Specificity and Commonality of the Phosphoinositide-Binding Proteome Analyzed by Quantitative Mass Spectrometry

Stephanie Jungmichel, Kathrine B. Sylvestersen, Chunaram Choudhary, Steve Nguyen, Matthias Mann, and Michael L. Nielsen

1Department of Proteomics, The Novo Nordisk Foundation Center for Protein Research, Faculty of Health Sciences, University of Copenhagen, Blegdamsvej 3, 2200 Copenhagen, Denmark
2Department of Proteomics and Signal Transduction, Max Planck Institute for Biochemistry, Martinsried 82152, Germany
*Correspondence: mmann@biochem.mpg.de (M.M.), michael.lund.nielsen@cpr.ku.dk (M.L.N.)
http://dx.doi.org/10.1016/j.celrep.2013.12.038

This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-No Derivative Works License, which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

SUMMARY

Phosphoinositides (PIPs) play key roles in signaling and disease. Using high-resolution quantitative mass spectrometry, we identified PIP-interacting proteins and profiled their binding specificities toward all seven PIP variants. This analysis revealed 405 PIP-binding proteins, which is greater than the total number of phospho- or ubiquitin-binding domains. Translocation and inhibitor assays of identified PIP-binding proteins confirmed that our methodology targets direct interactors. The PIP interactome encompasses proteins from diverse cellular compartments, prominently including the nucleus. Our data set revealed a consensus motif for PI(3,4,5)P3-interacting pleckstrin homology (PH) domains, which enabled in silico identification of phosphoinositide interactors. Members of the dedicator of cytokinesis family C exhibited specificity toward both PI(3,4,5)P3 and PI(4,5)P2. Structurally, this dual specificity is explained by a decreased number of positively charged residues in the L1 subdomain compared with DOCK1. The presented PIP-binding proteome and its specificity toward individual PIPs should be a valuable resource for the community.

INTRODUCTION

Phosphatidylinositol is a negatively charged phospholipid that represents less than 5% of the total phospholipid pool at the cytosolic side of eukaryotic cell membranes. Phosphatidylinositol can be phosphorylated by a variety of kinases on position 3, 4, or 5 of the inositol ring in seven different combinations. Phosphorylated forms of phosphatidylinositols, known as phosphoinositides (PIPs), play important roles in lipid-mediated cell signaling, membrane trafficking, and diseases involving these processes (Di Paolo and De Camilli, 2006). PIPs can act as precursors for secondary messengers or interact directly with proteins to orchestrate spatiotemporal activation of downstream signaling components (Berridge and Irvine, 1984; Cantley, 2002).

Despite biological interest in the PIP signaling pathways, our knowledge about the proteins that specifically interact with PIPs is limited. This is mainly due to the absence of “unbiased” technologies for detecting PIP interactions on a proteome-wide scale. Affinity matrices carrying tethered PIP variants have been used to isolate PIP-interacting proteins (Painter et al., 2001; Krugmann et al., 2002). Although this constituted an elegant biochemical approach, only a few PIPs could be investigated at a given time, and the specificity of the identified PIP-interacting candidates was unclear. Mass spectrometry (MS)-based proteomics has emerged as a key technology for comprehensive mapping of proteomes (de Godoy et al., 2008; Altelaar et al., 2013) and posttranslational modifications (PTMs) (Jensen, 2006; Witze et al., 2007), and is frequently employed to identify proteins bound to a “bait” such as peptides, RNA, or DNA. A central challenge in these experiments is to distinguish proteins that bind nonspecifically to the bait (background binders) from genuine interactors (specific binders). For example, previous studies performed pull-down experiments with immobilized PIPs to determine PI(3,5)P2, PI(4,5)P2, and PI(3,4,5)P3 interactors (Pasquali et al., 2007; Catimel et al., 2008, 2009; Rowland et al., 2011). Using low-resolution ion traps for data acquisition, these reports identified some of the known PIP binders, but also many additional proteins that are unlikely to be specific PIP binders. One can address this challenge by performing interaction screens in a quantitative format, most accurately by using stable isotope labeling approaches such as stable isotope labeling by amino acids in cell culture (SILAC) (Ong and Mann, 2005; Bantscheff et al., 2007). A strategy employing double-encoded quantitation for PIP studies was recently demonstrated, but the analysis was limited to a specific cellular compartment and only protein interactors for a single PIP were probed (Dixon et al., 2011; Lewis et al., 2011).
Here, we used triple-encoded SILAC quantification and high-resolution MS to systematically profile protein interaction specificities for all seven PIP variants. The data reveal an extensive catalog of PIP interactors and a quantitative estimate of their preferences for different PIP isoforms. We identified a large number of specific PIP interactors and validated several candidates by membrane translocation assays, which we discuss further below with regard to known PIP-binding domains (Seet et al., 2006). By probing highly specific PIP interactors in our data set, we established an extended PI(3,4,5)P3 consensus motif, allowing for in silico identification of PI(3,4,5)P3-binding proteins. Moreover, our data reveal insights into known PIP-binding protein families, as demonstrated by the identification of PIP-binding proteins, and extend the known PIP-binding specificity of dedicator of cytokinesis (DOCK) family members. Collectively, the presented data set provides a valuable resource for identifying PIP-binding proteins in human cells.

RESULTS

PIP Interaction Profiling Strategy
To identify PIP-interacting proteins and quantitatively profile the binding specificities for the different PIP isoforms, we used a SILAC-based proteomics approach (Figure 1A). HeLaS3 cells were grown under identical conditions in media containing “light,” “medium,” or “heavy” isotope-labeled variants of lysine and arginine (Ong et al., 2002). For an unbiased investigation of the human PIP interactome, all PIP isoforms were individually immobilized on agarose beads (PIP beads; Echelon) and incubated separately with light-, medium-, or heavy-labeled SILAC cell lysates. In the triple-SILAC approach, the interaction profile of three conditions can be analyzed simultaneously in a single experiment. However, to distinguish “background binders” from PIP-specific interactors, light-labeled SILAC lysates were incubated with beads coupled to PI(0), whereas phosphorylated

Figure 1. Experimental Setup
(A) Triple-encoded SILAC cell lysates were individually incubated with various PIPs.
(B) Three biological replicate experiments were performed to assess reliability and reproducibility. Each experiment contained four triple-encoded SILAC experiments (a total of 12 samples, labeled A1–C4).
(C) Each triple-encoded SILAC sample yielded quantitative information that was used to establish the PIP specificity for each identified protein. Data for two peptides from sample B1 (PLCD3 and PDPK1 peptide) and one PDPK1 peptide derived from sample C4 are depicted. Combining the quantitative information from all three experiments yielded a relative PIP specificity ratio (PSR) that corresponds to the PIP specificity for each quantified protein.
PIPs were incubated with either medium or heavy SILAC-labeled lysates. As a result, four SILAC experiments were needed to address the entire PIP interactome comprehensively (Figure 1B). The bound proteins were subsequently eluted, separated by SDS-PAGE, digested into peptides, and analyzed by liquid chromatography coupled to high-resolution tandem MS (LC MS/MS).

The protein-specific PIP affinity was determined by a quantitative comparison of peptide abundances across experiments as each peptide occurred as a SILAC triplet (Figure 1C). If the relative peptide abundance between light and medium or heavy SILAC peaks was 1:1, the identified protein was considered a background binder. In contrast, proteins that were significantly enriched in heavy-isotope-encoded samples appeared with high SILAC ratios and hence were considered as PIP-specific interactors. An example of the triple-SILAC intensity profile is illustrated in Figure 1C for the peptide sequence NLDLAAP TAEAEQR belonging to the protein PLCD3. No affinity toward PI(0) is observed in the light SILAC state (marked by a black arrow), whereas an abundant peak in the medium SILAC state signifies a strong PI(4,5)P2 interaction. For the heavy SILAC state, a peak of medium abundance demonstrates a relatively weaker affinity toward PI(3,4,5)P3. These data not only confirm the well-known affinity of PLCD3 for PI(4,5)P2 (Pawelczyk and Matecki, 1999) but also reveal the protein’s affinity for PI(3,4,5)P3, showing that PLCD3 exhibits dual PIP specificity. In comparison, the light and medium SILAC abundances for the peptide sequence LGCEEMEGYGPLK belonging to PDPK1 demonstrate that the protein has no specificity for PI(0) or PI(4,5)P2 (Figure 1C, middle section). Instead, an abundant signal present only in the heavy SILAC state demonstrates a strong affinity for PI(3,4,5)P3. Likewise, another peptide sequence LGCEEMEGYGPLK from PDPK1 protein confirms the strong affinity for PI(3,4,5)P3, whereas no affinity is observed for PI(4)P and PI(0) (Figure 1C, right panel). Combined, these data confirm the well-known PIP specificity of PDPK1 (Alessi et al., 1997; Stephens et al., 1998; Table S1) and provide a positive control for our screen. Moreover, since protein quantification was performed on the peptide level, every acquired mass spectrum of the identified peptides contributed to quantitative information for assessment of PIP affinity. In all experiments, we required that PIP interactors must be identified and quantified by a minimum of two unique peptides. Since many peptides from PIP interactors were quantified several times during their chromatographic elution time, the described PIP interactions were based on extensive quantitative data with high statistical significance.

**Identifying the PIP Interactome**

To obtain a high-quality PIP interactome and also assess the reproducibility of the approach, we analyzed three independent biological replicates for PIP interactors (Figure 1B). In total, our experiments identified 3,551 proteins representing a collection of potential PIP-interacting proteins (Table S2). Statistical analysis of the data set shows that the results from these unfiltered replicate experiments were highly consistent. For example, a comparison of the PI(3,4,5)P3 affinity data derived from sample A1 with the corresponding data from samples A4, B1, and C4 (Figure 1B, labeled Exp.A, Exp.B, and Exp.C) yielded a Pearson correlation above 0.70 between all three individual experiments.

It is worth noting that in interaction experiments, the actual SILAC ratios do not have to be identical between experiments, as they primarily serve to segregate background from specific interactors. Still, the correlation coefficients demonstrate that high quantitative reproducibility was achieved between the analyzed replicates (Figure 2A).

Having established that our data set contains highly reproducible protein ratios, we next determined specific PIP interactors by a “significant-outlier strategy,” as only a small subset of the 3,551 identified proteins are expected to be true PIP-interacting proteins. Each identified protein yielded a total of three SILAC ratios corresponding to light/medium, light/heavy, and medium/heavy from each SILAC sample analyzed (Figure 1B). These ratios were all extracted and normalized according to the summed peptide intensity of each protein relative to the signal for PI(0) (Figure 2B). Since high reproducibility was achieved in our analyses (Figure 2A), all PIP affinity ratios derived from the same protein across different samples were combined and normalized using the common control PI(0) (e.g., the PI(3,4,5)P3 affinity data derived from sample A1 were combined with the same PIP ratios from samples A4, B1, and C4). This allowed us to merge results from all experiments and to determine seven PIP specificity ratios (PSRs) for all identified proteins (one PSR for each investigated PIP isoform).

For statistical filtering, all PSR values were initially analyzed by boxplot analysis (Figure 2C), and only significant outliers (p < 0.01) were further considered. Then, the SD for each PSR value was calculated and only proteins that exhibited a PSR > 2 SDs were finally classified as PIP-interacting proteins. These filtered and significant PIP-interacting proteins are represented in Figure 2A by green (p < 0.01) and red (p < 0.001) dots. In this way, we identified a total of 405 proteins, constituting the largest experimentally determined PIP interactome reported to date (Tables S1 and S2). Among these PIP-binding proteins, we identified several well-characterized PIP-interacting proteins, including GAB1 (Rodrigues et al., 2000), RASA2 (Cozier et al., 2003), RASA3 (Cozier et al., 2003), PHLD1B (Zhou et al., 2010), CYTH3 (Klarlund et al., 1997), and PREX1 (Welch et al., 2002). Our data set confirms the known specificity of these proteins as visualized by their “PIP affinity and specificity chart” (Figure S1). This chart is based upon the derived PSR values for each protein (see above) and depicts the relative affinity and overall specificity of each PIP interactor toward all PIP isoforms. Each bar refers to a protein’s affinity toward a specific PIP, with a taller bar representing a relatively higher affinity. In addition to well-characterized interactors, several PIP-interacting proteins identified in our proteomics screen were recently validated by other groups through standard biochemical approaches (Feese et al., 2010; Komaba and Coluccio, 2010; Krick et al., 2012).

For further validation of our identified PIP-binding proteome, we performed membrane-translocation studies for several proteins upon platelet-derived growth factor (PDGF) stimulation in serum-starved SUM159 cells. Activation of the phosphatidylinositol 3-kinase (PI3K) pathway by PDGF causes a rapid buildup of PI(3,4,5)P3 at the plasma membrane (PM), and many PIP3 interactors have been reported to translocate to the PM in PDGF-stimulated cells, where they bind to PI(3,4,5)P3 (Auger et al., 1989; Hawkins et al., 1992). Such PDGF-dependent
translocation is demonstrated in Figure 3A for the well-characterized PI(3,4,5)P3-binding protein AKT1. In addition, we performed translocation assays for several PIP-interacting proteins, including ESYT1, PHLDA2, CCDC120, and TBC1D2 (Figures 3B–3E). ESYTs belong to a family of evolutionarily conserved mammalian proteins (referred to as extended synaptotagmin-like proteins) that are known to contain C2 domains (Min et al., 2007), and our analysis identified both ESYT1 and ESYT2 as PIP-binding proteins (Table S1). Since ESYT2 was recently demonstrated to be localized to the PM in a PI(4,5)P2-dependent manner (Giordano et al., 2013), we decided to investigate the PM translocation of ESYT1. To this end, we transfected SUM159 cells with full-length FLAG-tagged ESYT1, and after 5 min of PDGF stimulation, we observed membrane translocation (Figure 3B), which did not occur in unstimulated cells. Similar translocation assays on additional PIP-binding proteins also revealed clear PM translocation for PHLDA2, CCDC120, and TBC1D2, thereby confirming their affinity for PI(3,4,5)P3 (Figures 3 and S2).

To validate the specificity of our findings, we performed a PI3K inhibitor-dependent translocation study. SUM159 cells transfected separately with the PIP3-binding proteins TBC1D2 and PHLDA2, along with the positive control AKT1, were stimulated with PDGF in the presence or absence of the PI3K inhibitors Wortmannin or LYS294002 (Figure 3F). The data confirmed the PIP3-dependent translocation of the three targets upon PDGF stimulation, and a residual translocation of TBC1D2 could be observed upon PI3K inhibition. Since other PIPs besides PIP3 are known to accumulate at the PM upon PDGF stimulation (Auger et al., 1989; Hawkins et al., 1992), these results suggest that TBC1D2 (but not PHLDA2 or AKT1) would bind to these PIPs. This is in agreement with our proteomics data, which report that TBC1D2 exhibits dual specificity for both PI(4,5)P2 and PI(3,4,5)P3, whereas PHLDA2 exclusively binds PI(3,4,5)P3 (Table S1). Accordingly, these results demonstrate that the proteomics data can be utilized to decipher PIP specificities that are not easily entangled by common biochemical approaches.

Additionally, we performed western blot analysis for several full-length glutathione S-transferase (GST)-purified PIP-interacting proteins (not preferentially binding to PIP3) after pulldown with each of the seven PIP isoforms conjugated to agarose.
beads (Figure 4A). These results confirm the known and observed specificity of DAB2 toward both PI(4,5)P2 and PI(3,4,5)P3, and corroborate various PIP affinities reported in our SILAC screen for BANF1, SNX24, and ACAP2 (Table S1).

As our proteomics approach allows for concurrent determination of PIP specificity and affinity toward all seven PIPs within the same experiment, our data set reveals that PIP-binding proteins often exhibit selectivity toward more than one PIP. Such PIP-binding stereospecificity has been demonstrated for several PIP-binding proteins, such as PHLDB1 (Zhou et al., 2010) and DAB1 (Howell et al., 1999), highlighting that PIP-interacting proteins indeed exhibit stereospecificity in vivo and that these proteins are likely to encompass biological function. Our proteomic analysis confirms the known stereospecificity for PHLDB1 and DAB1, and together with in vivo evaluation of several PIP-interacting proteins, we conclude that our data set contains physiologically relevant PIP-binding proteins. Although many known PIP-interacting proteins were identified and validated, we did not find all known PIP interactors in our screen. As an example, DOCK2 and DAPP1 were not identified in our data set because these proteins have very cell-type-specific expression patterns (Dowler et al., 1999; Nishihara et al., 1999). Unfortunately, such cell-specific PIP-interacting proteins would require analyses of a large number of various cellular systems and consequently are not identifiable in our current setup.

To evaluate whether the identified proteins were direct PIP interactors, we assessed all PIP-interacting proteins for physical and/or functional interactions using the STRING software tool, which separately reports known and predicted protein interactors based upon direct and indirect associations (Snel et al., 2000). Although STRING extracts experimental data from databases such as BIND, DIP, GRID, HPRD, IntAct, MINT, and PID, as well as curated data from Biocarta, BioCyc, Gene Ontology (GO), KEGG, and Reactome, it does not encompass all direct or secondary PIP interactors. Still, STRING contains information on thousands of protein interactions, including the most prominent and well-studied PIP interactors. Secondary interaction partners...
Figure 4. Properties of PIP-Binding Proteins

(A) Pulldown of the recombinant, GST-tagged, full-length proteins DAB2, BANF1, SNX24, and ACAP2 with each of the seven PIP isoforms and PtdIns as a control bound to agarose beads. PIP-binding was detected by immunoblot using a GST-specific antibody.

(B) Domain architecture of PIP-interacting proteins, indicating Pfam domains that are significantly overrepresented.

(C) Cellular distribution of PIP-interacting proteins. Proteins were assigned based on exclusive GO annotations.

(D) Distribution of identified proteins according to their PIP specificity. Dark-brown bars represent the total pool of PIP-binding proteins, and light-brown bars represent nuclear annotated proteins only. A significant enrichment for P(4,5)P-interacting proteins in the nucleus is observed.
A

TAPP1

TAPP2

B

C

PIP(3,4,5)P3 binding PH motif
G-X(3)-K-X(7,13)-[FW]-X(2)-R-X-F-X(30,80)-[KR]-X(12,14)-W

D

Untreated cells

5 min after PDGF treatment
of these well-characterized PIP proteins are extensively described in the literature and therefore are present in the database. Importantly, STRING does not need to be comprehensive to assess the presence of secondary binding proteins in our data set. This secondary PIP interaction analysis, based upon experimental prediction methods and high-confidence STRING interactors (confidence score > 0.9 [von Mering et al., 2005]), revealed that the vast majority of proteins identified in our PIP interactome are likely to be direct PIP interactors, since only a few functional interactions are recorded in STRING among the PIP-interacting proteins (Figure S3). Pfam domain analysis revealed that PIP-interacting domains, such as the pleckstrin homology (PH) (Haslam et al., 1993), calponin homology (CH) (Fukami et al., 1996), RNA recognition motif (RBD) (Okada and Ye, 2009), and DOCK homology region (DHR) (Côté et al., 2005) domains, were highly enriched among our identified PIP-interacting proteins (Figure 4B). Taken together, these data confirm that our approach permits a stringent and high-confidence identification of the PIP interaction proteome (Table S1).

Next, we assessed the cellular distribution of identified PIP-interacting proteins and compared their distribution against the entire human proteome (Figure 4C). Surprisingly, a significant enrichment (p < 7.0 × 10^{-7}) of nuclear PIP interactors was evident, constituting the largest portion of identified PIP-interacting proteins. In contrast, the number of cytosolic interactors was significantly decreased (p < 5.5 × 10^{-7}), demonstrating that PIPs occupy several functionally distinct subcellular compartments. Our data are consistent with emerging evidence that PIPs play important biological roles in the nucleus (Barlow et al., 2010). To investigate whether interactors of any specific PIP might be enriched within the nucleus, we compared the distribution of the entire PIP interactome with its nuclear constituents (Figure 4D). PIP4P interactors were significantly (p < 4.0 × 10^{-7}) enriched in the nucleus. These results are consistent with previous reports of PIP4P in nuclear matrix preparations (Peyrastre et al., 1992), along with observations that membrane-free nuclei retain the capability to produce significant amounts of PIP4P (Boronenkov et al., 1998). Moreover, the clear preference for singly phosphorylated PIP4P, and less binding to multiphosphorylated PIPs (with increased negative charge) demonstrate that our data set entails specific PIP interactors and is not biased toward electrostatic interactions.

**TAPP1 and TAPP2 Exhibit Dual PIP Specificity**

To assess the sensitivity and dynamic range of the developed approach, we explored the PIP-binding properties of the cytoplasmic adaptor proteins tandem-PH-domain-containing-protein 1 (TAPP1) and TAPP2. These proteins assemble signaling complexes at the PM, and TAPP1 was previously demonstrated to bind strongly to PIP(3,4)P2 through its PH2 domain (Hogan et al., 2004).

It has been demonstrated that besides having affinity for PIP(3,4)P2, the PH2 domain of TAPP1 (referred to as AA054961 in this study) can bind the headgroup of Ins(1,3,4,5)P4, which shares structural similarity to PIP(3,4,5)P3 (Ferguson et al., 2000). Moreover, PM translocation was reported for the PH2 domain of TAPP2 upon PDGF stimulation of membrane-bound receptors, whereas the PH1 domain did not translocate under the same conditions (Supplemental Experimental Procedures; Park et al., 2008). These results demonstrate that the PH2 domain of TAPP1 exhibits affinity for PIP(3,4,5)P3 along with its affinity for PIP(3,4)P2. A recent study on the membrane penetration of various PH domains by quantitative surface plasmon resonance established that the PH2 domain of TAPP1 binds PIP(3,4)P2 roughly 20 times more strongly than it binds PIP(3,4,5)P3 (Manna et al., 2007). Such distinct preferences are often difficult to determine through standard biochemical approaches, which demonstrate the limited dynamic range of standard protein-lipid experiments (Dowler et al., 2000). In contrast, our data confirm that both TAPP1 and TAPP2 exhibit dual affinity toward PIP(3,4)P2 and PIP(3,4,5)P3 (Figure 5A) in agreement with previous observations, validating the ability of our proteomics approach to detect even subtle and distinct PIP-binding properties.

The observed dual PIP-binding specificity is due to the properties of the PH2 domains of TAPP1 and TAPP2 as described above. This is further highlighted by previous results demonstrating that the PH1 domains of TAPP1 and TAPP2 are only able to bind PIP(3,4,5)P3 upon site-directed mutagenesis (Thomas et al., 2001). To further investigate the PIP(3,4,5)P3-binding properties of the PH2 domain of TAPP1 and TAPP2, we superimposed its 3D structure onto known PIP(3,4,5)P3-specific PH domains (DAP1P, PDPK1, and AKT1). In this way, we identified three positively charged residues that are highly conserved (Figure 5B, residues highlighted in blue). These positively charged residues are in close proximity to the negatively charged phospho-groups of PIP(3,4,5)P3 and are important for stabilizing the PIP interaction through their positive charges. These observations are in full agreement with the fact that PH domains are electronically polarized and exhibit strong positive electrostatic potential in the proximity of the PIP-binding site (Kutateladze, 2010).

A sequence alignment analysis of all PH domains identified in our data set that exhibited PIP(3,4,5)P3 specificity revealed a strong conservation of these positively charged residues, which highlights the importance of these residues for affinity toward PIP(3,4,5)P3 (Figure 5C, residues highlighted in red and marked with an asterisk). In contrast, the PH1 domains of TAPP1 and...
TAPP2 do not show a similar conservation of positively charged residues (Figure 5C, marked by arrows). Instead, they contain several negatively charged amino acids (aspartic and glutamic acids) at these positions. These negatively charged residues are expected to have a counter-interacting effect on PIP interaction, which explains why the PH1 domain is only able to bind PI(3,4,5)P3 after site-directed mutagenesis. In conclusion, our data support previous observations that the PH2 domain of TAPP1 and TAPP2 exhibit subtle affinity toward PI(3,4,5)P3 due to a conserved sequence in the PH2 domain harboring positively charged residues, a signature that is also found in other known PI(3,4,5)P3-interacting PH domains.

Identification of an Extended PI(3,4,5)P3-Interacting PH Domain Motif

Besides the conserved positively charged residues across PI(3,4,5)P3 specific PH domains, the sequence alignment revealed the presence of a highly conserved PI(3,4,5)P3 interaction motif. The motif contains several determining residues and stretches across the entire PH domain from a conserved glycine at position 4 to a tryptophan at the C terminus of the PH domain (Figure 5C). This motif is more extensive than the previously reported PPBM motif (putative PI(3,4,5)P3-binding motif [Isakoff et al., 1998]) and the signature motif (Lietzke et al., 2000), which both are confined to the N-terminal part of PH domains. An in silico search identified 38 proteins harboring this PI(3,4,5)P3-binding motif in the human genome (Table 1). For 36 out of the 38 proteins, the extended sequence motif falls within a known PH domain, validating the specificity of the derived motif. Only 18 of the identified candidates have previously been reported as PIP interactors.

Among interactors, the in silico analysis revealed that all of the PLEKHA family proteins, except for PLEKHA3 (FAPP1) and PLEKHA8 (FAPP2), harbor a PH domain with specificity toward PI(3,4,5)P3. Recent studies confirmed that the PH domains of FAPP1 and FAPP2 exhibit affinity for PI(4)P, which further validates the PI(3,4,5)P3 specificity of our extended PH motif (Godi et al., 2004). Notably, PLEKHA proteins often contain more than one PIP-binding domain, and therefore may exhibit stereospecificity toward several PIPs besides PI(3,4,5)P3, as demonstrated by TAPP1 and TAPP2. However, our analysis reveals that the majority of PLEKHA proteins contain at least one domain that is responsible for interaction with PI(3,4,5)P3.

Among the 38 candidates identified in silico, 15 were identified in our MS screen. Two possible explanations for the fact that not all 38 candidates were identified are that (1) some of these proteins are likely to be expressed in a cell-type-specific manner and therefore may not be present in the HeLa cells used in this screen, and (2) some of these proteins are expressed at low levels. For example, a published HeLa proteome (Olsen et al., 2010) comprising 6,695 proteins identified only eight of the 38 candidates derived from the in silico analysis; however, all eight proteins were identified in our MS screen.

Overall, the in silico analysis identified PI(3,4,5)P3-binding PH domains from a wide range of proteins, including signaling adaptor proteins, regulators of small GTPases, kinases, and phosphatases. One PI(3,4,5)P3-binding protein identified through the in silico analysis was ADRBK1, a kinase that specifically phosphorylates the agonist-occupied form of the beta-adrenergic receptor (Aziziyeh et al., 2009). ADRBK1 was previously suggested to bind PI(4,5)P2 (Touhara et al., 1995), but was never predicted by other PIP motif analyses or demonstrated to exhibit PI(3,4,5)P3-binding properties. To validate the predicted affinity, we performed a membrane-translocation assay as previously described. When transfected SUM159 cells were stimulated with PDGF, a noticeable PM translocation of GFP-tagged ADRBK1 was observed, confirming the in silico-derived PIP-binding properties of ADRBK1 (Figures 5D and S2). Since ADRBK1 is coupled to G protein-mediated signaling, efficient translocation might require both PI3K and other G protein subunits in a manner similar to that previously reported for the PM translocation of PREX1 (Barber et al., 2007). Such preferences are not easily discernible in the presented translocation setup; however, our results demonstrate the specificity of the presented PI(3,4,5)P3-binding motif for in silico discovery of PI(3,4,5)P3-binding proteins.

Analysis of DOCK Family Proteins as PIP Interactors

Our proteomics screen identified several members of the DOCK protein family as PIP-interacting proteins. The DOCK family consists of related proteins involved in intracellular signaling networks, where they act as guanine nucleotide exchange factors (GEFs) for small G proteins of the Rho family, such as Rac and/or Cdc42. The DOCK family proteins consist of 11 proteins that are categorized into four subfamilies (DOCK-A, DOCK-B, DOCK-C, and DOCK-D) based upon their sequence homology (Figure 6A; Côté and Vuori, 2002). Among the 11 DOCK proteins, only DOCK1, DOCK4, and DOCK5 have previously been experimentally validated as PIP-interacting proteins (Côté et al., 2005), and our proteomic analysis confirms the reported affinity of these three DOCK proteins toward PI(3,4,5)P3 (Table S1). Moreover, our PIP analysis reveals that all DOCK protein families act as PIP interactors, and that DOCK6, DOCK7, and DOCK8 preferentially bind to PI(4,5)P2 and to a lesser extent to PI(3,4,5)P3 (Figure 6B).

The interaction of DOCK1, DOCK4, and DOCK5 proteins with PI(3,4,5)P3 is known to occur through the DHR-1 domain, which is conserved throughout all DOCK proteins. The 3D structure of the DOCK1 DHR-1 domain was recently determined, and it was discovered that the PI(3,4,5)P3 interaction with DOCK1 occurs through the L1 subdomain of DHR-1. Within this L1 subdomain, three positively charged lysine residues were identified as being...
responsible for the PIP specificity of DOCK1 (Premkumar et al., 2010). Sequence alignment of DHR-1 domains from all DOCK family members reveals that these lysine residues are highly conserved across DOCK proteins except for the DOCK-C family members (DOCK6, DOCK7, and DOCK8), which contain only two charged lysine residues in their L1 subdomain (Figure 6C). Considering our previous observations regarding conservation of charged residues in PH domains, and the known importance of positive electrostatic interaction in PIP-binding domains (Kutateladze, 2010), we suggest that the PIP specificity of DOCK-C family members stems from the reduced number of charged residues in their L1 domains. Since DOCK6, DOCK7, and DOCK8 have only two charged lysine residues within their L1 domain, this would allow for strong interaction with either

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Protein Name</th>
<th>Known PH Domain</th>
<th>AA Region</th>
<th>Known PI(3,4,5)P3 Interactor</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACAP1</td>
<td>Arf-GAP with coiled-coil, ANK repeat, and PH-domain-containing protein 1</td>
<td>yes</td>
<td>270–351</td>
<td>yes</td>
</tr>
<tr>
<td>ACAP2</td>
<td>Arf-GAP with coiled-coil, ANK repeat, and PH-domain-containing protein 2</td>
<td>yes</td>
<td>271–352</td>
<td>yes</td>
</tr>
<tr>
<td>ACAP3</td>
<td>Arf-GAP with coiled-coil, ANK repeat, and PH-domain-containing protein 3</td>
<td>yes</td>
<td>273–354</td>
<td>yes</td>
</tr>
<tr>
<td>ADAP1</td>
<td>Arf-GAP with dual PH-domain-containing protein 1</td>
<td>yes</td>
<td>134–221</td>
<td>yes</td>
</tr>
<tr>
<td>ADAP2</td>
<td>Arf-GAP with dual PH-domain-containing protein 2</td>
<td>yes</td>
<td>137–224</td>
<td>yes</td>
</tr>
<tr>
<td>AKT1</td>
<td>RAC-alpha serine/threonine-protein kinase</td>
<td>yes</td>
<td>10–99</td>
<td>yes</td>
</tr>
<tr>
<td>AKT2</td>
<td>RAC-beta serine/threonine-protein kinase</td>
<td>yes</td>
<td>10–99</td>
<td>yes</td>
</tr>
<tr>
<td>AKT3</td>
<td>RAC-gamma serine/threonine-protein kinase</td>
<td>yes</td>
<td>10–98</td>
<td>yes</td>
</tr>
<tr>
<td>ADRBK1</td>
<td>β-adrenergic receptor kinase 1</td>
<td>yes</td>
<td>563–643</td>
<td>no</td>
</tr>
<tr>
<td>ADRBK2</td>
<td>β-adrenergic receptor kinase 2</td>
<td>yes</td>
<td>563–643</td>
<td>no</td>
</tr>
<tr>
<td>DAPP1</td>
<td>dual adaptor for phosphotyrosine and 3-phosphotyrosine and 3-PIP</td>
<td>yes</td>
<td>169–256</td>
<td>yes</td>
</tr>
<tr>
<td>DEF6</td>
<td>differentially expressed in FDCP 6 homolog</td>
<td>yes</td>
<td>221–303</td>
<td>yes</td>
</tr>
<tr>
<td>DGKH</td>
<td>diacylglycerol kinase eta</td>
<td>yes</td>
<td>70–149</td>
<td>no</td>
</tr>
<tr>
<td>DGKK</td>
<td>diacylglycerol kinase kappa</td>
<td>yes</td>
<td>221–303</td>
<td>no</td>
</tr>
<tr>
<td>EPHA3</td>
<td>ephrin type-A receptor 3</td>
<td>yes</td>
<td>367–457</td>
<td>no</td>
</tr>
<tr>
<td>GAB1</td>
<td>GRB2-associated-binding protein 1</td>
<td>yes</td>
<td>10–107</td>
<td>yes</td>
</tr>
<tr>
<td>GAB2</td>
<td>GRB2-associated-binding protein 2</td>
<td>yes</td>
<td>11–108</td>
<td>yes</td>
</tr>
<tr>
<td>GAB3</td>
<td>GRB2-associated-binding protein 3</td>
<td>yes</td>
<td>10–108</td>
<td>yes</td>
</tr>
<tr>
<td>GAB4</td>
<td>GRB2-associated-binding protein 4</td>
<td>yes</td>
<td>44–143</td>
<td>yes</td>
</tr>
<tr>
<td>SBF1</td>
<td>myotubularin-related protein 5</td>
<td>yes</td>
<td>1,766–1,856</td>
<td>yes</td>
</tr>
<tr>
<td>SBF2</td>
<td>myotubularin-related protein 13</td>
<td>yes</td>
<td>1,748–1,838</td>
<td>no</td>
</tr>
<tr>
<td>MYO10</td>
<td>myosin-X</td>
<td>yes</td>
<td>1,217–1,301</td>
<td>yes</td>
</tr>
<tr>
<td>OSBP9</td>
<td>oxysterol-binding protein-related protein 9</td>
<td>yes</td>
<td>7–90</td>
<td>no</td>
</tr>
<tr>
<td>PHLD1</td>
<td>PH-like domain family B member 1</td>
<td>yes</td>
<td>1,261–1,361</td>
<td>yes</td>
</tr>
<tr>
<td>PHLD2</td>
<td>PH-like domain family B member 2</td>
<td>yes</td>
<td>1,148–1,237</td>
<td>yes</td>
</tr>
<tr>
<td>PLEKHA1</td>
<td>PH-domain-containing family A member 1</td>
<td>yes</td>
<td>196–280</td>
<td>no</td>
</tr>
<tr>
<td>PLEKHA2</td>
<td>PH-domain-containing family A member 2</td>
<td>yes</td>
<td>203–289</td>
<td>no</td>
</tr>
<tr>
<td>PLEKHA4</td>
<td>PH-domain-containing family A member 4</td>
<td>yes</td>
<td>59–144</td>
<td>no</td>
</tr>
<tr>
<td>PLEKHA5</td>
<td>PH-domain-containing family A member 5</td>
<td>yes</td>
<td>174–259</td>
<td>no</td>
</tr>
<tr>
<td>PLEKHA6</td>
<td>PH-domain-containing family A member 6</td>
<td>yes</td>
<td>64–149</td>
<td>no</td>
</tr>
<tr>
<td>PLEKHA7</td>
<td>PH-domain-containing family A member 7</td>
<td>yes</td>
<td>169–273</td>
<td>no</td>
</tr>
<tr>
<td>PLEKHH1</td>
<td>PH-domain-containing family H member 1</td>
<td>yes</td>
<td>583–663</td>
<td>no</td>
</tr>
<tr>
<td>RASA2</td>
<td>Ras GTPase-activating protein 2</td>
<td>yes</td>
<td>609–724</td>
<td>yes</td>
</tr>
<tr>
<td>RASA3</td>
<td>Ras GTPase-activating protein 3</td>
<td>yes</td>
<td>581–696</td>
<td>yes</td>
</tr>
<tr>
<td>ARHGAP22</td>
<td>Rho GTPase-activating protein 22</td>
<td>yes</td>
<td>36–145</td>
<td>no</td>
</tr>
<tr>
<td>SWAP70</td>
<td>switch-associated protein 70</td>
<td>yes</td>
<td>215–297</td>
<td>yes</td>
</tr>
<tr>
<td>TAF1</td>
<td>transcription initiation factor TFIID subunit 1</td>
<td>yes</td>
<td>793–862</td>
<td>no</td>
</tr>
</tbody>
</table>

The predicted location of the motif falls within a known PH domain for 36 of the candidates.
Among some of the PIP-recognizing domains, the DHR-1 explain the unexpectedly large size of the PIP interactome. This may the protein as warranted by cellular conditions, without going is a powerful mechanism to increase the local concentration of have PIP-recognizing domains. The presence of such a domain prominently including the nucleus. A large proportion of them interactors. Data analysis revealed that the PIP interactome actome and identified a total of 405 PIP-interacting proteins, SILAC-based strategy. Our approach provides an extended view of the human PIP inter-actome using quantitative MS. The combination of high-resolu-
tion MS with triple-encoded SILAC allowed for reliable protein identification and concurrent distinction between true PIP inter-
actors and background binding proteins. Triplicate analysis of our PIP interaction screen revealed high reproducibility between experiments.

Our approach efficiently assesses PIP interactions on a global scale and reveals additional information about PIP-inter-
acting proteins. In previous studies, many PIP interactions were established through in vitro assays, in which binding is often analyzed using only a particular region and/or domain under investigation rather than the full-length protein. Additionally, inter-
actions have often been characterized using strategies in which all seven PIPs are concurrently introduced to a protein of interest. Thus, subtle PIP specificities may be masked if the affinity of a protein for any single PIP is particularly strong. Furthermore, these approaches lack global data that can help to disentangle interactions that occur through secondary pro-
teins. These limitations are greatly circumvented or alleviated by individually introducing PIPs to whole-cell lysates in a SILAC-based strategy.

Our approach provides an extended view of the PIP inter-
actome and identified a total of 405 PIP-interacting proteins, including many well-known interactors as well as additional interactors. Data analysis revealed that the PIP interactome encompasses proteins from diverse cellular compartments, prominently including the nucleus. A large proportion of them have PIP-recognizing domains. The presence of such a domain is a powerful mechanism to increase the local concentration of the protein as warranted by cellular conditions, without going through relatively slow gene-expression changes. This may explain the unexpectedly large size of the PIP interactome. Among some of the PIP-recognizing domains, the DHR-1 domain of the DOCK family was revealed to exhibit dual specificity according to the number of lysine residues residing within its L1 subdomain. A higher affinity toward PI(3,4,5)P3 was observed for DHR-1 domains harboring three lysine residues in L1, whereas dual specificity toward PI(4,5)P2 and PI(3,4,5)P3 occurred when it contained only two lysines. This dual specificity follows the DOCK-C family’s dual GEF activity for Rac1 and Cdc42.

In conclusion, our data reveal that the PIP interactome is remarkably expanded, such that it contains more protein inter-
actors than the total number containing either phospho- or ubiquitin-binding domains in the genome (Venter et al., 2001). We anticipate that this data set will be a valuable resource for the community, and may provide novel insights into the PIP inter-
actome and allow functional testing of proteins that were previ-
ously not known to be PIP interactors.

**DISCUSSION**

We have carried out a proteomic analysis of the human PIP inter-
actome using quantitative MS. The combination of high-resolu-
tion MS with triple-encoded SILAC allowed for reliable protein identification and concurrent distinction between true PIP inter-
actors and background binding proteins. Triplicate analysis of our PIP interaction screen revealed high reproducibility between experiments.

Our approach efficiently assesses PIP interactions on a global scale and reveals additional information about PIP-inter-
acting proteins. In previous studies, many PIP interactions were established through in vitro assays, in which binding is often analyzed using only a particular region and/or domain under investigation rather than the full-length protein. Additionally, inter-
actions have often been characterized using strategies in which all seven PIPs are concurrently introduced to a protein of interest. Thus, subtle PIP specificities may be masked if the affinity of a protein for any single PIP is particularly strong. Furthermore, these approaches lack global data that can help to disentangle interactions that occur through secondary pro-
teins. These limitations are greatly circumvented or alleviated by individually introducing PIPs to whole-cell lysates in a SILAC-based strategy.

Our approach provides an extended view of the PIP inter-
actome and identified a total of 405 PIP-interacting proteins, including many well-known interactors as well as additional interactors. Data analysis revealed that the PIP interactome encompasses proteins from diverse cellular compartments, prominently including the nucleus. A large proportion of them have PIP-recognizing domains. The presence of such a domain is a powerful mechanism to increase the local concentration of the protein as warranted by cellular conditions, without going through relatively slow gene-expression changes. This may explain the unexpectedly large size of the PIP interactome. Among some of the PIP-recognizing domains, the DHR-1 domain of the DOCK family was revealed to exhibit dual specificity according to the number of lysine residues residing within its L1 subdomain. A higher affinity toward PI(3,4,5)P3 was observed for DHR-1 domains harboring three lysine residues in L1, whereas dual specificity toward PI(4,5)P2 and PI(3,4,5)P3 occurred when it contained only two lysines. This dual specificity follows the DOCK-C family’s dual GEF activity for Rac1 and Cdc42.

In conclusion, our data reveal that the PIP interactome is remarkably expanded, such that it contains more protein inter-
actors than the total number containing either phospho- or ubiquitin-binding domains in the genome (Venter et al., 2001). We anticipate that this data set will be a valuable resource for the community, and may provide novel insights into the PIP inter-
actome and allow functional testing of proteins that were previ-
ously not known to be PIP interactors.

**EXPERIMENTAL PROCEDURES**

A detailed summary of all materials and methods used in this study can be found in the accompanying the Supplemental Experimental Procedures.

**Enrichment of PIP Protein Interactors**

PIP beads were obtained from Echelon Biosciences. One PIP variant was incubated with one SILAC lysate state for 3 hr at 4°C. Beads were washed five times with washing buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 0.25% Ige-
pal). Bound protein interactors were eluted by boiling PIP beads in Laemmli Sample Buffer at 95°C for 5 min. Proteins were resolved on 4%–20% SDS-
PAGE and stained with colloidal Coomassie blue, cut into 20 slices, and pro-
cessed for MS analysis.

**MS Analysis**

All MS experiments were performed on a nanoscale high-performance liquid chromatography system connected to a hybrid linear trap quadrupole (LTQ) Orbitrap mass spectrometer. The MS instrument was operated in data-depen-
dent mode to automatically switch between full-scan MS and MS/MS acquisi-
tion. Survey full-scan MS spectra were acquired in the Orbitrap analyzer with resolution R = 60,000. The ten most intense peptide ions with charge states > 2 were sequentially isolated and fragmented by collision-induced dissociation (CID) in the LTQ mass spectrometer.

**SUPPLEMENTAL INFORMATION**

Supplemental Information includes Supplemental Experimental Procedures, three figures, and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2013.12.038.

**ACKNOWLEDGMENTS**

Dr. Alexei Degterev, Sackler School of Graduate Biomedical Sciences, Tuft University, is acknowledged for providing the SUM159 cell line. We thank members of the Novo Nordisk Foundation Center for Protein Research for fruitful discussions and careful readings of the manuscript. The work carried out in this study was supported in part by the Novo Nordisk Foundation, The Lundbeck Foundation, and the European Commission’s Seventh Framework Program PRIME-XS (contract no. 262067, FP7-INFRASTRUCTURES-2010-1).

**Figure 6. The Entire DOCK Family Entails PIP-Binding Proteins**

(A) Schematic structure of DOCK proteins and their division into subfamilies based upon structural differences.

(B) Sequence alignment of the DHR-1 domain from all 11 DOCK proteins. Green boxes refer to L1, L2, and L3 domains, and blue boxes refer to structural loops. Highlighted in red are the charged residues within the L1 domain that are responsible for PIP interaction.

(C) Specificity chart for various DOCK proteins. The individual specificities follow the structural differences within the DOCK subfamilies.
REFERENCES


