMAN	814	19	ILPRQQYFKQNCS	Stat_SH2	Stat3_SH2	10661409
TPOR_HUMAN	631	19	SYLPLSYWQQPXX	Stat_SH2	Stat3_SH2	9122198
PRLR_HUMAN	611	19	QLGGLDYLDPACF	Stat_SH2	Stat5_SH2	7537382
IL9R_HUMAN	407	19	RVQTLAYLPQEDW	Stat_SH2	Stat5_SH2	8756628
GHR_HUMAN	534	19	FLMDNAYFCEADA	Stat_SH2	Stat5_SH2	8647880
GHR_HUMAN	566	19	LNQEDIYITTESL	Stat_SH2	Stat5_SH2	8647880
GHR_HUMAN	627	19	FLSSCGYVSTDQL	Stat_SH2	Stat5_SH2	8647880
IL3RB_HUMAN	628	19	PPGSLEYLCLPAG	Stat_SH2	Stat5_SH2	10362354
IL2RB_HUMAN	418	19	SGEDDAYCTFPSR	Stat_SH2	Stat5_SH2	10214954
IL2RB_HUMAN	536	19	PLNTDAYLSLQEL	Stat_SH2	Stat5_SH2	10214954
IL4RA_HUMAN	575	19	SAPTSGYQEFVHA	Stat_SH2	Stat6_SH2	10358772
IL4RA_HUMAN	603	19	PPGEAGYKAFSSL	Stat_SH2	Stat6_SH2	9712048
IL4RA_HUMAN	631	19	SSGEEGYKPFQDL	Stat_SH2	Stat6_SH2	10358772
LCP2_HUMAN	113	4	SFEEDDYESPNDD	Vav_SH2	Vav_SH2	9047237
LCP2_HUMAN	128	4	GEDDGDYESPNEE	Vav_SH2	Vav_SH2	9047237
CD19_HUMAN	409	4	EEEGEGYEEPDSE	Vav_SH2	Vav_SH2	10706702
CD19_HUMAN	439	4	SQDGSGYENPEDE	Vav_SH2	Vav_SH2	10706702
91760	181	481	KGQRDLYSGLNQR	P		
91760	170	481	PVPNPDYEPIRKG	P		
4557431	152	481	TATKDTYDALHMQ	P		
4557431	83	481	LGRREEYDVLDKR	P		
4557431	141	481	KGHDGLYQGLSTA	P		
4557431	110	481	NPQEGLYNELQKD	P		
4557431	122	481	DKMAEAYSEIGMK	P		
4557431	72	481	QGQNQLYNELNLG	P		
7656965	453	481	SHVDNEYSQPPRN	P		
29849	15	481	KIGEGTYGVVYKA	P		
4503357	449	481	SHNSALYSQVQKS	P		
34289	27	481	LIEDNEYTAREGA	P		
22325381	940	481	EEPEALYAAVNKK	P		
107499	220	481	DISYKRYIPETLN	P		
107499	1034	481	IETDKEYYTVKDD	P		
107499	1035	481	ETDKEYYTVKDDR	P		
2780223	191	481	LDGSREYVNVSQE	P		

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

5031887	300	481	GRYYEGYYAAGPG	P
2554639	188	481	DDEMTGYVATRWY	P
7682684	227	481	KAEFAEYASVDRN	P
7682684	417	481	VPKENDYESISDL	P
7682684	341	481	TVPESTYTSIQGD	P
7682684	359	481	SSCNDLYATVKDF	P
7657577	148	481	PGTPVKYSEVVLD	P
32015	271	481	HFPLATYAPVISA	P
31543838	292	481	AEGEPCYIRDSGV	P
31543838	1054	481	VPEGHEYYRVRED	P
31543838	1055	481	PEGHEYYRVREDG	P
346421	493	481	GADDSYYTARSAG	P
346421	492	481	LGADDSYYTARSA	P
346421	248	481	KADGLIYCLKEAC	P
346421	292	481	TLNSDGYTPEPAR	P
125135	393	481	LMTGDTYTAHAGA	P
125135	185	481	ASDGKLYVSSESR	P
125135	264	481	EGVWKKYSLTVAV	P
125135	226	481	RNKPTVYGVSPNY	P
125135	469	481	GCPEKVYELMRAC	P
8567380	91	481	FGQQKLYQVPNSQ	P
8567380	188	481	HSTQGVYDIPPSS	P
115857	655	481	SSANAIYSLAARP	P
115857	681	481	SEEDTEYMTPTSR	P
6015042	321	481	GQRIYQYIQSRFY	P
7305329	273	481	VKAGESYTHIRYL	P
125135	253	481	KLGGGQYGEVYEG	P
125135	257	481	GQYGEVYEGVWKK	P
18553914	827	481	FASDPKYATPQVI	P
11038641	644	481	SLETLLYKPVDRV	P
11038641	591	481	ASQLGVYRAFVDN	P
4885205	421	481	LPSSPVYEDAASF	P
4557447	1068	481	KELEEIYMLPRMR	P
2245671	305	481	STSSGGYRRTPSV	P
4503851	259	481	SVDSSLYNLPRSY	P
18105042	266	481	EFRDSTYDLPRSL	P
4503881	10	481	SRGKSTYYWPRPR	P
4503359	299	481	RGQEGEYAVPFDA	P
4505843	482	481	ATYAEPYRPIQYR	P
4755142	1135	481	TLSEVDYAPAGPA	P
4755142	1162	481	RGLPSDYGRPLSF	P
4755142	671	481	RGSRDTYAWHKQK	P
4755142	886	481	GTRERLYEWISID	P
4755142	986	481	SFNNPAYYVLEGV	P
10835031	350	481	PPDHQYYNDFPGK	P
10835051	260	481	KIVDNGYGYVSNG	P
10835059	300	481	KLVDGCYSFWQAG	P
10835101	734	481	KDPDQLYTTLKNL	P
10835189	22	481	AGAVASYDYLVIG	P
10863977	35	481	LHSVDQYLNIKLT	P
11141893	44	481	ASSQETYGKSPFW	P
11230802	266	481	TYVSSFYHAFSGA	P
11321583	84	481	KSPDEAYAIAKKL	P
11321601	651	481	YTTDFIYQLYSEE	P
11321607	56	481	NIVFGDYTWTEFD	P
11321634	88	481	VQRISTYGLPAGG	P
11415026	63	481	SSGEIVYCGQVFE	P
1169475	29	481	TTGHLIYKCGGID	P
12083691	251	481	ERNTEFYHIPTHS	P
1220311	113	481	EHALLAYTLGVKQ	P
12232415	453	481	GKSELPYEELWLE	P
12644292	279	481	GEPNVSYICSRYY	P
12654583	24	481	IQLLDDYPKCFIV	P

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

12746446	586	481	IKRNRYYDEDEDE	P
12963489	342	481	LNGDTDYFGQQFD	P
12963491	25	481	TVEVDLYTAKGLF	P
12963491	44	481	GASTGIYEALELR	P
12963669	26	481	NRFSGWYDADLSP	P
12963847	444	481	REVDSEYEARSRL	P
13128860	221	481	GAGKGKYYAVNYP	P
132956	307	481	NNASTDYDLSDKS	P
1335288	144	481	LRIVEPYIAWGYP	P
1335288	200	481	DLIHEIYTVGKRF	P
13376058	158	481	SRGDNVYVVTEVL	P
13376058	37	481	STGFQPYCLVVRK	P
13376735	110	481	GTFDNIYLHVQIS	P
13385132	39	481	AIIVGDYGPMWVY	P
13385136	97	481	DEVEDLYELLKLV	P
13385164	215	481	LLDEPAYGLIQKI	P
13540492	213	481	LDRHAAYGGYSTP	P
13540492	216	481	HAAYGGYSTPEDR	P
13569841	131	481	VVYENAYGRFIGP	P
13569874	352	481	SARAFPYGNVAFP	P
14141161	454	481	FPIPEEYTFIQNV	P
14149680	1009	481	RDPPDPYVSLLLL	P
14149680	822	481	TKHLSPYATLTVG	P
14149771	33	481	PARNLTYFGVVHG	P
14327963	392	481	ETQKSIYYITGES	P
14861844	67	481	SKDKDAYSSFGSR	P
15082252	281	481	HTGEKPYECIKCG	P
15277856	274	481	YSNTRAYYFVKFK	P
15718704	351	481	GQEAPIYELTSQF	P
15778828	313	481	LFDDPSYVNIQNL	P
15808988	12	481	TYDDRAYSSFGGG	P
15808988	45	481	EPPYTAYVGNLPF	P
16359158	218	481	IKEKLCYVALDFE	P
16753233	70	481	AGKALDYYMLRNG	P
17384428	1233	481	PDTEEIYKLTGTG	P
17941283	48	481	SLPNPEYYTLRYA	P
17998694	477	481	DPDSNPYSLLDNT	P
17999535	849	481	DNDITPYLVSRFY	P
18375501	262	481	LYPNTPYAYTFWT	P
18375642	62	481	IQNTGDYYDLYGG	P
18426902	534	481	GGEDPLYVARRLV	P
18426902	562	481	TQAVAAYQGCHFI	P
18590404	122	481	HTGEKPYACLECG	P
18700002	338	481	SKLSDSYSNTLPV	P
18700002	506	481	SQGLDDYGARSVS	P
18765758	145	481	GYDDDNYDYIVKN	P
19526792	326	481	RLGEDPYYTENGG	P
19913412	13	481	IIRIPPYHYIHVL	P
19923497	226	481	INQEGEYIKMFMR	P
20824027	354	481	GDQIDTYELSGGA	P
20832385	214	481	SWYDNEYGYSNRV	P
20833867	690	481	YEPVSPYSCREAP	P
20884633	53	481	MGQKDSYVGDKAQ	P
20897586	607	481	ENTEDQYSLVEDD	P
20915086	253	481	ETSKVIYDFIEKT	P
20986529	187	481	TGFLTEYVATRWY	P
20988575	98	481	AHDGGIYAISWSP	P
2117873	466	481	ARQAHLYRGIFPV	P
21264341	147	481	YDEDSPYQNIKIL	P
21281693	88	481	PSPLPVYSSSAKN	P
21311861	337	481	PLPNCTYATRQAI	P
21313156	172	481	VSLEEIYSGCTKR	P
21313428	488	481	EFLQQKYKEEKKG	P

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

21313500	63	481	IQQEQIYQQSRAY	P
21361070	680	481	NEKCHDYYTTEFL	P
21361282	53	481	DADDAVYELNGKD	P
21361598	500	481	ARIQRAYNHYFDL	P
2145060	153	481	HTGEKPYTCPDCG	P
21542258	142	481	SAWFQDYSSGGGG	P
21614520	401	481	QPNEAVYTKMMTK	P
21618338	704	481	DPGAAPYLKTKFI	P
21618340	705	481	PGSAAPYLKTKFI	P
21703974	35	481	PSTDPLYDTPDTR	P
21706701	38	481	TGAKNLYIISVKG	P
21735417	424	481	SHEDGDYYEVDIN	P
21735596	152	481	DQENCVYETVVLP	P
21735621	56	481	VSRLTLYDIAHTP	P
21746169	508	481	GGKEAVYSGVQSL	P
22045052	45	481	KLTEGQYVLCRWT	P
22074743	352	481	ASPYNDYEGRKGD	P
22325385	858	481	PASVTDYQNVSFS	P
223582	51	481	RISGLIYEETRGV	P
22507319	666	481	EDRDLLYSIDAYR	P
22538467	102	481	LGASGDYADFQYL	P
22749307	110	481	NARLSIYGIWFYD	P
22749307	133	481	MKNLTQYEQLKAH	P
22749307	191	481	TSSSAIYDNPNLI	P
23397502	250	481	HTGEKPYKCLECG	P
23397584	352	481	KAVCSTYLQSRYY	P
23510283	202	481	HTSNNIYEAVKYS	P
23510303	130	481	KADLDHYTGRDSE	P
23956292	460	481	PLDGSPYAQVQRV	P
24119274	497	481	ALEDFVYKIFGPG	P
24308009	139	481	KLVQTTYECLMQA	P
24418901	615	481	GPSGALYARVARR	P
24431955	520	481	HTGENLYKCKVCA	P
25049644	470	481	ESSQSPYERQPLS	P
25054791	140	481	ELQANCYEEVKDR	P
25070577	765	481	GPEQEEYSTFVID	P
25072051	900	481	LRENPYYDSRVVG	P
25121938	1105	481	NEEENIYSVPHDS	P
25141235	104	481	DHGRNDYDGIGGR	P
25188179	195	481	EFGGSIYQKVNEK	P
26024215	295	481	GSPPALYAEPLDS	P
26024215	314	481	PSQDSVYSDPLGS	P
26024215	340	481	PLYWDLYGHVQQQ	P
26024215	361	481	SKEDPIYDEPEGL	P
26024215	376	481	APPRGLYDLPQEP	P
26024215	450	481	SSDTALYSQVQKS	P
27229221	79	481	SEQYSDYSYRTNP	P
27229221	81	481	QYSDYSYRTNPPK	P
27229238	394	481	LYEEITYSGISDG	P
27369543	131	481	LEEDVIYDDVPCE	P
27370398	196	481	VQPEPIYAESAKR	P
27370424	476	481	QPLCKPYTVVDVS	P
27370424	638	481	YDNLAIYKSFLGT	P
27370424	796	481	CSVEELYAVPPDA	P
27370424	879	481	RSTSSPYHASNLL	P
27881484	648	481	VWLLQQYSGMTHE	P
28076915	876	481	NLANNPYSGDVTK	P
28173550	319	481	KLAAVTYNGVDNN	P
28477118	154	481	SWYNNEYGYSNRE	P
28477132	1291	481	SSTETCYSAIPKS	P
28477382	202	481	ERENVQYSSVNGD	P
28477911	186	481	TPGSATYARYKPS	P
28477911	282	481	SKGEPQYSSHSSS	P

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

28485943	228	481	PGGSDNYGSLSRV	P
28485943	257	481	PSRQDVYGPQPQV	P
28485943	280	481	RFHPEPYGLEDDQ	P
28485943	334	481	EVPPDQYYWAPLA	P
28485943	96	481	DHNHLLYSTIPRM	P
28486522	894	481	KGYYSPYSVSGSG	P
28488252	659	481	GTAEPDYGALYEG	P
28491046	347	481	LGETASYSLFLPT	P
28492519	904	481	YDNANIYSRTREY	P
28493790	245	481	PSSGQIYGPGPRG	P
28493790	301	481	GRYYEPYYAAGPS	P
28493790	302	481	RYYEPYYAAGPSY	P
28493835	734	481	RESEGLYGQPGDE	P
28495148	419	481	KNRIAIYELLFKK	P
28495256	211	481	VDGDNIYVRHSNL	P
28495256	494	481	ETKTITYESPQID	P
28500488	764	481	RTYDETYTRYRPV	P
28506227	847	481	QKTEDAYTISSAG	P
28506227	881	481	SKDSNPYATLPRA	P
28506274	661	481	EDHVDIYNLRSKL	P
28510510	492	481	LEDDAIYSVHVPA	P
28510510	501	481	HVPAGLYRIRKGV	P
28514402	672	481	RTTDGVYEGVAIG	P
28546037	192	481	QGPPEIYSDTQFP	P
28549040	24	481	STPPSAYGSVKPY	P
29029632	1078	481	ELQSPEYKLSKLR	P
29029632	1096	481	TDYNPNYCFAGKT	P
29029632	1131	481	GAFGEVYEGQVSG	P
29029632	1278	481	GMARDIYRASYYR	P
29029632	1282	481	DIYRASYYRKGGC	P
29029632	1507	481	SLWNPTYGSWFTE	P
29029632	1584	481	PCGNVNYGYQQQG	P
29029632	1604	481	APGAGHYEDTILK	P
29294649	664	481	LPDGTFYSTLYLP	P
29469069	359	481	KTVCSTYLQSRYY	P
29648305	201	481	MDPSDIYAVIQIP	P
29789319	845	481	NRASQDYDQLPSS	P
30155975	313	481	HTGEKPYKCEERG	P
30156077	164	481	IKEKLCYVALDFK	P
30160191	360	481	ETAKGNYPLEAVH	P
31560645	309	481	LTADFLYDEVHPK	P
31652244	548	481	TSSDTTYKGGASE	P
3183544	54	481	TRRSLGYAYVNFQ	P
31981690	15	481	IDLGTTYSCVGVF	P
33239421	855	481	RFPLDYYSIPFPT	P
33942081	954	481	QDSIFEYFNTAPL	P
339578	369	481	EKSEGTYCCGPVP	P
34447209	336	481	LKKEDIYAVEIVG	P
34447209	89	481	EKSNLAYDIVQLP	P
37538013	113	481	NSHDPSYFACLIK	P
37747535	136	481	AMCRLDYCGGGGG	P
4501865	131	481	ATREGDYVLFHHE	P
4501881	55	481	MGQKDSYVGDEAQ	P
4501885	294	481	DIRKDLYANTVLS	P
4502101	207	481	SDARALYEAGERR	P
4502709	15	481	KIGEGTYGVVYKG	P
4502709	19	481	GTYGVVYKGRHKT	P
4503165	58	481	LSFEELYRNAYTM	P
4503277	79	481	NKLDTKYPYVCHA	P
4503429	174	481	DHLAYRYEVLKII	P
4503513	308	481	SGGEDGYVRIHYF	P
4503523	318	481	LAMEATYINHNFS	P
4503527	525	481	WQKNGDYLCVKVD	P

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

4503571	287	481	KSFIKDYPVVSIE	P
4503633	216	481	QDPNPQYSPIIKQ	P
4503633	305	481	EAPSNIYVEVEDE	P
4503633	39	481	SCQNLGYTAASPQ	P
4503915	348	481	KGYPGDYTKGVEI	P
4503971	333	481	NRKSDIYVCMISY	P
4504445	295	481	YGGGGQYFAKPRN	P
4504521	227	481	RGYISPYFINTSK	P
4504689	400	481	TEAPGEYFFSDGI	P
4504705	355	481	GGSDQNYDIVTIG	P
4504715	140	481	ENGSKGYAFVHFE	P
4504771	113	481	VTTCNDYVALVHP	P
4504965	124	481	YSEEEKYAFVNWI	P
4505185	37	481	TGKPPQYIAVHVV	P
4505571	148	481	CSVCPDYDLCSVC	P
4505575	361	481	VATKPLYVALAQR	P
4505763	196	481	MKKELNYFAKALE	P
4505839	105	481	ASDPILYRPVAVA	P
4505839	390	481	EAEAAIYHLQLFE	P
4505839	83	481	SHGTHEYHAETIK	P
4505939	145	481	KRLTHVYDLCKGK	P
4506053	381	481	PPSQGDYHGCPFR	P
4506181	24	481	KLVQIEYALAAVA	P
4506181	98	481	KLAQQYYLVYQEP	P
4506183	161	481	SGVSYGYWGCAIG	P
4506203	154	481	VTGPHLYSIYPHG	P
4506275	309	481	LTADFLYEEVHPK	P
4506431	615	481	KHFTNPYCNIYLN	P
4506597	14	481	NEIKVVYLRCTGG	P
4506633	103	481	DSPNKLYTLVTYV	P
4506633	108	481	LYTLVTYVPVTTF	P
4506663	133	481	ARASGNYATVISH	P
4506679	12	481	KNRIAIYELLFKE	P
4506685	38	481	DVKEQIYKLAKKG	P
4507123	151	481	PDYPPNYILFLNN	P
4507241	311	481	NMSGSLYEMVSRV	P
4507499	50	481	HTSNDPYCFVEFY	P
4507521	275	481	QIIQEIYSQIQSK	P
4507713	40	481	YNEAYNYYTKAID	P
4507747	11	481	EDLPKSYDYDLII	P
4507747	13	481	LPKSYDYDLIIIG	P
4507747	131	481	VVYENAYGQFIGP	P
4507747	422	481	RDNNKCYAKIICN	P
4507869	39	481	FSRVQIYHNPTAN	P
4507879	195	481	EFGGSIYQKVNKK	P
4507895	117	481	NDRFANYIDKVRF	P
4507909	291	481	ETSKLIYDFIEDQ	P
4508007	382	481	HTGEKPYQCGQCG	P
4557032	240	481	MVVESAYEVIKLK	P
4557305	204	481	DLKRCQYVTEKVL	P
4557469	737	481	HRQGHIYMEMNFT	P
4557491	367	481	FNHTEDYVLLPDE	P
4557619	173	481	EGDNVNYDWIHWD	P
4557681	785	481	SLISSDYELLSDP	P
4585643	231	481	HSGVKPYECTECG	P
4585643	343	481	HSGENPYECLECG	P
4585643	483	481	HTGEKPYECVECG	P
4585643	511	481	HSGEKPYECVECG	P
4758030	249	481	DTCRGHYNNVSCA	P
4758138	297	481	EDFLKDYIHINIG	P
4758442	84	481	DDGRVSYPLCFIF	P
4758528	216	481	RVCEPCYEQLNRK	P
4758662	385	481	QNGQTIYLRGTGD	P

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

4826898	129	481	LINKKCYEMASHL	P
4826960	491	481	GRLNLHYAVVSKR	P
4827065	278	481	SKFDTIYQILLKK	P
4885153	207	481	PEPAHAYAQPQTT	P
4885409	437	481	LINRMDYVEINID	P
5031573	231	481	KAVKERYSYVCPD	P
5031753	306	481	ATENDIYNFFSPL	P
5031799	914	481	EIINPNYMGVGPF	P
5031805	46	481	GPARLEYYENEKK	P
5031805	662	481	RVDPNGYMMMSPS	P
5031857	239	481	QVVESAYEVIKLK	P
5031977	188	481	NLDGLEYKLHDFG	P
5031981	32	481	DTAEQVYISSLAL	P
5032069	56	481	TGQHQGYGFVEFL	P
5032087	759	481	GKQKLQYEGIFIK	P
5107727	43	481	PAGTEDYIHIRIQ	P
5174447	228	481	NEGKHLYTLDGGD	P
5453555	147	481	KKNLQYYDISAKS	P
5453998	311	481	LKVLYQYKEKQYM	P
5729840	83	481	NLDPLVYLLSKLT	P
5803167	414	481	CEICGNYTYRGPK	P
5901926	40	481	ERTINLYPLTNYT	P
5901948	225	481	EVKRNVYDLTSIP	P
5902076	189	481	HEGETAYIRVKVD	P
5902162	279	481	HSGEKPYGCVECG	P
5902162	335	481	HTGEKPYECVQCG	P
6005727	505	481	DTYLGKYWAIKLA	P
6005749	67	481	AKKEGPYDVVVLP	P
6005798	222	481	TTPTKVYSEVHFT	P
6005846	309	481	LTAEFLYDEVHPK	P
6005854	128	481	QRLGLDYEERVLP	P
6671509	166	481	THTVPIYEGYALP	P
6671666	31	481	SDMHCGYGDSPSK	P
6678129	424	481	LIEDNEYTARQGA	P
6678179	473	481	RISEQTYQLSRWT	P
6678351	1323	481	SHRLSIYEEWDPF	P
6678355	1164	481	EQPAPAYEVHNRY	P
6678355	1177	481	RPEAQPYSSTGPK	P
6678359	275	481	QIIQEIYSQVQSK	P
6679321	458	481	REYDQLYEEYTRT	P
6679567	310	481	TPDHVVYARSKTA	P
6679599	183	481	ETEVELYNEFPEP	P
6679669	562	481	HRSLEQYDQVPDG	P
6679673	492	481	PADGYAYSSSMYH	P
6679673	490	481	YPPADGYAYSSSM	P
6679741	397	481	VSETDDYAEIIDE	P
6679741	576	481	YMEDSTYYKASKG	P
6679773	345	481	QEDGGVYSSSGLK	P
6680217	216	481	RVCEPCYEQLNKK	P
6681017	251	481	KRVPCAYDKTALA	P
6681087	141	481	SAVGFEYQGKTEK	P
6681087	421	481	PPSSPIYEDAAPF	P
6753524	132	481	KTQQGLYQAPGPN	P
6753524	238	481	QTEQDEYDTPRHL	P
6753524	331	481	LLREETYDVPPAF	P
6753524	376	481	PPAPDLYDVPPGL	P
6753524	414	481	VVDDGVYAVPPPA	P
6753620	104	481	DRGRGDYDGIGGR	P
6753926	520	481	DPDSNPYSLLDTS	P
6754472	29	481	KDPVGVYCRVRPL	P
6754816	17	481	NPETPGYVGFANL	P
6754968	139	481	PLTEASYVNLPTI	P
6755502	921	481	GEPEALYAAVTKK	P

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

6755809	1116	481	AQGNENYAGIAAR	P
6755975	53	481	STSRSLYSSSPGG	P
6755975	61	481	SSPGGAYVTRSSA	P
6857826	205	481	PHLSKGYAYVEFE	P
6996913	238	481	YKSYSPYDMLESI	P
7106303	453	481	TYDEIFYTLSPVN	P
7108367	826	481	WWRGEIYGRVGWF	P
7305031	479	481	EVGTGQYATTKGI	P
7305485	315	481	PSCKALYDFEPEN	P
7305503	124	481	GVEDPEYAVTQLA	P
7549795	554	481	REDAVLYLLEIPK	P
7656977	47	481	NLNTITYETLKYI	P
7656991	301	481	ITDESPYVHYLNT	P
7657403	74	481	LQRTLHYECIVLV	P
7657431	84	481	TPQGREYGMIYLG	P
7661646	122	481	TLQQAEYEFLSFV	P
7661646	184	481	RWYPEEYEFAPKK	P
7661972	43	481	RKEIADYLAAGKD	P
7662260	237	481	LSLANCYEVLSAA	P
7669492	42	481	PFIDLNYMVYMFQ	P
7706007	65	481	QPDSAGYTALHYA	P
7706326	86	481	FNVCNRYLVVLYY	P
7710042	1510	481	VKLQQTYSALNSK	P
7710042	172	481	PQIQDLYGKVDFT	P
7949121	139	481	SGSQGGYDRYSGG	P
8400706	239	481	LQNGPIYARVIQK	P
8922075	518	481	GVKKPTYDPVSED	P
8922856	25	481	LKGSNLYSQLERQ	P
8922987	286	481	FKQAEEYLSFAFE	P
9055302	948	481	KSKQPLYSSIVTV	P
9055324	197	481	APDKRIYQFTAAS	P
9055324	260	481	PIDDEIYEELPEE	P
9055324	75	481	AEDGDEYDDPFAG	P
9624990	192	481	DVGFSSYELVATP	P
9789975	374	481	APVYSVYSKLHPA	P
9910460	145	481	STFDTPYCRVGLG	P
5803137	129	481	GGYGYGYGRSRDY	P
4885405	198	481	KGFGGQYGIQKDR	P
20482720	94	481	GTENRTYQASSAA	P
4885405	103	481	SAVGHEYVAEVEK	P
4503533	211	481	TDSFDDYPPRRGD	P
4503821	757	481	EEGKYGYVLRSYL	P
4758476	45	481	LGSQEGYVPKNFI	P
21361670	162	481	APVGSVYQKTNAV	P
3023518	98	481	GARLGEYEDVSRV	P
4503821	755	481	RNEEGKYGYVLRS	P
5031887	301	481	RYYEGYYAAGPGY	P
4505871	753	481	RDINSLYDVSRMY	P
21361757	165	481	RNMKSQYEVFRSS	P
21265031	374	481	APVYSVYSKLHPP	P
4505317	911	481	LGRSGSYSYLEER	P
8923710	244	481	LQLDNQYAVLENQ	P
4505317	901	481	TSAGDRYDSLLGR	P
8923909	1104	481	QIPEGDYLSYREF	P
4505317	446	481	LRKTGSYGALAEI	P
130225	1253	481	GSFESRYQQPFED	P
1905874	321	481	VTPPEGYEVVTVF	P
19526773	18	481	ALEVAEYLTPVLK	P
20532395	110	481	SGADRIYDLNIPA	P
5453611	209	481	RKHYQPYAPPRDF	P
4502669	149	481	LRNDQVYQPLRDR	P
12003435	602	481	YVDPHTYEDPTQA	P
8572221	972	481	QSAPQIYGPPQYN	P

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

6166197	571	481	TAVEIDYDSLKLK	P
8922679	114	481	QTQPERYIHLENL	P
961446	346	481	AVGSSNYASTSAG	P
29791421	421	481	YEGRTYYSPVYRS	P
5803137	118	481	YGSGRYYDSRPGG	P
16445419	86	481	PKNYGSYSTQASA	P
16445419	35	481	HRPSRQYATLDVY	P
7657577	95	481	LYGNLHYLQTGRL	P
13186251	278	481	EFMCKVYSDPQPH	P
13186251	461	481	LAGVSEYELPEDP	P
13186251	583	481	PGLEYCYNPSHNP	P
13186251	651	481	DIHHIDYYKKTTN	P
13186251	728	481	NCTNELYMMMRDC	P
6740102	249	481	PGPQDIYDVPPVR	P
6740102	410	481	VVDSGVYAVPPPA	P
4759050	707	481	GAMAATYSALNRN	P
4885205	446	481	TEPEPVYSMEAAD	P
14751580	373	481	VKKLKDYAFIHFD	P
33598948	783	481	GRNPGFYVEANPM	P
182530	652	481	IHHIDYYKKTTNG	P
24476013	1007	481	MKLAQQYVMTSLQ	P
4504733	921	481	ADSSSDYVNMDFT	P
13325136	339	481	DSMRVLYYMEKEL	P

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

Peptides

#Exp	IPI	PEP_SEQ	MCR	Charge	Score	Num of Peps
20060324MM_15N	IPI00130621.3	ATTLASTLMEQYMK	798.4	2	76	1 of 7
20060324MM_15N	IPI00130621.3	EPVPMQDQGQVLNDTVDGK	1039.5	2	85	2 of 7
20060324MM_15N	IPI00130621.3	GHATDEWFLLSSHIPLK	653.68	3	16	3 of 7
20060324MM_15N	IPI00130621.3	LLAITELLQQK	639.4	2	43	4 of 7
20060324MM_15N	IPI00130621.3	MFNIISDSPSPIAAR	814.92	2	56	5 of 7
20060324MM_15N	IPI00130621.3	NEDVNTNLAHLLSILSELVEK	787.09	3	32	6 of 7
20060324MM_15N	IPI00130621.3	SVQNLANLVEFGAK	749.41	2	59	7 of 7
20060324MM_19N	IPI00116554.2	ESQSHPGDFVLSVR	784.39	2	65	1 of 12
20060324MM_19N	IPI00116554.2	FATLAELVQYYMEHHGQLK	572.29	4	27	2 of 12
20060324MM_19N	IPI00116554.2	FDSLTDLVEHYK	737.87	2	74	3 of 12
20060324MM_19N	IPI00116554.2	IQNTGDYYDLYGGEK	872.4	2	96	4 of 12
20060324MM_19N	IPI00116554.2	NILPFDHTR	374.87	3	31	5 of 12
20060324MM_19N	IPI00116554.2	NPMVETLGTVLQLK	775.94	2	64	6 of 12
20060324MM_19N	IPI00116554.2	RIEEEQK	475.26	2	24	7 of 12
20060324MM_19N	IPI00116554.2	RNGAVTHIK	338.54	3	56	8 of 12
20060324MM_19N	IPI00116554.2	SNPGDFTLSVR	601.81	2	31	9 of 12
20060324MM_19N	IPI00116554.2	TIQMVR	379.22	2	16	10 of 12
20060324MM_19N	IPI00116554.2	VTHVMIR	433.25	2	30	11 of 12
20060324MM_19N	IPI00116554.2	WFHPNITGVEAENLLLTR	1060.56	2	57	12 of 12
20060324MM_19X	IPI00116554.2	ESQSHPGDFVLSVR	519.93	3	30	1 of 13
20060324MM_19X	IPI00116554.2	FATLAELVQYYMEHHGQLK	1140.07	2	28	2 of 13
20060324MM_19X	IPI00116554.2	FDSLTDLVEHYK	489.58	3	53	3 of 13
20060324MM_19X	IPI00116554.2	FDSLTDLVEHYKK	399.46	4	21	4 of 13
20060324MM_19X	IPI00116554.2	GVDGSFLAR	461.24	2	54	5 of 13
20060324MM_19X	IPI00116554.2	INAAEIESR	501.76	2	62	6 of 13
20060324MM_19X	IPI00116554.2	KGHEYTNIK	545.29	2	32	7 of 13
20060324MM_19X	IPI00116554.2	NILPFDHTR	556.8	2	36	8 of 13
20060324MM_19X	IPI00116554.2	NPMVETLGTVLQLK	514.96	3	19	9 of 13
20060324MM_19X	IPI00116554.2	RNGAVTHIK	332.53	3	26	10 of 13
20060324MM_19X	IPI00116554.2	SNPGDFTLSVR	596.8	2	41	11 of 13
20060324MM_19X	IPI00116554.2	TIQMVR	374.21	2	22	12 of 13
20060324MM_19X	IPI00116554.2	VYENVGLMQQQR	488.91	3	28	13 of 13
20060324MM_21N	IPI00130621.3	EPVPMQDQGQVLNDTVDGK	1039.5	2	64	1 of 8
20060324MM_21N	IPI00130621.3	GHATDEWFLLSSHIPLK	653.68	3	52	2 of 8
20060324MM_21N	IPI00130621.3	IFMASEILPPTLR	749.42	2	28	3 of 8
20060324MM_21N	IPI00130621.3	MFNIISDSPSPIAAR	814.92	2	55	4 of 8
20060324MM_21N	IPI00130621.3	MIMFLDELGNVPELPDTTEHSR	852.41	3	20	5 of 8
20060324MM_21N	IPI00130621.3	NEDVNTNLAHLLSILSELVEK	787.42	3	39	6 of 8
20060324MM_21N	IPI00130621.3	TNENIQR	442.73	2	27	7 of 8
20060324MM_21N	IPI00130621.3	YYNSIGDIIDHYR	546.93	3	35	8 of 8
20060324MM_21X	IPI00130621.3	ATTLASTLMEQYMK	794.39	2	76	1 of 5
20060324MM_21X	IPI00130621.3	IMPEEEYSEFK	701.31	2	20	2 of 5
20060324MM_21X	IPI00130621.3	NEDVNTNLAHLLSILSELVEK	784.42	3	70	3 of 5
20060324MM_21X	IPI00130621.3	VPDTDEISFLK	632.33	2	16	4 of 5
20060324MM_21X	IPI00130621.3	YYNSIGDIIDHYR	543.6	3	43	5 of 5
20060324MM_3N	IPI00130621.3	ATTLASTLMEQYMK	798.4	2	72	1 of 13
20060324MM_3N	IPI00130621.3	EGQNPVWSEEFVFDDLPPDINR	1185.58	3	52	2 of 13
20060324MM_3N	IPI00130621.3	EPVPMQDQGQVLNDTVDGK	1039.5	2	69	3 of 13
20060324MM_3N	IPI00130621.3	FEITLSNK	480.27	2	32	4 of 13
20060324MM_3N	IPI00130621.3	GHATDEWFLLSSHIPLK	653.68	3	38	5 of 13
20060324MM_3N	IPI00130621.3	LLAITELLQQK	639.4	2	59	6 of 13
20060324MM_3N	IPI00130621.3	LLYPVAPPEPVEDR	802.93	2	55	7 of 13
20060324MM_3N	IPI00130621.3	MIMFLDELGNVPELPDTTEHSR	639.31	4	18	8 of 13
20060324MM_3N	IPI00130621.3	TNENIQR	442.73	2	24	9 of 13
20060324MM_3N	IPI00130621.3	VPDTDEISFLK	636.33	2	40	10 of 13
20060324MM_3N	IPI00130621.3	VVSGFVFLR	517.31	2	45	11 of 13
20060324MM_3N	IPI00130621.3	YYNSIGDIIDHYR	546.93	3	35	12 of 13
20060324MM_3N	IPI00130621.3	YYNSIGDIIDHYRK	592.3	3	37	13 of 13
20060324MM_3N	IPI00309413.4	DSESSPSDFSVSLK	746.85	2	58	1 of 4
20060324MM_3N	IPI00309413.4	GVEGDFLIR	508.28	2	36	2 of 4
20060324MM_3N	IPI00309413.4	IYDLNIPAFVK	650.87	2	36	3 of 4
20060324MM_3N	IPI00309413.4	LWLLDDSK	499.28	2	39	4 of 4
20060324MM_3X	IPI00130621.3	EIYNTIR	454.75	2	28	1 of 16
20060324MM_3X	IPI00130621.3	EPVPMQDQGQVLNDTVDGK	1035.5	2	79	2 of 16
20060324MM_3X	IPI00130621.3	EQIVEGYYLK	621.32	2	19	3 of 16
20060324MM_3X	IPI00130621.3	FEITLSNK	476.26	2	36	4 of 16
20060324MM_3X	IPI00130621.3	GHATDEWFLLSSHIPLK	651.01	3	52	5 of 16
20060324MM_3X	IPI00130621.3	IFMASEILPPTLR	744.42	2	39	6 of 16
20060324MM_3X	IPI00130621.3	IMPEEEYSEFK	701.32	2	38	7 of 16
20060324MM_3X	IPI00130621.3	LLYPVAPPEPVEDR	797.93	2	32	8 of 16
20060324MM_3X	IPI00130621.3	MFNIISDSPSPIAAR	809.92	2	60	9 of 16
20060324MM_3X	IPI00130621.3	MIMFLDELGNVPELPDTTEHSR	1272.61	2	73	10 of 16
20060324MM_3X	IPI00130621.3	NEDVNTNLAHLLSILSELVEK	784.42	3	81	11 of 16

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

20060324MM_3X	IPI00130621.3	PSDNTPGDYSLYFR	816.37	2	40	12 of 16
20060324MM_3X	IPI00130621.3	TLLASILLK	486.33	2	21	13 of 16
20060324MM_3X	IPI00130621.3	VVSGFVFLR	512.3	2	51	14 of 16
20060324MM_3X	IPI00130621.3	YYNSIGDIIDHYR	543.59	3	59	15 of 16
20060324MM_3X	IPI00130621.3	YYNSIGDIIDHYRK	439.97	4	21	16 of 16
20060324MM_3X	IPI00309413.4	DSESSPSDFSVSLK	742.84	2	48	1 of 4
20060324MM_3X	IPI00309413.4	EWYYGNVTR	594.28	2	26	2 of 4
20060324MM_3X	IPI00309413.4	GVEGDFLIR	503.27	2	41	3 of 4
20060324MM_3X	IPI00309413.4	LWLLDDSK	495.27	2	46	4 of 4
20060324MM_7N	IPI00312067.4	ALQDMSSTAPPAPLQPSIR	996.01	2	27	1 of 12
20060324MM_7N	IPI00312067.4	GSYGLDLEAVR	595.31	2	68	2 of 12
20060324MM_7N	IPI00312067.4	HVHTYR	411.72	2	22	3 of 12
20060324MM_7N	IPI00312067.4	ISHVSTSSVK	351.53	3	26	4 of 12
20060324MM_7N	IPI00312067.4	LDVTLGDLTK	541.81	2	70	5 of 12
20060324MM_7N	IPI00312067.4	LGVVFEK	400.24	2	37	6 of 12
20060324MM_7N	IPI00312067.4	LLLDTLQLSK	576.36	2	39	7 of 12
20060324MM_7N	IPI00312067.4	LVLLR	312.23	2	26	8 of 12
20060324MM_7N	IPI00312067.4	NVTSWFTSK	539.28	2	23	9 of 12
20060324MM_7N	IPI00312067.4	RLHSEVDK	334.52	3	16	10 of 12
20060324MM_7N	IPI00312067.4	SFENDAQSSDNINFLK	918.93	2	76	11 of 12
20060324MM_7N	IPI00312067.4	VLSGLEILSK	533.83	2	19	12 of 12
20060324MM_7X	IPI00312067.4	DFIFVSAR	477.76	2	26	1 of 11
20060324MM_7X	IPI00312067.4	DFLSGIQK	454.25	2	21	2 of 11
20060324MM_7X	IPI00312067.4	GSYGLDLEAVR	590.3	2	58	3 of 11
20060324MM_7X	IPI00312067.4	LDVTLGDLTK	537.81	2	26	4 of 11
20060324MM_7X	IPI00312067.4	LLLDTLQLSK	572.35	2	38	5 of 11
20060324MM_7X	IPI00312067.4	LYEWISIDK	583.81	2	42	6 of 11
20060324MM_7X	IPI00312067.4	NSFNNPAYYVLEGVPHQLLPLE	1046.21	3	33	7 of 11
20060324MM_7X	IPI00312067.4	SFENDAQSSDNINFLK	914.92	2	80	8 of 11
20060324MM_7X	IPI00312067.4	TLSEVDYAPGPGR	681.34	2	36	9 of 11
20060324MM_7X	IPI00312067.4	VDQLNLER	493.77	2	57	10 of 11
20060324MM_7X	IPI00312067.4	VLSGLEILSK	529.83	2	45	11 of 11
20060324MM_9N	IPI00119058.1	AEEMLSK	408.21	2	50	1 of 12
20060324MM_9N	IPI00119058.1	ATADDELSFK	552.77	2	60	2 of 12
20060324MM_9N	IPI00119058.1	DGFIPK	342.69	2	40	3 of 12
20060324MM_9N	IPI00119058.1	FGNDVQHFK	367.19	3	37	4 of 12
20060324MM_9N	IPI00119058.1	FNSLNELVDYHR	758.88	2	57	5 of 12
20060324MM_9N	IPI00119058.1	GDFIHVMDNSDPNWWK	656.97	3	30	6 of 12
20060324MM_9N	IPI00119058.1	NQQIFLR	464.77	2	36	7 of 12
20060324MM_9N	IPI00119058.1	NYIEMKPHPWFFGK	603.98	3	15	8 of 12
20060324MM_9N	IPI00119058.1	NYVTPVNR	486.76	2	16	9 of 12
20060324MM_9N	IPI00119058.1	PHPWFFGK	341.85	3	35	10 of 12
20060324MM_9N	IPI00119058.1	RGDFIHVMDNSDPNWWK	1068	2	48	11 of 12
20060324MM_9N	IPI00119058.1	YFLWVVK	481.78	2	39	12 of 12
20060324MM_9N	IPI00130621.3	EPVPMQDQGQVLNDTVDGK	1039.51	2	55	1 of 11
20060324MM_9N	IPI00130621.3	EQIVEGYYLK	625.33	2	18	2 of 11
20060324MM_9N	IPI00130621.3	FEITLSNK	480.27	2	21	3 of 11
20060324MM_9N	IPI00130621.3	GHATDEWFLLSSHIPLK	653.68	3	31	4 of 11
20060324MM_9N	IPI00130621.3	GVQQHVLK	458.78	2	15	5 of 11
20060324MM_9N	IPI00130621.3	IFMASEILPPTLR	749.42	2	20	6 of 11
20060324MM_9N	IPI00130621.3	LLAITELLQQK	639.4	2	50	7 of 11
20060324MM_9N	IPI00130621.3	LLYPVAPPEPVEDR	802.93	2	50	8 of 11
20060324MM_9N	IPI00130621.3	MFNIISDSPSPIAAR	814.92	2	28	9 of 11
20060324MM_9N	IPI00130621.3	NEDVNTNLAHLLSILSELVEK	787.09	3	83	10 of 11
20060324MM_9N	IPI00130621.3	TNENIQR	442.73	2	34	11 of 11
20060324MM_9X	IPI00119058.1	AEEMLSK	412.2	2	33	1 of 13
20060324MM_9X	IPI00119058.1	ATADDELSFK	548.76	2	61	2 of 13
20060324MM_9X	IPI00119058.1	DGFIPK	338.69	2	30	3 of 13
20060324MM_9X	IPI00119058.1	DIEQMPQQPTYVQALFDFDPQE	1138.53	3	83	4 of 13
20060324MM_9X	IPI00119058.1	ESESAPGDFSLSVK	726.85	2	41	5 of 13
20060324MM_9X	IPI00119058.1	FGNDVQHFK	546.27	2	46	6 of 13
20060324MM_9X	IPI00119058.1	FNSLNELVDYHR	753.87	2	65	7 of 13
20060324MM_9X	IPI00119058.1	GDFIHVMDNSDPNWWK	654.29	3	31	8 of 13
20060324MM_9X	IPI00119058.1	NYIEMKPHPWFFGK	598.63	3	16	9 of 13
20060324MM_9X	IPI00119058.1	NYVTPVNR	481.76	2	16	10 of 13
20060324MM_9X	IPI00119058.1	PHPWFFGK	339.18	3	26	11 of 13
20060324MM_9X	IPI00119058.1	QRHDGAFLIR	399.22	3	38	12 of 13
20060324MM_9X	IPI00119058.1	YFLWVVK	477.78	2	35	13 of 13
20060324MM_9X	IPI00130621.3	ATTLASTLMEQYMK	794.39	2	76	1 of 13
20060324MM_9X	IPI00130621.3	EPVPMQDQGQVLNDTVDGK	1035.5	2	58	2 of 13
20060324MM_9X	IPI00130621.3	FEITLSNK	476.26	2	18	3 of 13
20060324MM_9X	IPI00130621.3	GHATDEWFLLSSHIPLK	651.01	3	60	4 of 13
20060324MM_9X	IPI00130621.3	GVQQHVLK	454.77	2	39	5 of 13
20060324MM_9X	IPI00130621.3	IFMASEILPPTLR	744.42	2	36	6 of 13

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

20060324MM_9X	IPI00130621.3	IMPEEEYSEFK	701.31	2	26	7 of 13
20060324MM_9X	IPI00130621.3	LLYPVAPPEPVEDR	797.93	2	47	8 of 13
20060324MM_9X	IPI00130621.3	MFNIISDSPSPIAAR	540.28	3	27	9 of 13
20060324MM_9X	IPI00130621.3	MIMFLDELGNVPELPDTTEHSR	636.81	4	27	10 of 13
20060324MM_9X	IPI00130621.3	NEDVNTNLAHLLSILSELVEK	784.42	3	87	11 of 13
20060324MM_9X	IPI00130621.3	SVQNLANLVEFGAK	497.27	3	21	12 of 13
20060324MM_9X	IPI00130621.3	VVSGFVFLR	512.3	2	33	13 of 13
20060405MM_25N	IPI00119058.1	AEEMLSK	408.21	2	49	1 of 15
20060405MM_25N	IPI00119058.1	ATADDELSFK	552.77	2	61	2 of 15
20060405MM_25N	IPI00119058.1	DGFIPK	342.69	2	32	3 of 15
20060405MM_25N	IPI00119058.1	DIEQMPQQPTYVQALFDFDPQE	1141.87	3	58	4 of 15
20060405MM_25N	IPI00119058.1	ESESAPGDFSLSVK	730.85	2	46	5 of 15
20060405MM_25N	IPI00119058.1	FGNDVQHFK	550.27	2	59	6 of 15
20060405MM_25N	IPI00119058.1	FNSLNELVDYHR	506.25	3	73	7 of 15
20060405MM_25N	IPI00119058.1	GDFIHVMDNSDPNWWK	656.97	3	27	8 of 15
20060405MM_25N	IPI00119058.1	HDGAFLIR	469.76	2	37	9 of 15
20060405MM_25N	IPI00119058.1	NQQIFLR	464.77	2	27	10 of 15
20060405MM_25N	IPI00119058.1	NYIEMKPHPWFFGK	609.31	3	30	11 of 15
20060405MM_25N	IPI00119058.1	PHPWFFGK	341.85	3	24	12 of 15
20060405MM_25N	IPI00119058.1	QRHDGAFLIR	608.33	2	43	13 of 15
20060405MM_25N	IPI00119058.1	RGDFIHVMDNSDPNWWK	534.5	4	16	14 of 15
20060405MM_25N	IPI00119058.1	YFLWVVK	481.78	2	28	15 of 15
20060405MM_25X	IPI00119058.1	ATADDELSFK	548.76	2	67	1 of 12
20060405MM_25X	IPI00119058.1	DGFIPK	338.69	2	23	2 of 12
20060405MM_25X	IPI00119058.1	ESESAPGDFSLSVK	726.85	2	47	3 of 12
20060405MM_25X	IPI00119058.1	FGNDVQHFK	364.51	3	38	4 of 12
20060405MM_25X	IPI00119058.1	FNSLNELVDYHR	502.92	3	86	5 of 12
20060405MM_25X	IPI00119058.1	HDGAFLIR	464.75	2	26	6 of 12
20060405MM_25X	IPI00119058.1	NQQIFLR	459.76	2	38	7 of 12
20060405MM_25X	IPI00119058.1	NYIEMK	399.2	2	32	8 of 12
20060405MM_25X	IPI00119058.1	PHPWFFGK	339.18	3	25	9 of 12
20060405MM_25X	IPI00119058.1	QRHDGAFLIR	404.89	3	34	10 of 12
20060405MM_25X	IPI00119058.1	RGDFIHVMDNSDPNWWK	706.33	3	33	11 of 12
20060405MM_25X	IPI00119058.1	YFLWVVK	477.78	2	31	12 of 12
20060405MM_29N	IPI00116554.2	ESQSHPGDFVLSVR	523.26	3	30	1 of 25
20060405MM_29N	IPI00116554.2	FATLAELVQYYMEHHGQLK	762.72	3	88	2 of 25
20060405MM_29N	IPI00116554.2	FDSLTDLVEHYK	492.25	3	76	3 of 25
20060405MM_29N	IPI00116554.2	FDSLTDLVEHYKK	403.47	4	20	4 of 25
20060405MM_29N	IPI00116554.2	FIYMAVQHYIETLQR	961.99	2	28	5 of 25
20060405MM_29N	IPI00116554.2	HGSFLVR	413.23	2	28	6 of 25
20060405MM_29N	IPI00116554.2	KNPMVETLGTVLQLK	843.99	2	45	7 of 25
20060405MM_29N	IPI00116554.2	LLYSR	331.2	2	22	8 of 25
20060405MM_29N	IPI00116554.2	NGAVTHIK	424.25	2	49	9 of 25
20060405MM_29N	IPI00116554.2	NGDVIELK	448.25	2	73	10 of 25
20060405MM_29N	IPI00116554.2	NILPFDHTR	374.87	3	32	11 of 25
20060405MM_29N	IPI00116554.2	NPMVETLGTVLQLK	775.94	2	77	12 of 25
20060405MM_29N	IPI00116554.2	RIEEEQK	475.26	2	20	13 of 25
20060405MM_29N	IPI00116554.2	RNGAVTHIK	338.87	3	42	14 of 25
20060405MM_29N	IPI00116554.2	RWFHPNITGVEAENLLLTR	572.31	4	19	15 of 25
20060405MM_29N	IPI00116554.2	SGMVQTEAQYR	648.3	2	53	16 of 25
20060405MM_29N	IPI00116554.2	TIQMVR	379.22	2	28	17 of 25
20060405MM_29N	IPI00116554.2	TVWQYHFR	573.79	2	29	18 of 25
20060405MM_29N	IPI00116554.2	VGQALLQGNTER	648.35	2	56	19 of 25
20060405MM_29N	IPI00116554.2	VIVMTTK	400.25	2	33	20 of 25
20060405MM_29N	IPI00116554.2	VTHVMIR	433.25	2	34	21 of 25
20060405MM_29N	IPI00116554.2	VYENVGLMQQQR	737.87	2	57	22 of 25
20060405MM_29N	IPI00116554.2	WFHPNITGVEAENLLLTR	707.38	3	57	23 of 25
20060405MM_29N	IPI00116554.2	YDVGGGER	431.7	2	26	24 of 25
20060405MM_29N	IPI00116554.2	YDVGGGERFDSLTDLVEHYK	773.38	3	23	25 of 25
20060405MM_29N	IPI00312067.4	LLLDTLQLSK	576.36	2	39	1 of 2
20060405MM_29N	IPI00312067.4	VLSGLEILSK	533.83	2	21	2 of 2
20060405MM_29X	IPI00116554.2	ESAAHDYTLR	581.78	2	32	1 of 27
20060405MM_29X	IPI00116554.2	ESQSHPGDFVLSVR	779.39	2	46	2 of 27
20060405MM_29X	IPI00116554.2	FATLAELVQYYMEHHGQLK	1139.57	2	55	3 of 27
20060405MM_29X	IPI00116554.2	FDSLTDLVEHYK	489.58	3	66	4 of 27
20060405MM_29X	IPI00116554.2	FDSLTDLVEHYKK	399.46	4	15	5 of 27
20060405MM_29X	IPI00116554.2	FIYMAVQHYIETLQR	638	3	47	6 of 27
20060405MM_29X	IPI00116554.2	GVDGSFLAR	461.24	2	58	7 of 27
20060405MM_29X	IPI00116554.2	HGSFLVR	408.23	2	41	8 of 27
20060405MM_29X	IPI00116554.2	INAAEIESR	501.76	2	53	9 of 27
20060405MM_29X	IPI00116554.2	IQNTGDYYDLYGGEK	579.27	3	22	10 of 27
20060405MM_29X	IPI00116554.2	KNPMVETLGTVLQLK	557.66	3	26	11 of 27
20060405MM_29X	IPI00116554.2	NGAVTHIK	424.25	2	45	12 of 27
20060405MM_29X	IPI00116554.2	NGDVIELK	444.25	2	50	13 of 27

Molecular & Cellular Proteomics ASBMB

Downloaded from www.mcponline.org by on January 9, 2009

20060405MM_29X	IPI00116554.2	NPMVETLGTVLQLK	779.93	2	39	14 of 27
20060405MM_29X	IPI00116554.2	QPLNTTR	415.23	2	35	15 of 27
20060405MM_29X	IPI00116554.2	RNGAVTHIK	332.86	3	34	16 of 27
20060405MM_29X	IPI00116554.2	SGMVQTEAQYR	635.3	2	60	17 of 27
20060405MM_29X	IPI00116554.2	SNPGDFTLSVR	596.8	2	27	18 of 27
20060405MM_29X	IPI00116554.2	TIQMVR	374.21	2	30	19 of 27
20060405MM_29X	IPI00116554.2	TVWQYHFR	568.79	2	16	20 of 27
20060405MM_29X	IPI00116554.2	VGQALLQGNTER	643.35	2	40	21 of 27
20060405MM_29X	IPI00116554.2	VIVMTTK	396.24	2	28	22 of 27
20060405MM_29X	IPI00116554.2	VTHVMIR	428.25	2	34	23 of 27
20060405MM_29X	IPI00116554.2	VYENVGLMQQQR	732.87	2	71	24 of 27
20060405MM_29X	IPI00116554.2	WFHPNITGVEAENLLLTR	1055.56	2	92	25 of 27
20060405MM_29X	IPI00116554.2	YDVGGGER	426.7	2	30	26 of 27
20060405MM_29X	IPI00116554.2	YWPDEYALK	592.79	2	24	27 of 27
20060405MM_31N	IPI00116554.2	ESQSHPGDFVLSVR	784.39	2	48	1 of 14
20060405MM_31N	IPI00116554.2	FATLAELVQYYMEHHGQLK	572.29	4	36	2 of 14
20060405MM_31N	IPI00116554.2	FDSLTDLVEHYK	737.87	2	56	3 of 14
20060405MM_31N	IPI00116554.2	FDSLTDLVEHYKK	403.47	4	23	4 of 14
20060405MM_31N	IPI00116554.2	GVDGSFLAR	466.25	2	58	5 of 14
20060405MM_31N	IPI00116554.2	HGSFLVR	413.23	2	23	6 of 14
20060405MM_31N	IPI00116554.2	NGAVTHIK	424.25	2	37	7 of 14
20060405MM_31N	IPI00116554.2	NGDVIELK	448.25	2	25	8 of 14
20060405MM_31N	IPI00116554.2	NILPFDHTR	374.87	3	20	9 of 14
20060405MM_31N	IPI00116554.2	NPMVETLGTVLQLK	775.94	2	81	10 of 14
20060405MM_31N	IPI00116554.2	SNPGDFTLSVR	601.81	2	27	11 of 14
20060405MM_31N	IPI00116554.2	TVWQYHFR	382.86	3	20	12 of 14
20060405MM_31N	IPI00116554.2	VGQALLQGNTER	648.35	2	58	13 of 14
20060405MM_31N	IPI00116554.2	VYENVGLMQQQR	737.87	2	69	14 of 14
20060405MM_31N	IPI00312067.4	ASGLGEAGMGAWLR	693.35	2	80	1 of 12
20060405MM_31N	IPI00312067.4	DFIFVSAR	482.76	2	24	2 of 12
20060405MM_31N	IPI00312067.4	ELTDLDYRPIAMQSLWNIK	775.41	3	23	3 of 12
20060405MM_31N	IPI00312067.4	GSYGLDLEAVR	595.31	2	53	4 of 12
20060405MM_31N	IPI00312067.4	LDVTLGDLTK	541.81	2	48	5 of 12
20060405MM_31N	IPI00312067.4	LYEWISIDK	587.81	2	22	6 of 12
20060405MM_31N	IPI00312067.4	NVTSWFTSK	539.28	2	17	7 of 12
20060405MM_31N	IPI00312067.4	SFENDAQSSDNINFLK	918.93	2	88	8 of 12
20060405MM_31N	IPI00312067.4	SMIGSTAQQFLTFLSHR	645.33	3	22	9 of 12
20060405MM_31N	IPI00312067.4	VDQLNLER	498.77	2	62	10 of 12
20060405MM_31N	IPI00312067.4	VFDQQSSPMVTR	702.84	2	49	11 of 12
20060405MM_31N	IPI00312067.4	VLSGLEILSK	533.83	2	24	12 of 12
20060405MM_31X	IPI00312067.4	ASGLGEAGMGAWLR	688.34	2	60	1 of 8
20060405MM_31X	IPI00312067.4	DFIFVSAR	477.76	2	29	2 of 8
20060405MM_31X	IPI00312067.4	DFLSGIQK	454.25	2	29	3 of 8
20060405MM_31X	IPI00312067.4	GSYGLDLEAVR	590.31	2	54	4 of 8
20060405MM_31X	IPI00312067.4	SALLPNPLELQPPR	772.94	2	43	5 of 8
20060405MM_31X	IPI00312067.4	SFENDAQSSDNINFLK	914.92	2	84	6 of 8
20060405MM_31X	IPI00312067.4	VDQLNLER	493.77	2	65	7 of 8
20060405MM_31X	IPI00312067.4	VLSGLEILSK	529.83	2	39	8 of 8
20060405MM_31X	IPI00116554.2	ESQSHPGDFVLSVR	779.39	2	36	1 of 15
20060405MM_31X	IPI00116554.2	FATLAELVQYYMEHHGQLK	570.29	4	28	2 of 15
20060405MM_31X	IPI00116554.2	FDSLTDLVEHYK	489.58	3	33	3 of 15
20060405MM_31X	IPI00116554.2	FDSLTDLVEHYKK	532.28	3	18	4 of 15
20060405MM_31X	IPI00116554.2	HGSFLVR	408.23	2	46	5 of 15
20060405MM_31X	IPI00116554.2	INAAEIESR	501.77	2	49	6 of 15
20060405MM_31X	IPI00116554.2	KNPMVETLGTVLQLK	557.66	3	21	7 of 15
20060405MM_31X	IPI00116554.2	NILPFDHTR	556.8	2	29	8 of 15
20060405MM_31X	IPI00116554.2	QPLNTTR	415.23	2	27	9 of 15
20060405MM_31X	IPI00116554.2	SGMVQTEAQYR	635.3	2	55	10 of 15
20060405MM_31X	IPI00116554.2	SNPGDFTLSVR	596.8	2	34	11 of 15
20060405MM_31X	IPI00116554.2	TVWQYHFR	379.53	3	16	12 of 15
20060405MM_31X	IPI00116554.2	VGQALLQGNTER	643.35	2	28	13 of 15
20060405MM_31X	IPI00116554.2	VVLHDGDPNEPVSDYINANIIMP	1052.85	3	30	14 of 15
20060405MM_31X	IPI00116554.2	WFHPNITGVEAENLLLTR	704.04	3	33	15 of 15
20060405MM_35N	IPI00130621.3	ATTLASTLMEQYMK	798.4	2	73	1 of 16
20060405MM_35N	IPI00130621.3	EIYNTIR	459.75	2	27	2 of 16
20060405MM_35N	IPI00130621.3	EPVPMQDQGQVLNDTVDGK	1039.5	2	57	3 of 16
20060405MM_35N	IPI00130621.3	EPYMEGVNPFIK	724.35	2	23	4 of 16
20060405MM_35N	IPI00130621.3	FEITLSNK	480.27	2	32	5 of 16
20060405MM_35N	IPI00130621.3	GHATDEWFLLSSHIPLK	490.51	4	26	6 of 16
20060405MM_35N	IPI00130621.3	IFMASEILPPTLR	749.42	2	30	7 of 16
20060405MM_35N	IPI00130621.3	IMPEEEYSEFK	705.32	2	21	8 of 16
20060405MM_35N	IPI00130621.3	LLAITELLQQK	639.4	2	47	9 of 16
20060405MM_35N	IPI00130621.3	LLYPVAPPEPVEDR	802.93	2	50	10 of 16
20060405MM_35N	IPI00130621.3	MFNIISDSPSPIAAR	814.92	2	54	11 of 16

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

20060405MM_35N	IPI00130621.3	MIMFLDELGNVPELPDTTEHSR	852.08	3	51	12 of 16
20060405MM_35N	IPI00130621.3	NEDVNTNLAHLLSILSELVEK	787.09	3	31	13 of 16
20060405MM_35N	IPI00130621.3	SVQNLANLVEFGAK	749.41	2	69	14 of 16
20060405MM_35N	IPI00130621.3	VPDTDEISFLK	636.34	2	26	15 of 16
20060405MM_35N	IPI00130621.3	YYNSIGDIIDHYR	546.93	3	54	16 of 16
20060405MM_35X	IPI00130621.3	EISMEDEATTLFR	771.36	2	59	1 of 12
20060405MM_35X	IPI00130621.3	EPYMEGVNPFIK	720.35	2	25	2 of 12
20060405MM_35X	IPI00130621.3	EQIVEGYYLK	621.32	2	22	3 of 12
20060405MM_35X	IPI00130621.3	GHATDEWFLLSSHIPLK	651.01	3	33	4 of 12
20060405MM_35X	IPI00130621.3	LLAITELLQQK	635.4	2	55	5 of 12
20060405MM_35X	IPI00130621.3	LLYPVAPPEPVEDR	797.92	2	35	6 of 12
20060405MM_35X	IPI00130621.3	MFNIISDSPSPIAAR	809.92	2	56	7 of 12
20060405MM_35X	IPI00130621.3	MIMFLDELGNVPELPDTTEHSR	854.07	3	55	8 of 12
20060405MM_35X	IPI00130621.3	SKDPDILFMR	407.88	3	34	9 of 12
20060405MM_35X	IPI00130621.3	TLTLVAK	373.24	2	18	10 of 12
20060405MM_35X	IPI00130621.3	VPDTDEISFLK	632.33	2	27	11 of 12
20060405MM_35X	IPI00130621.3	VVSGFVFLR	512.3	2	70	12 of 12
20051002SH_IRS1_Y0608cı	IPI00263878	IMHNHDK	451.72	2	26	1 of 6
20051002SH_IRS1_Y0608cı	IPI00263878	RTAIEAFNETIK	705.89	2	63	2 of 6
20051002SH_IRS1_Y0608cı	IPI00263878	TATGYGFAEPYNLYSSLK	995.49	2	44	3 of 6
20051002SH_IRS1_Y0608cı	IPI00263878	TEADTEQQALPLPDLAEQFAPP	1112.59	3	22	4 of 6
20051002SH_IRS1_Y0608cı	IPI00263878	VLSEIFSPVLFR	703.91	2	53	5 of 6
20051002SH_IRS1_Y0608cı	IPI00263878	YGFSDPLTFNSVVELINHYR	794.4	3	70	6 of 6
20060324MM_11N	IPI00119058.1	ATADDELSFK	552.77	2	37	1 of 9
20060324MM_11N	IPI00119058.1	DGFIPK	342.69	2	26	2 of 9
20060324MM_11N	IPI00119058.1	ESESAPGDFSLSVK	726.85	2	24	3 of 9
20060324MM_11N	IPI00119058.1	FNSLNELVDYHR	758.87	2	59	4 of 9
20060324MM_11N	IPI00119058.1	GDFIHVMDNSDPNWWK	656.97	3	20	5 of 9
20060324MM_11N	IPI00119058.1	NQQIFLR	464.77	2	26	6 of 9
20060324MM_11N	IPI00119058.1	PHPWFFGK	341.85	3	22	7 of 9
20060324MM_11N	IPI00119058.1	QRHDGAFLIR	411.56	3	21	8 of 9
20060324MM_11N	IPI00119058.1	YFLWVVK	481.78	2	28	9 of 9
20060324MM_11X	IPI00119058.1	ATADDELSFK	548.76	2	50	1 of 11
20060324MM_11X	IPI00119058.1	DGFIPK	338.69	2	25	2 of 11
20060324MM_11X	IPI00119058.1	DIEQMPQQPTYVQALFDFDPQE	1190.88	3	42	3 of 11
20060324MM_11X	IPI00119058.1	ESESAPGDFSLSVK	726.85	2	23	4 of 11
20060324MM_11X	IPI00119058.1	FGNDVQHFK	546.27	2	69	5 of 11
20060324MM_11X	IPI00119058.1	FNSLNELVDYHR	753.87	2	65	6 of 11
20060324MM_11X	IPI00119058.1	NQQIFLR	459.76	2	32	7 of 11
20060324MM_11X	IPI00119058.1	NYIEMKPHPWFFGK	449.23	4	28	8 of 11
20060324MM_11X	IPI00119058.1	PHPWFFGK	339.18	3	25	9 of 11
20060324MM_11X	IPI00119058.1	RGDFIHVMDNSDPNWWK	706.33	3	40	10 of 11
20060324MM_11X	IPI00119058.1	YFLWVVK	481.78	2	39	11 of 11
20060607MM_49N	IPI00312067.4	DFIFVSAR	482.76	2	31	1 of 12
20060607MM_49N	IPI00312067.4	DSQEDWTTFTHDR	824.35	2	46	2 of 12
20060607MM_49N	IPI00312067.4	GSYGLDLEAVR	595.31	2	59	3 of 12
20060607MM_49N	IPI00312067.4	LLQQQSLPQTGEQELESLVLK	797.11	3	69	4 of 12
20060607MM_49N	IPI00312067.4	LYEWISIDK	587.82	2	23	5 of 12
20060607MM_49N	IPI00312067.4	NQNYLDILR	579.81	2	15	6 of 12
20060607MM_49N	IPI00312067.4	NSFNNPAYYVLEGVPHQLLPLE	1049.22	3	38	7 of 12
20060607MM_49N	IPI00312067.4	SALLPNPLELQPPR	777.95	2	27	8 of 12
20060607MM_49N	IPI00312067.4	SFENDAQSSDNINFLK	918.93	2	86	9 of 12
20060607MM_49N	IPI00312067.4	SMIGSTAQQFLTFLSHR	645.33	3	33	10 of 12
20060607MM_49N	IPI00312067.4	TLSEVDYAPGPGR	686.34	2	74	11 of 12
20060607MM_49N	IPI00312067.4	VLSGLEILSK	533.83	2	58	12 of 12
20060607MM_49N	IPI00125298.2	AAEDGEGSAGAALR	642.81	2	108	1 of 15
20060607MM_49N	IPI00125298.2	DLFDMK	388.69	2	25	2 of 15
20060607MM_49N	IPI00125298.2	DLFDMKPFEDALR	538.94	3	28	3 of 15
20060607MM_49N	IPI00125298.2	DSGLLPLLQDMNK	726.39	2	56	4 of 15
20060607MM_49N	IPI00125298.2	EAEALLQLNGDFLVR	849.46	2	62	5 of 15
20060607MM_49N	IPI00125298.2	ELFDDPSYVNIQNLDK	959.47	2	80	6 of 15
20060607MM_49N	IPI00125298.2	EPPLGGVVDMR	590.31	2	47	7 of 15
20060607MM_49N	IPI00125298.2	ESTTTPGQYVLTGLQSGQPK	1050.54	2	57	8 of 15
20060607MM_49N	IPI00125298.2	HLLLVDPEGVVR	678.9	2	52	9 of 15
20060607MM_49N	IPI00125298.2	MAGFDGSAWDEEEEEPPDHQY	1090.11	3	59	10 of 15
20060607MM_49N	IPI00125298.2	MDLLPPKPK	548.82	2	15	11 of 15
20060607MM_49N	IPI00125298.2	PLSSILGR	426.76	2	57	12 of 15
20060607MM_49N	IPI00125298.2	QIIANHHMQSISFASGGDPDTAE	1071.18	3	57	13 of 15
20060607MM_49N	IPI00125298.2	VEGGQLGGEEWTR	714.34	2	93	14 of 15
20060607MM_49N	IPI00125298.2	VPPPPQSMSMAEQLQGEPWFF	647.31	4	36	15 of 15
20060607MM_49X	IPI00125298.2	AAEDGEGSAGAALR	637.8	2	68	1 of 19
20060607MM_49X	IPI00125298.2	DLFDMK	384.68	2	28	2 of 19
20060607MM_49X	IPI00125298.2	DLFDMKPFEDALR	532.93	3	43	3 of 19
20060607MM_49X	IPI00125298.2	DSGLLPLLQDMNK	722.38	2	67	4 of 19

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

20060607MM_49X	IPI00125298.2	EAEALLQLNGDFLVR	844.46	2	65	5 of 19	
20060607MM_49X	IPI00125298.2	ELFDDPSYVNIQNLDK	955.46	2	67	6 of 19	
20060607MM_49X	IPI00125298.2	EPPLGGVVDMR	585.3	2	62	7 of 19	
20060607MM_49X	IPI00125298.2	ESTTTPGQYVLTGLQSGQPK	1046.53	2	79	8 of 19	
20060607MM_49X	IPI00125298.2	GWLHPNDK	483.74	2	25	9 of 19	
20060607MM_49X	IPI00125298.2	HGSFVNKPTR	571.81	2	51	10 of 19	
20060607MM_49X	IPI00125298.2	HLLLVDPEGVVR	673.89	2	38	11 of 19	
20060607MM_49X	IPI00125298.2	MDLLPPKPK	540.81	2	16	12 of 19	
20060607MM_49X	IPI00125298.2	PFEDALR	424.22	2	37	13 of 19	
20060607MM_49X	IPI00125298.2	QAGGGAGPPNPSLNGSAPR	852.92	2	94	14 of 19	
20060607MM_49X	IPI00125298.2	QIIANHHMQSISFASGGDPDTAE	1068.51	3	55	15 of 19	
20060607MM_49X	IPI00125298.2	TRVEGGQLGGEEWTR	837.91	2	45	16 of 19	
20060607MM_49X	IPI00125298.2	VEGGQLGGEEWTR	709.34	2	105	17 of 19	
20060607MM_49X	IPI00125298.2	VMGPGVSYLVR	589.32	2	66	18 of 19	
20060607MM_49X	IPI00125298.2	VPPPPQSMSMAEQLQGEPWFH	645.31	4	37	19 of 19	
20060607MM_53N	IPI00312067.4	GEETGNIR	443.22	2	34	1 of 10	
20060607MM_53N	IPI00312067.4	GSYGLDLEAVR	595.31	2	66	2 of 10	
20060607MM_53N	IPI00312067.4	LDVTLGDLTK	541.81	2	48	3 of 10	
20060607MM_53N	IPI00312067.4	LLQQQSLPQTGEQELESLVLK	797.11	3	59	4 of 10	
20060607MM_53N	IPI00312067.4	LYEWISIDK	587.82	2	37	5 of 10	
20060607MM_53N	IPI00312067.4	NSFNNPAYYVLEGVPHQLLPLE	1049.21	3	53	6 of 10	
20060607MM_53N	IPI00312067.4	NVTSWFTSK	539.28	2	24	7 of 10	
20060607MM_53N	IPI00312067.4	SFENDAQSSDNINFLK	918.93	2	95	8 of 10	
20060607MM_53N	IPI00312067.4	SMIGSTAQQFLTFLSHR	645.33	3	22	9 of 10	
20060607MM_53N	IPI00312067.4	VAITVPAPQLGR	616.37	2	15	10 of 10	
20060607MM_53X	IPI00312067.4	ASGLGEAGMGAWLR	688.35	2	31	1 of 8	
20060607MM_53X	IPI00312067.4	GSYGLDLEAVR	590.3	2	53	2 of 8	
20060607MM_53X	IPI00312067.4	LLQQQSLPQTGEQELESLVLK	794.44	3	41	3 of 8	
20060607MM_53X	IPI00312067.4	LYEWISIDK	583.81	2	37	4 of 8	
20060607MM_53X	IPI00312067.4	SMIGSTAQQFLTFLSHR	642	3	22	5 of 8	
20060607MM_53X	IPI00312067.4	TLSEVDYAPGPGR	681.34	2	24	6 of 8	
20060607MM_53X	IPI00312067.4	VDQLNLER	493.77	2	47	7 of 8	
20060607MM_53X	IPI00312067.4	VLSGLEILSK	529.83	2	52	8 of 8	
20060607MM_55N	IPI00312067.4	LLQQQSLPQTGEQELESLVLK	797.11	3	46	1 of 1	1
20060607MM_55X	IPI00312067.4	ELTDLDYRPIAMQSLWNIK	769.4	3	24	1 of 6	
20060607MM_55X	IPI00312067.4	GSYGLDLEAVR	590.3	2	57	2 of 6	
20060607MM_55X	IPI00312067.4	LLQQQSLPQTGEQELESLVLK	794.44	3	42	3 of 6	
20060607MM_55X	IPI00312067.4	LYEWISIDK	583.81	2	34	4 of 6	
20060607MM_55X	IPI00312067.4	TLSEVDYAPGPGR	681.34	2	36	5 of 6	
20060607MM_55X	IPI00312067.4	VLSGLEILSK	529.83	2	70	6 of 6	
20061102_SH_catenin	IPI00312067.4	ELTDLDYRPIAMQSLWNIK	780.74	3	22	1 of 7	
20061102_SH_catenin	IPI00312067.4	GSYGLDLEAVR	595.31	2	64	2 of 7	
20061102_SH_catenin	IPI00312067.4	LDVTLGDLTK	541.81	2	36	3 of 7	
20061102_SH_catenin	IPI00312067.4	LGVVFEK	400.24	2	27	4 of 7	
20061102_SH_catenin	IPI00312067.4	LLQQQSLPQTGEQELESLVLK	797.11	3	42	5 of 7	
20061102_SH_catenin	IPI00312067.4	SMIGSTAQQFLTFLSHR	650.66	3	64	6 of 7	
20061102_SH_catenin	IPI00312067.4	VLSGLEILSK	533.83	2	35	7 of 7	
20061102_SH_catenin	IPI00116554.2	SNPGDFTLSVR	601.81	2	31	1 of 2	
20061102_SH_catenin	IPI00116554.2	WFHPNITGVEAENLLLTR	707.38	3	41	2 of 2	
20061102_SH_catenincross	IPI00312067.4	ELTDLDYRPIAMQSLWNIK	774.73	3	32	1 of 8	
20061102_SH_catenincross	IPI00312067.4	GSYGLDLEAVR	590.3	2	38	2 of 8	
20061102_SH_catenincross	IPI00312067.4	LGVVFEK	396.24	2	31	3 of 8	
20061102_SH_catenincross	IPI00312067.4	LLQQQSLPQTGEQELESLVLK	794.44	3	60	4 of 8	
20061102_SH_catenincross	IPI00312067.4	LVLLR	307.22	2	16	5 of 8	
20061102_SH_catenincross	IPI00312067.4	LYEWISIDK	583.81	2	30	6 of 8	
20061102_SH_catenincross	IPI00312067.4	SMIGSTAQQFLTFLSHR	647.33	3	49	7 of 8	
20061102_SH_catenincross	IPI00312067.4	VLSGLEILSK	529.83	2	58	8 of 8	
20061102_SH_catenincross	IPI00309413.4	LWLLDDSK	495.27	2	29	1 of 2	
20061102_SH_catenincross	IPI00309413.4	VLHVVQTLYPFSSVTEEELSFEK	894.46	3	16	2 of 2	
20061102_SH_MM78	IPI00312067.4	ELTDLDYRPIAMQSLWNIK	775.41	3	36	1 of 2	
20061102_SH_MM78	IPI00312067.4	VLSGLEILSK	533.83	2	56	2 of 2	
20061102_SH_MM78	IPI00309413.4	LWLLDDSK	499.28	2	31	1 of 1	1
20061102_SH_MM78cross	IPI00309413.4	FAYVAEREDELSLVK	590.31	3	21	1 of 4	
20061102_SH_MM78cross	IPI00309413.4	IYDLNIPAFVK	646.87	2	39	2 of 4	
20061102_SH_MM78cross	IPI00309413.4	LWLLDDSK	495.27	2	42	3 of 4	
20061102_SH_MM78cross	IPI00309413.4	VLHVVQTLYPFSSVTEEELSFEK	894.46	3	19	4 of 4	
20061102_SH_MM78cross	IPI00312067.4	ELTDLDYRPIAMQSLWNIK	774.73	3	23	1 of 2	
20061102_SH_MM78cross	IPI00312067.4	VLSGLEILSK	529.83	2	44	2 of 2	
20061102_SH_wasp	IPI00312067.4	LLQQQSLPQTGEQELESLVLK	797.11	3	34	1 of 1	1
20061102_SH_wasp	IPI00309413.4	LWLLDDSK	499.28	2	31	1 of 2	
20061102_SH_wasp	IPI00309413.4	VLHVVQTLYPFSSVTEEELSFEK	897.13	3	24	2 of 2	
20061102_SH_waspcross	IPI00312067.4	ELTDLDYRPIAMQSLWNIK	774.73	3	17	1 of 3	
20061102_SH_waspcross	IPI00312067.4	SMIGSTAQQFLTFLSHR	647.33	3	29	2 of 3	

Downloaded from www.mcponline.org by on January 9, 2009

Peptides

20061102_SH_waspcross	IPI00312067.4	VLSGLEILSK	529.83	2	46	3 of 3
20061102_SH_waspcross	IPI00309413.4	FAYVAEREDELSLVK	590.31	3	19	1 of 2
20061102_SH_waspcross	IPI00309413.4	LWLLDDSK	495.27	2	32	2 of 2

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009

		7 of 14	4 of 6	(pY[ED]x[IV])		TWF2_MOUSE, 309	
18	32	Stat SH2 15 of 19	pY[ILMV]PQ 4 of 4	STAT SH2 (pYxxQ)	14966128	SGENFVPpYMPQFQTC LEPR_MOUSE, 1138	none
19	33		H[ST]GxKPpYxCxxCG 10 of 10	new motif		RIHTGEKPpYECVQCGK ZNF24_MOUSE, 335	none
20	36		[AG]xpYxx[ILMV]x[KR] 8 of 15	new motif		KKNRIAIpYELLFKEG RS10_MOUSE, 12	none

Molecular & Cellular Proteomics
ASBMB
Downloaded from www.mcponline.org by on January 9, 2009