A novel proteomic screen for peptide-protein interactions

Waltraud X. Schulze and Matthias Mann*

Center for Experimental BioInformatics (CEBI), Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark

Correspondence should be addressed to MM (Email: mann@bmb.sdu.dk; Tel: +45 65502364; Fax: +45 65933929)

Manuscript information: 5 figures, 1 table

Keywords: protein-protein interaction, mass spectrometry, phosphorylation, proline-rich motif, EGF signaling, Sos, Grb2, quantitative proteomics

Running head: peptide-protein interaction screen

Abbreviations: $^{12}C_6$-Arg – $^{12}C_6$ or ‘normal’ arginine, $^{13}C_6$-Arg – $^{13}C_6$ arginine, EGFR Epidermal Growth Factor Receptor, Grb2- Growth factor associated protein 2, LeuD3 – L-leucine-5,5,5-d3, LC/MS – high performance liquid chromatography/mass spectrometry, Pacsin – Protein kinase C and casein kinase substrate in neurons, SH2 – Src homology domain 2, SH3 – Src homology domain 3, SILAC – stable isotope labeling with amino acids in cell culture, Snx9 – Sorting nexin 9, Sos – son of sevenless

Copyright 2003 by The American Society for Biochemistry and Molecular Biology, Inc.
SUMMARY

Regulated interactions between short, unstructured amino acid sequences and modular protein domains are central to cell signaling. Here we use synthetic peptides in 'active' (e.g. phosphorylated) and 'control' (e.g. non-phosphorylated) forms as baits in affinity pull-down experiments to determine such interactions by quantitative proteomics. Stable isotope labeling by amino acids in cell culture (SILAC) distinguishes specific binders directly by the isotope ratios determined by mass spectrometry (Blagoev, et al Nature Biotech, 2003, 21, 315-318). A tyrosine-phosphorylated peptide of the Epidermal Growth Factor Receptor (EGFR) specifically retrieved the SH2- and SH3 domain-containing adapter protein Grb2. A proline-rich sequence of Son of Sevenless also specifically bound Grb2, demonstrating that the screen maintains specificity with low affinity interactions. The proline-rich Sos peptide retrieved only SH3 domain containing proteins as specific binding partners. Two of these, Pacsin 3, and Sorting Nexin 9, were confirmed by immunoprecipitation. Our data are consistent with a change in the role of Sos from Ras-dependent signaling to actin remodeling/endocytic signaling events by a proline - SH3 domain switch.
INTRODUCTION

Important protein-protein interactions in cell signaling are frequently mediated by short, unstructured sequences, which specifically interact with peptide binding domains (1). Typical examples are the binding of tyrosyl-phosphorylated peptides to proteins containing Src homology domain 2 (SH2) or phosphotyrosyl binding domain (PTB) domain (2,3). Peptides with certain proline motifs constitutively bind to proteins containing Src homology domain 3 (SH3) at lower affinities (4,5). Despite the central importance of motif-dependent interactions in cell signaling, there are few methods to identify peptide-binding partners. Phage display (6), array technologies (7), and peptide libraries (8) are powerful strategies to learn about general features of optimal peptide sequences binding to a given domain or protein. Information from these experiments has in turn been helpful in the development of bioinformatic algorithms to predict occurrence of these sequences in proteins (9). However, these methods cannot efficiently identify specific in vivo interaction partners of a given modified peptide sequence. The yeast two-hybrid system (10) has been successfully applied to large scale protein-protein interaction mapping (11-13) but it is not well suited to determine the interaction partners of short and modified peptide sequences. In addition, the yeast two-hybrid system is a heterologous system for studying interactions of mammalian or plant proteins.

The most straightforward method for determining interaction partners of a peptide is to use it as bait in affinity pull-down experiments followed by direct detection of binding proteins. General use of this approach has been limited by two factors: Firstly, the affinity of a given binding domain to the peptide bait is often weak, and can further be reduced since the peptide bait is used outside of the context of the whole protein. Secondly, binding partners of interest in signal transduction are typically of low abundance. Thus, small amounts of protein binding specifically to the bait are masked by more abundant, but non-specific binders to both the peptides and the matrix employed in the pull-down. This has often prevented unbiased methods, such as peptide sequencing, to be employed for the identification of interacting proteins and synthetic peptides are usually only
employed in the verification of postulated protein-protein interactions by competitive disruption of binding.

Recently our group described a proteomic method to detect signaling complexes formed upon cell stimulation (14). Using Stable Isotope Labeling by Amino Acids in Cell culture, SILAC (15,16), all the proteins in one protein population were metabolically labeled with a heavy isotope- form of Arginine ($^{13}$C$_6$ Arg, leading to a difference of six mass units). In this way, two forms of each arginine-containing peptide are measured in subsequent mass spectrometric experiments. The two forms differ only in their mass and can therefore be used to directly quantify the ratio of the proteins from one population in relation to the other. Using the SH2 domain of Grb2 as an affinity probe, tyrosine phosphorylated EGF receptor and associated proteins were purified from labeled, stimulated cells and non-labeled, non-stimulated cells. After trypsin digestion of eluted proteins, arginine-containing peptides were observed as light (the normal $^{12}$C$_6$ form of arginine) and heavy ($^{13}$C$_6$-arginine) forms in mass spectra. One-to-one ratios between the isotope forms indicated proteins unaffected by EGF stimulation, whereas ratios greater than one indicated proteins recruited to the signaling complex upon EGF stimulation. A related quantitative proteomics strategy has independently been described (17), where ICAT technology (18) was used to distinguish two cellular states, leading to the identification of specific members of the yeast polymerase II complex in a large background of non-specific proteins.

Our previous work allowed characterization of multi-protein complexes formed as a result of cell stimulation. Here, we employ the principle of SILAC in affinity pull-down experiments for a different purpose; the determination of the interaction partners of short unstructured and potentially modified amino acid sequences, which requires a completely different strategy. We find that synthetic peptides can efficiently be used as baits in such experiments. In contrast to the previous approach, the current screen is designed to primarily reveal direct binders, rather than multiprotein complexes. Furthermore, cloning and expression of domains, or prior knowledge of the signaling events leading to the interaction are not required, more easily adaptable to a large-scale approach. The strategy was
validated in a classical part of growth factor signaling, namely in the interactions of the adaptor protein Grb2. Its role in EGF signaling is to bring the guanine exchange factor Sos to the activated receptor, so that Sos can initiate signaling through Ras at the plasma membrane. We selected a phospho-tyrosine containing peptide of EGFR, which is known to bind the SH2 domain of Grb2, and proline-rich motif containing peptides of Sos1 and Sos2 that are known to bind the SH3 domains of Grb2. Both proline-rich Sos peptides retrieved only proteins containing SH3 domains as specific binding partners, demonstrating exquisite specificity in the affinity pull-downs from whole cell lysates. Two of the identified proteins binding to the proline-rich Sos peptide represent novel interactions, and were confirmed by co-immunoprecipitation. These proteins, Pacsin3 and Snx9, also have potential roles in growth factor signaling.

**EXPERIMENTAL PROCEDURES**

**Metabolic labeling in cell culture**

Human HeLa cells were grown in arginine-deficient Dulbecco's Modified Eagle Medium (DMEM) with 10% dialyzed fetal bovine serum. One cell population was supplemented with normal isotopic abundance L-arginine (Sigma), and the other with 99% isotopic abundance $^{13}$C$_6$-arginine (Aldrich) as described (16). Each population was grown for at least a minimum of five population doublings. Four plates (14 cm) of each cell population were used per experiment. In some experiments, L-leucine-3,3,3-D$_3$ (Aldrich) and Leu were used for metabolic stable isotope labeling in leucine-deficient DMEM supplemented with 10% fetal bovine serum (15).

**Peptide synthesis and pull-downs**

Biotinylated peptides were synthesized on a solid-phase peptide synthesizer using amide resin (Intavis, Germany). Peptides were designed as 15-mers bearing a N-terminal biotin on a tetrapeptide linker SGSG (19). Biotinylation was done using Biotinylloxy-succimide (NovaBiochem) or Sulfo-NHS-SS-Biotin (Pierce), the latter allowing reductive cleavage of the 15-mer from the biotin group. The identity and purity of the synthesized peptides was confirmed by
mass spectrometric analysis. Peptides were synthesized as pairs as 'active' and 'control' forms. For affinity pull-downs, 30 nmol of immobilized peptide was added to approximately 6 mg of cell lysate. Under such conditions, sub-picomole amounts of specific binding proteins were detected during mass spectrometric analysis, indicating a 1000-fold excess of bait molecules over endogenous binding proteins. Thus, no competition for binding sites is expected.

Immobilized streptavidin beads (Pierce) were loaded with biotinylated peptide prior to incubation with cell lysates. Cells were lysed in 1% (v/v) Nonident P-40, 150 mM sodium chloride, 50 mM Tris-HCl pH 7.5, protease inhibitors (Complete Tablets, Roche), and 1 mM sodium orthovanadate as phosphatase inhibitor. Equal amount of protein was incubated with the respective immobilized peptides at 4°C for 6 h. After extensive washes, bound proteins were either eluted from the immobilized peptides by boiling in SDS sample buffer. Alternatively, the bait peptide with its bound proteins was cleaved off the beads using 50 mM dithiothreitol. Eluates from the 'active' and 'control' bait peptide pull-downs were combined for further analysis.

**Immunoprecipitation**

HeLa cells were lysed in 1% (v/v) Nonident P-40, 150 mM sodium chloride, 50 mM Tris-HCl pH 7.5, 1 mM sodium orthovanadate and protease inhibitors (Complete Tablets, Roche), and incubated at 4°C with a rabbit antibody against Sos isoforms (Santa Cruz). Protein-A-separose beads were added after 4 h and incubated for additional 4h. After extensive washes, co-precipitated proteins were eluted in sample buffer for SDS gel electrophoresis.

**LC/MS/MS, database searching and quantitation**

Combined eluted proteins were resolved by SDS-PAGE (15% w/v). After silver staining, the entire lane was excised in up to six equally spaced slices and digested enzymatically with trypsin prior to LC-MS analysis. Tryptic peptide mixtures were then desalted (20) and loaded onto reversed phase analytical columns for liquid chromatography (21). Peptides were eluted from the analytical column by a multistep linear gradient running from 5% to 60% acetonitrile in 60 minutes and sprayed directly into the orifice of a QSTAR-Pulsar quadrupole
Time-Of-Flight hybrid mass spectrometer (PE-Sciex, USA). Proteins were identified by high mass accuracy tandem mass spectrometry (LC MS/MS) by information-dependent acquisition of fragmentation spectra of doubly, triply or quadruply charged peptides. Acquired spectra were then searched against the human International Protein Index Database (www.ebi.ac.uk/IPI) using the Mascot algorithm (22) with search parameters as described (23). All protein identifications were manually verified against the raw mass spectrometric data using in-house software.

For quantitation, ratios between the centroids of the MS peaks of 'heavy' and 'light' forms of the peptide were calculated and averaged over consecutive MS cycles for the duration of the respective MS peaks in the total ion chromatogram (16). Heavy and light forms of peptides co-elute for 13C based labels and the quantification is based on the average of a number of independently determined ratios for each peptide as the peak elutes from the chromatography column. Labeled and non-labeled peptide pairs were identified by their charge state and mass difference, for example doubly charged peptide pairs were identified through a difference in mass-to-charge ratio of three. In-house developed software allowed automation of the extraction of corresponding MS peaks. Each ratio was subsequently calculated from the extracted ion chromatograms after manually verifying that the MS spectra containing the respective peaks were at a sufficient level above background and separated from interfering peaks of other peptides. Ratios obtained from different peptides identifying the same protein were averaged. A final standard deviation for protein ratios was calculated from the ratios of the individual peptides or, when the identification was based on a single peptide, from the ratios obtained from the different mass spectra of this peptide pair.

We have employed custom made software for our instrument. The quantification task is very similar to those encountered in the popular ICAT method (18) or other quantitative proteomics schemes. Our peptide hit verification and quantitation software will be made open-source at Sourceforge (http://sourceforge.net). Furthermore, while SILAC has certain advantages in this
screen, it would also be possible to perform it with other quantitative proteomic approaches, for which software is already available.

**RESULTS**

**The proteomic peptide-protein interaction screen**

First, two cell populations were metabolically labeled with normal arginine ($^{12}\text{C}_6$-Arg) and arginine in which the carbon atoms were replaced by the stable $^{13}\text{C}$ isotope of carbon ($^{13}\text{C}_6$-Arg) (Fig. 1). HeLa cells were adapted to the medium for a minimum of five cell doublings to ensure complete labeling of even long-lived proteins. This ‘SILAC’ labeling (15,16) makes arginine-containing peptides distinguishable between the two populations without introducing any other chemical or biological differences. In particular, both forms of each peptide are detected with equal efficiency by the mass spectrometer so that the measured ratio between the two peptides accurately reflects the ratio of the proteins in the two populations.

Cell lysates from the $^{12}\text{C}_6$-Arg-labeled and the $^{13}\text{C}_6$-Arg-labeled cells were employed in affinity purification using synthetic peptides, representing the non-phosphorylated and the phosphorylated form of a peptide, respectively. The unphosphorylated peptide was exposed to the $^{12}\text{C}_6$-Arg-encoded cell lysate and the phosphorylated peptide was exposed to the $^{13}\text{C}_6$-Arg-encoded lysate. Lysates were then mixed and analyzed together (Fig. 1). Background proteins – for example ribosomal proteins binding to the beads to which the peptides are attached – are present in equal amounts in both pull-downs experiments. Upon analysis by mass spectrometry, peptides from such proteins will appear in the two forms with equal abundance. If a protein binds specifically to the phosphopeptide and not to the non-phosphorylated peptide then it will only be represented in the $^{13}\text{C}_6$-labeled cell lysate. Consequently, in this case we will observe only the higher molecular weight form of the peptide. More generally, if a protein binds specifically but not exclusively to one of the two forms of the bait then this will lead to a ratio between the pair of peaks that is related to the specificity of the binding to one form over the other. While the above discussion
assumes phosphorylation, any other modifications that can be synthesized can be investigated for protein binding as well. Interactions of unmodified peptides can, furthermore, be studied by using mutated peptides as the control.

**Phosphopeptide of EGFR as bait**

For pull-down experiments according to the strategy outlined in Figure 1, a bait peptide was chosen flanking tyrosine 1068 of EGF Receptor, which is known to bind the adapter protein Grb2 upon auto-phosphorylation of the tyrosine residue (19,24,25). The peptide biotinSS-SGSGLPVPETpYINQSV, containing a cleavable biotin (biotinSS) for coupling to streptavidin coated beads and a spacer SGSG (19), was synthesized in the phosphotyrosine (pY) and tyrosine form. Lysates of \(^{13}\)C\(_6\)-Arg-labeled HeLa cells were incubated with the phosphorylated peptide while lysates of normal amino acid containing HeLa cells (\(^{12}\)C\(_6\)-Arg labeled) were incubated with the unphosphorylated control peptide. Eluted proteins were combined, digested to peptides, and analyzed by mass spectrometry as described in Methods.

The phosphotyrosine-1068 peptide bait removed Grb2 from the cell lysate (Fig.2) indicating a sufficient excess of bait peptide. Out of 148 proteins identified, only Grb2 displayed pairs of tryptic peptides with a significant ratio as expected for those proteins specifically binding to the phosphopeptide (Tab. 1). In fact, only the \(^{13}\)C\(_6\)-labeled forms of the Grb2 peptides were detected above noise levels (Fig. 3A), indicating a ratio of more than 10 between labeled and unlabeled forms. In contrast, nonspecifically binding proteins, such as ribosomal proteins, were detected in labeled and unlabeled form with ratios close to unity. In addition, they were present in amounts that led to ten fold higher ion currents during mass spectrometry (Fig. 3B). This suggests that it would have been very difficult to identify the low-abundant specific binder in a conventional affinity pull-down experiment through staining of proteins after gel electrophoresis. To confirm the result, a separate experiment was performed using deuterated leucine instead of arginine for metabolic labeling and, again, Grb2 was detected as the only specific binder (Tab. 1).
To eliminate any potential variability in the binding of background proteins, ‘cross over’ experiments were performed according to the strategy described in Fig. 1. In these experiments, the ‘active’ peptide is incubated with the cell lysate of the light amino acid containing proteins whereas the ‘control’ peptide is incubated with the heavy amino acid containing proteins. Therefore, proteins that specifically bind to the bait should have $^{13}C_6$ to $^{12}C_6$ ratios equal to the $^{12}C_6$ to $^{13}C_6$ ratios in the cross over experiment (Fig. 3C). There were only 19 such proteins since most of the background proteins were identified only in one of the two experiments. Grb2 was detected with a ratio of greater than 10 in both cases and is therefore confirmed as a specific binder to the phosphopeptide. Some other proteins had large ratios in one, but not the other experiment. Especially cytoskeletal proteins show these variations, possibly due to these proteins being especially susceptible to minor variations in lysis conditions. We have found that the requirement for significant, compatible ratios in the cross over experiment is a stringent criterion for specific binding and removed other proteins in all cases.

Specificity of phosphotyrosine binding to SH2 domains is mainly determined by the amino acids immediately C-terminal to the phospho-tyrosine, and it is known that exchange of the amino acid in the +2 position of the EGFR peptide used here is sufficient to abolish binding (26). As a further test for the specificity of our peptide-protein binding assay we synthesized a peptide where the +2 residue was mutated to an alanine (biotinSS-SGSGLPVPEpYAIAQSV). This single amino acid change led to the complete absence of Grb2 in the list of identified proteins and none of the identified proteins had an isotopic ratio indicative of a specific binder (Fig. 3D).

We cannot exclude that other specific binding proteins exist that were present at amounts below detection limit. However, the fact that Grb2 was quantified in multiple experiments (Tab. 1), with several different peptides and with a signal-to-noise ratio of more than ten to one, argues that such proteins would have had to be present at much lower abundance than Grb2.
Proline-rich peptides of Sos1 and Sos2 as bait

In addition to its SH2 domain, Grb2 has two SH3 domains that can bind proline-rich regions of the Ras guanine exchange factor Sos. In a second experiment, the specificity of the peptide-protein screen for lower affinity interactions was tested using one of the four proline-rich motifs of the two isoforms Sos1 and Sos2. The peptides representing amino acids 1150 to 1158 of Sos1 (biotinSS-SSGSGVPVPVR) and amino acids 1146 to 1154 of Sos2 (biotinSS-SSGSGIPPLPPPR) contain the classical poly-proline type II motif PxΦPxR (Φ: hydrophobic residue) for SH3 domain recognition (27). Both peptides bind Grb2 with an affinity of about 4 µM, which is approximately six fold weaker than the interaction of the SH2 domain of Grb2 with the phosphotyrosine peptide used above (25,27,28). To obtain control baits according to the strategy presented in Fig. 1, prolines were replaced with alanines to prevent proline-directed folding (29).

Both proline-rich peptides specifically enriched a small number of proteins out of more than 100 identified proteins. Grb2 was clearly identified exclusively in the labeled form, indicating that the screen performs well with motifs typical of SH3 domain interactions (Fig. 4A). In addition, Pacsin3 and Snx9, which also contain SH3 domains, were detected with high ratios of labeled to unlabeled tryptic peptides (Fig. 4B, Tab. 1).

The identification of the interacting proteins was confirmed in independent ‘cross-over’ experiments, in which the bait peptides were incubated with unlabeled cell lysate and the control peptides with labeled cell lysate (Fig. 1). Grb2 and Pacsin3 were identified with similar ratios (>10) in all four experiments. In contrast, Snx9 was consistently identified with a lower ratio (Fig. 4B,C,D; Table 1). Proteins with ratios less than two were not considered to be specific.

The bait peptides of Sos1 and Sos2 differ by a conservative substitution of two hydrophobic residues and two consecutive arginines by lysines. Sos1 and Sos2 have been reported to display differences in binding of Grb2 (28), and to have distinct roles during signaling events (30). Our data only indicate that the proline-rich motifs examined here both specifically bind Grb2 and certain other SH3


domain containing proteins, like Pacsin3 and Snx9. Differences in binding partners of the full-length proteins may be caused by residues outside the core binding motif as well as through interactions with other regions of Sos.

To biologically validate the peptide-protein interactions, we performed co-immunoprecipitation experiments with an antibody against Sos, which recognizes both isoforms. Grb2 was detected in the co-immunoprecipitation by western blot and mass spectrometric analysis (Fig. 5A). The mass spectrometric experiment also identified Pacsin3 and Snx9 (Fig. 5B,C), as well as the bait protein, Sos. We therefore conclude that both the proline-rich peptides examined here, as well as the full length Sos protein can bind these three proteins in vivo.

**DISCUSSION**

We present a novel proteomic screen to find binding partners to short peptide stretches comprising typical sites of interaction with peptide binding domains. The screen uses synthetic peptides which can be custom made in modified and non-modified forms, and takes advantage of differential metabolic labeling of cell populations using stable isotopes, which allows distinguishing the proteins by mass spectrometry. Specific binding partners are apparent by their isotopic ratios when comparing mass spectra of tryptic peptides from proteins bound to 'active' versus 'control' peptides. This approach can detect modification-mediated protein-protein interactions as demonstrated by a phosphorylated synthetic peptide representing an SH2-domain binding site of the EGF-receptor. Furthermore, the peptide-protein interaction screen works well for low-affinity interactions, as shown with proline-rich peptides, which retrieved known SH3-domain containing binding partners.

In this work we have been especially concerned about demonstrating specificity and reproducibility of the novel method. In summary, we have described six lines of evidence: (1) Determination of the ratio between labeled and unlabeled peptide for several – not only one - mass spectra across the peptide peak as it elutes from the liquid chromatography column. The requirement here is that the peptide ratios must be consistent over the peak. (2) If the differential ratio of a protein is determined from several peptides, then these peptide ratios must be
consistent. (3) The ‘cross-over’ experiment, in which we expose the peptide to the unlabeled lysate and the control peptide to the labeled lysate must be consistent with the original experiment. (4) We have used the phosphotyrosine peptide with a single point mutation to alanine at the +2 position relative to pY. This experiment specifically abrogated any detectable binding to Grb2. Importantly, this ‘non-natural’ peptide had no binding partners with significant ratio at all. (5) We have performed the screen with the proline rich peptide from Sos for both the Sos1 and the Sos2 isoform. Both experiments retrieved the same significant novel interactors. (6) The interaction with Pacsin3 was independently verified by immuno-precipitation of Sos.

**Novel interactions with Sos proline-rich peptides**

Besides validating the method, the results using bait peptides binding the adapter protein Grb2 have biological interest. Grb2 is known as an adapter protein that recruits Sos to the plasma membrane in response to growth factor stimulation. This then allows Sos to activate the small GTPase Ras, which in turn activates the MAP kinase pathway (31,32). In this report, two novel interaction partners of the proline-rich region of Sos were identified, both of whom are involved in growth factor receptor endocytosis (33,34). Since all the proteins identified in the screen bind to the same proline-rich motif of Sos, the SH3-domain mediated interactions between Sos and Grb2 or Sos and the proteins involved in endocytosis could be sequential, similar to the association of Sos with Grb2 or E3b1, which depends on the activation status of the cell (35). The Grb2-Sos complex, which is stimulatory through the role of Sos as a guanine nucleotide exchange factor for Ras and Rac (31) is down-regulated by feedback phosphorylation through the MAP kinase pathway (36). The Grb2-Sos complex dissociates upon phosphorylation of Grb2 or Sos (36), leaving the proline-rich sequence of Sos open to associate with new factors, such as some of the interaction partners we describe here. Thus, the data presented is consistent with the idea of a proline-rich-SH3 domain switch, changing the role of Sos through the assembly of different multi-protein signaling complexes.
Interaction of Sos with Pacsin3

Pacsin3 belongs to the Pacsin/Syndapin family of proteins, which contain an N-terminal FCH domain (37), followed by a coiled-coil region possibly involved in oligomerization and a single SH3 domain at the extreme C-terminus (38). Pacsins/syndapins have been shown to interact with dynamin, synaptojanin, and N-WASP. Their over-expression inhibits endocytosis, suggesting they function as a critical link between endocytosis and the actin cytoskeleton (39,40). Pacsins 1 and 2 have recently been shown to interact with Sos1, and based on the localization of Pacsin to sites of rapid actin turnover it was suggested that Pacsins could function as adaptors guiding Sos to sites of actin filament reorganization (34). The interaction of Sos1 and Sos2 with Pacsin3 isoform has not been reported previously, but is not surprising given the conservation of the domain structure and the high sequence identity between the three isoforms. It is also interesting to note that Pacsin 3 and Grb2 can bind to the same motif of Sos isoforms.

Interaction of Sos with Snx9

Sorting nexins are a diverse family of proteins defined by their common POX homology (PX) domain binding different phosphatidylinositols (33). They are involved in various steps in endocytosis, mainly in vesicle trafficking. Snx9 interacts with the endocytic clathrin adapter complex 2 (41) and like Pacsin it is potentially linked to actin filaments by an interaction with WASP. In response to EGFR stimulation, Snx9 is phosphorylated by the AP-2 associated tyrosine kinase Ack2 and participates in EGFR signal attenuation. In contrast to the other proteins, Snx9 was found with a relatively low isotopic ratio of 4.3 ± 0.7, indicating less discrimination between the proline and alanine peptide form and suggesting lower affinity. Quantitative measurements, such as surface plasmon resonance experiments, could be used to validate this finding. Since Snx9 co-immunoprecipitated with endogenous full length Sos, it is also possible that additional motifs may be involved in the interaction of Sos and Snx9.

Possible further interaction partners
Three other proteins were found interacting with the proline-rich motifs of Sos1 or Sos2, either in the peptide pull-down or in the cross over experiment (Tab. 1). Remarkably, without exception these proteins have SH3 domains suggesting specific interaction with the bait peptides. One of these proteins, Snx18 is grouped with the sorting nexins because of its POX domain but has not been described in the literature so far. Since both Snx9 and Snx18 appear to bind Sos1, and Snx9 and Snx18 are the only ones among the sorting nexin family to share the same domain structure, they may have related functions. Another protein is CD2AP (42) belonging to the CMS family of adapter proteins (43). CD2AP binds the classical type II proline-rich domain in CD2 (44), and also binds the endocytotic adapter protein complex AP2 directly (45). Interestingly, Cbl, Grb2 and Sos were found in a CIN85 complex (46), a protein closely related to CD2AP, but direct interaction of Sos and CD2AP has not yet been reported.

Since these proteins were not validated, they are only listed as possible interaction partners in Tab. 1 and are not further discussed. With ongoing developments of mass spectrometric techniques, especially in their sensitivity, we expect that many of such possible interactors can more readily be verified.

**General aspects of the peptide-protein interaction screen**

The use of synthetic modified and unmodified peptides to find direct protein interaction partners appears to be a powerful method to dissect the role of certain motifs during signaling events. The peptide-protein interaction screen is specific and reproducible as shown by results of different experiments using different labeling methods. ‘Cross-over’ experiments define the variation of ratios for unspecific binding proteins, and identify specific interacting proteins by the fact that they yield the same ratio in both experiments. The method is unbiased and only requires the binding partner to be present in the cell lysate in an amount compatible with the mass spectrometric read out.

The peptide-protein interaction screen is exceedingly discriminating. In both examples tested, only a very small number of proteins were retrieved as significant binders from the thousands of proteins present in total cell lysates. This specificity is a result of the quantitative proteomic strategy employed.
Specific and non-specific binders can be distinguished without the need for stringent washing steps, which may eliminate weak binders. In the case of the proline-rich peptides from Sos, all retrieved proteins had SH3 domains, several of which were known to bind Sos or had been found in multiprotein complexes together with Sos. Most importantly, they all had previously been associated with endocytosis in connection with growth factor signaling. This fact and independent co-immunoprecipitation of Grb2, Pacsin3, and Snx9 together with Sos strongly argues that the interaction also occurs in vivo.

The method can complement approaches using full-length proteins as baits to purify multi-protein complexes (47,48). Those approaches are powerful in defining protein assemblies, but cannot assign binary interactions, let alone determine the ‘point of contact’ between two proteins. Furthermore, other methods can often not detect signal-dependent or modification-dependent interactions. Our peptide-protein interaction screen would also complement the yeast two-hybrid screen, which is less efficient for membrane proteins and interactions mediated by post-translational modifications. Likewise, it can be a follow-up for interactions found in genetic screens - for example synthetic lethal screens - or interactions predicted by bioinformatic methods. The method described here only applies to a subset of protein-protein interactions, namely the ones mediated by short peptide stretches. However, many regulated interactions are of this type, and it is only required that the interaction with the peptide has the same specificity, not the same affinity as the interaction with the whole protein.

Although during the pull-down experiments the bait peptide itself is not present at endogenous amounts, it bears the necessary modifications allowing specific binding of partner proteins. The potential partner proteins are present at endogenous levels in cell lysates bearing their endogenous modifications, helping to avoid non-physiological interactions.

Finally, the screen is relatively simple to perform given high sensitivity mass spectrometric equipment and associated software. Experiments such as the ones reported here can be ‘scaled up’ and eventually be performed in a large-scale
format. This would be very useful particularly for elucidating modification-dependent interactions, for which there are very few direct methods available at present.

**ACKNOWLEDGEMENT**

WS was supported by the Reimar-Lüst-Stipend of the Körber Foundation. The Center for Experimental BioInformatics is supported by the Danish National Research Foundation. The authors thank members of CEBI for constructive comments and discussion, especially Leonard Foster and Blagoy Blagoev.

**REFERENCES**

**FIGURE LEGENDS**

**Fig.1:** Principle of peptide-protein interaction screen

Protein populations are metabolically labeled to allow discrimination based on increased peptide masses (6 Da with $^{13}$C$_6$-Arg labeling or 3 Da with LeuD3 labeling). Peptides corresponding to regions potentially involved in interactions are synthesized in two closely related forms. In the figure, a phosphorylated peptide (indicated by the letter P) is incubated with the $^{13}$C$_6$-Arg labeled lysate whereas the unphosphorylated form of the same peptide is incubated with $^{12}$C$_6$-Arg lysate. Proteins bound to the bait peptides are eluted and mixed prior to analysis. After digestion of eluted proteins with trypsin, differentially labeled forms of tryptic peptides are detected by mass spectrometry. Tryptic peptides from proteins specifically binding the phosphorylated bait will have a larger peak intensity of the $^{13}$C-labeled form. Nonspecific binders will have a one to one ratio of both isotopic forms. Dotted lines indicate a ‘cross over’ experiment, in which specific binders should have high ratios of $^{12}$C to $^{13}$C labeled peptides.
**Fig.2**: Quantitative removal of Grb2 from cell lysate by Y-1068 phosphopeptide. Western blots of whole cell lysate before peptide pull-down and wash solution and eluate after a pull-down with either phosphorylated or non-phosphorylated Y-1068 peptide.
Fig. 3: Results of experiments using a phosphopeptide of EGF receptor as bait. (A) Mass spectrum of a doubly charged $^{13}$C$_6$-Arg-labeled tryptic peptide of Grb2 as identified through the interaction with the tyrosine-phosphorylated peptide of EGFR Y1068. (B) Mass spectrum of a doubly charged tryptic peptide of ribosomal protein L7, a non-specific binder. (C) Isotopic ratios of proteins identified also in 'cross over' experiment. Proteins with ratios of 10 were found exclusively with the form incubated with the phosphopeptide (D) Isotopic ratios of proteins identified in pull-down and 'cross over' experiment using a mutated phosphopeptide. No specific binding protein was detected in this experiment. Accession numbers refer to Swissprot.
Fig. 4: Results of a pull-down experiment using proline-rich motifs of Sos1 and Sos2 bait.

(A) Mass spectrum of a doubly charged $^{13}$C$_6$-Arg labeled tryptic peptide of Grb2 as identified through the interaction with the proline-rich domain of Sos1. (B) $^{13}$C$_6$-Arg labeled tryptic peptide of Snx9 as identified through the interaction with the proline-rich domain of Sos1. (C) Ratios in pull-down and 'cross over' experiments using the proline-rich motif of Sos1. (D) Ratios in pull-down and 'cross over' experiments using the proline-rich motif of Sos2.
**Fig. 5:** Co-immunoprecipitation of Grb2, Pacsin3 and Snx9 with Sos. Co-immunoprecipitating proteins were identified by western blot or mass spectrometry.

(A) Grb2 was co-immunoprecipitated with Sos and detected by western blot (insert) and mass spectrometry. (B) Fragmentation spectrum of a tryptic peptide of Pacsin3. (C) Fragmentation spectrum of a tryptic peptide of Snx9 as identified from a co-immunoprecipitated peptide mixture.
**Tables**

Table 1: Proteins identified in peptide affinity pull-downs using either a tyrosine-phosphorylated peptide of EGFR or a proline-rich motif of Sos1 and Sos2. Ratios are displayed as $^{13}$C$_6$ to $^{12}$C$_6$ for the pull down experiment and as $^{12}$C$_6$ to $^{13}$C$_6$ for the cross over experiments. Ratios ‘$>$ 10’ indicate that only one isotopic form was detected at the given signal-to-noise level. Proteins in bold were verified in co-immunoprecipitation experiments. 'no R': no Arginine containing peptide, 'n.d.' not detected. Entries in bold were verified by co-immunoprecipitation.

<table>
<thead>
<tr>
<th>Swissprot Nr.</th>
<th>Name</th>
<th>pull-down</th>
<th>'cross-over'</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>no. peptides</td>
<td>ratio</td>
</tr>
<tr>
<td><strong>EGFR:</strong> (biotin)-SGGLPVPEpYNQSV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P29354</td>
<td>Grb2</td>
<td>2</td>
<td>&gt;10 ($^{13}$C$_6$-Arg)</td>
</tr>
<tr>
<td><strong>Sos1:</strong> biotinSS-SGSGVPVPPPRRR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P29354</td>
<td>Grb2</td>
<td>4</td>
<td>&gt;10</td>
</tr>
<tr>
<td>Q9UKS6</td>
<td>Pacsin3</td>
<td>5</td>
<td>&gt;10</td>
</tr>
<tr>
<td>Q9Y5X1</td>
<td>Snx9</td>
<td>5</td>
<td>4.0 ± 0.5</td>
</tr>
<tr>
<td>Q9Y5K6</td>
<td>CD2AP</td>
<td>2</td>
<td>&gt;10</td>
</tr>
<tr>
<td>Q96RF0</td>
<td>Snx18</td>
<td>1</td>
<td>&gt;10</td>
</tr>
<tr>
<td><strong>Sos2:</strong> biotinSS-SGGSIPPLPRKK</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P29354</td>
<td>Grb2</td>
<td>3</td>
<td>&gt;10</td>
</tr>
<tr>
<td>Q9UKS6</td>
<td>Pacsin3</td>
<td>5</td>
<td>&gt;10</td>
</tr>
<tr>
<td>Q9Y5X1</td>
<td>Snx9</td>
<td>5</td>
<td>3.7 ± 0.2</td>
</tr>
<tr>
<td>Q9Y5K6</td>
<td>CD2AP</td>
<td>2</td>
<td>no R</td>
</tr>
<tr>
<td>P00519</td>
<td>Abl1</td>
<td>1</td>
<td>6.0 ± 0.3</td>
</tr>
</tbody>
</table>
A novel proteomic screen for peptide-protein interactions
Waltraud X. Schulze and Matthias Mann

J. Biol. Chem. published online December 16, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M309909200

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

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