Chapter from the book Integrative Proteomics
Downloaded from: http://www.intechopen.com/books/integrative-proteomics

Interested in publishing with InTechOpen?
Contact us at book.department@intechopen.com
Functional Proteomics: 
Mapping Lipid-Protein Interactomes

Clive D’Santos¹ and Aurélia E. Lewis²

¹Cancer Research UK Cambridge, Cambridge
²Department of Molecular Biology,
University of Bergen, Bergen
¹United Kingdom
²Norway

1. Introduction

Cell function is dependent upon the co-ordinated and dynamic formation of complex interaction networks between molecules of diverse biochemical properties. These networks, or interactomes, are comprised of macromolecular biopolymers; proteins, DNA, RNA and polysaccharides, in addition to non-polymer compounds such as small molecular metabolites. This myriad of interactions is highly regulated and any perturbation or alteration has potential to result in disease.

Profiling protein-protein interactions has been the major focus of interactomics in the past few years (Charbonnier et al. 2008) largely due to the advances in technological platforms that have the capacity to probe globally. Early efforts have included two-hybrid screens to identify binary binding interactions; more recent studies have used a range of mass spectrometry based methods to identify protein complexes that are a better reflection of multi-interactive nature of such complexes. Protein/small molecule interactions are equally important in modulating the function of their target proteins but few studies have analyzed these interactions on a large scale. The field is indeed still in its infancy due to difficulties in identifying metabolites but has recently benefitted from technological advances in mass spectrometry, data analysis software and metabolites database development for the measurement and identification of metabolites. The next step is to integrate metabolomic profiling to functional characterization of metabolic pathways by identifying systematically metabolite-protein interactions.

Research efforts have in general been more focused on lipid-mediated interactions and this chapter reviews the global methods as well as their applications used to map lipid-protein interactomes based on mass spectrometry or arrays. The potential of these studies to deepen our understanding on the biological function of metabolites as protein effectors is also discussed.

2. Metabolomics

Metabolites are defined as small organic molecules produced and modified by a living organism as a result of cellular and physiological metabolism. These molecules constitute an
important fraction of the dry weight of a living cell ranging from 17 to 27% in bacteria and mammalian cells respectively. They consist of a wide variety of small molecules with a vast chemical diversity, including amino acids, nucleotides, sugars and fatty acids that are central to all metabolic pathways existing in the cell. The precise number of metabolites produced in a cell at a certain time point is unknown but is estimated to range from a few hundreds in bacteria to a few thousands in plant and animal cells. Metabolic networks reconstructed from studies in yeast have indicated up to 1494 different metabolic compounds (Herrgard et al. 2008). The human metabolome database (version 2.5) embraces 7982 compounds that have been experimentally confirmed (Wishart et al. 2009). These compounds have been further divided into 52 different classes. The number and diversity of possible metabolites is extensive and this entails that a significant proportion of proteins may form functional but also opportunistic interactions with metabolites.

Overall metabolites constitute the metabolome of a cell, tissue or organism at a specific time and changes in metabolic profiles have enormous potential to understand cellular function and for clinical diagnostics (Vinayavekhin et al. 2010). To this effect, metabolomics has been applied to the general profiling of metabolites in biological samples, the discovery of biomarkers in diseases and the clinical screening of targeted compounds. Metabolomics gives an additional biologically relevant dimension to transcriptomics and proteomics and the integration of these data allows for a deeper understanding of physiological processes in normal and pathological states. While transcriptomics, proteomics and metabolomics allow the cellular inventory of biochemicals, an additional layer of data integration is still necessary to assess the mechanisms of regulation leading to a specific metabolic status. Computational-based metabolic flux analysis provides information on intracellular flux distributions of metabolic processes of a cell or an organism that can be integrated to data generated through transcriptomics, proteomics and metabolomics (Blank & Kuepfer 2010). The range and diversity of functional metabolites has been highlighted and it is beyond the scope of this review to critique the methods that are currently used for all metabolites. Rather we have chosen to focus on a subset of metabolites, lipids, and in particular inositol-phospholipids, which are key regulators of numerous signalling pathways and which have been the most studied in recent years.

3. Functional lipidomics: From lipidomics to lipid-protein interactomics

The last 5 to 10 years have witnessed an incredible advancement in mass spectrometry and bioinformatics to analyse and identify systematically most lipids existing in biological samples at any one time and under different conditions in a specific entity (Wenk 2010; van Meer & de Kroon 2011). Cells contain thousands of lipids with a large chemical diversity and in light of recent advancement in lipid research, their classification has recently been updated by the LIPID MAPS initiative (Fahy et al. 2009). Lipidomics provide snapshots of the biochemical status of a cell and have the potential to complement transcriptomic and proteomic profiles. Integration of metabolomics, including lipidomics, to transcriptomics and proteomics analyses is expected to improve our understanding of metabolic pathways in health and diseases.

Although metabolite profiling provides important information on the status of a cell or organism, there is still a lack of functional data. Recently a shift has occurred from profiling all existing metabolites in a cell, tissue or an organism (metabolomics) to understanding how they may affect cellular functions (functional metabolomics) by the identification of
metabolites-protein interaction networks. In this chapter, we have highlighted the methods as well as their applications used to map lipid-protein interactions in biological systems.

3.1 Methods to identify lipid-protein interactions
Small scale and large scale mapping of lipid-protein interactions methods have been developed using either targeted strategies studying a specific lipid or protein of interest (Figure 1) or large-scale strategies using lipid arrays or protein arrays (Figure 2). Protein capture using affinity-based pull down has been widely used in combination with mass spectrometry to identify lipid interactomes in particular (Figure 1). In these cases cell extracts are incubated with lipid conjugated to affinity matrices and bound proteins are identified by mass spectrometry (Krugmann et al. 2002; Scholten et al. 2006; Osborne et al. 2007; Pasquali et al. 2007; Catimel et al. 2008; Catimel et al. 2009; Lewis et al. 2011).

Opposite strategies to identify lipids bound to a selection of proteins or a particular protein of interest have also been developed (Tagore et al. 2008; Kim et al. 2011; Li & Snyder 2011) (Figure 1). Recombinant proteins can be purified and immobilised onto a solid support and exposed to a metabolite mixture obtained from cells or tissues where the protein is known to be expressed. Additionally, endogenous proteins or tagged proteins can be immunoprecipitated from a cell or tissue extract (Li et al. 2007; Urs et al. 2007). Metabolites that are bound to the isolated protein are eluted and identified by mass spectrometry.

High-throughput screening strategies of these interactions have also been established using protein and small molecules microarrays (Lueking et al. 2005; Chen & Snyder 2010; Wu et al.)

Fig. 1. Methods to identify lipid-protein interactions using targeted methods
Fig. 2. Methods to identify lipid-protein interactions using large-scale methods

Microarrays are collections of hundreds to thousands of molecules immobilised on planar surfaces such as glass slides or nitrocellulose coated slides. Protein microarrays consist of individually expressed and purified proteins representing the complete or partial proteome known for a particular organism. Small molecules arrays consist of synthetic or naturally occurring molecules printed or spotted onto solid surfaces. To assess metabolites-protein interactions, protein arrays are exposed to fluorescently labelled metabolites (Zhu et al. 2001). Small molecules arrays are exposed to individual proteins (Rogers et al. 2011) or cellular lysates containing tagged proteins and interactions are detected using antibody recognising the specific tag (Gallego et al. 2010).

3.2 Lipid-protein interactomes
Lipids represent the largest class of metabolites in cells and are involved in a wide variety of cellular functions (van Meer & de Kroon 2011). They are essential structural components of cellular membranes and function as energy stores, cellular signalling molecules and regulators of transcription factor. Recent lipidomics analyses in mammalian cells have highlighted the dynamic remodelling of different lipid molecules (Dennis et al. 2010). These molecules have therefore been the focus of un-biased and systematic interactome studies in an effort to further clarify the functions of these molecules.

3.2.1 Phosphoinositide-protein interactomes mapping using lipid affinity matrices capture combined with MS of proteins
Many studies have focused on the identification of phosphoinositides (PIs)-protein interactomes. Inositol phospholipids, a small subset of the total lipid pool function as key...
regulators of numerous regulatory pathways (Toker 2002; Janmey & Lindberg 2004; Di Paolo & De Camilli 2006; Poccia & Larijani 2009). They function predominantly but not exclusively as sensors that recruit proteins and protein complexes to sites of synthesis in response to external cues (Lindmo & Stenmark 2006; Lemmon 2008). Target proteins possess well-characterized domains within their structure that bind with varying affinity and specificity to the phosphorylated inositol head group. In addition, the hydrolysis of these lipids by phospholipase activities generate further second messengers such as diacylglycerols and polyphosphorylated inositols extending the influence of these lipids on cellular function and highlighting the need for further efforts to understand molecular mechanisms. A first step toward this would be to identify specific effector protein complexes that are regulated directly via binding and from this point of view proteomic methods and their applications are well placed to characterize globally the macromolecular complexes.

A number of studies have focused on the identification of PI-binding proteins using affinity matrices to pull down potential PI interacting proteins from cellular lysates and subsequent mass spectrometry analyses. These studies are summarized in Table 1. In a study using a combination of PI affinity matrices, competitive lipid pull down and protein fractionation

<table>
<thead>
<tr>
<th>PI interactome analysed</th>
<th>Method</th>
<th>Cell type/subcellular compartment</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>PtdIns(3,4,5)P₃</td>
<td>PI conjugated beads and MS</td>
<td>Pig leukocyte cytosolic extract</td>
<td>(Krugmann et al. 2002)</td>
</tr>
<tr>
<td>Mostly PtdIns(3,4)P₂</td>
<td>PI conjugated to cleavable S-S bond biotin + streptavidin beads and MS</td>
<td>Primary macrophage cytosolic extract</td>
<td>(Pasquali et al. 2007)</td>
</tr>
<tr>
<td>PtdIns(4,5)P₂</td>
<td>Biotinylated PI, streptavidine conjugated beads and MS</td>
<td>Secretory granules from bovine adrenal chromaffin cells</td>
<td>(Osborne et al. 2007)</td>
</tr>
<tr>
<td>PtdIns(3,5)P₂ &amp; PtdIns(4,5)P₂</td>
<td>PI conjugated beads or liposomes and MS</td>
<td>LIM1215 colon cancer cell cytosolic extract</td>
<td>(Catimel et al. 2008)</td>
</tr>
<tr>
<td>PtdIns(3,4,5)P₃</td>
<td>PI conjugated beads or liposomes and MS</td>
<td>LIM1215 colon cancer cell cytosolic extract</td>
<td>(Catimel et al. 2009)</td>
</tr>
<tr>
<td>PtdIns(4,5)P₂</td>
<td>PI conjugated beads and quantitative MS</td>
<td>Neomycin extracted nuclear proteins isolated from murine erythroleukemia (MEL) cells</td>
<td>(Lewis et al. 2011)</td>
</tr>
<tr>
<td>PtdIns(3,4)P₂</td>
<td>Stimulation of class I PI3K +/- wortmannin, biotinylated PI coupled to streptavidin beads and SILAC-based quantitative MS</td>
<td>1321N1 astrocytoma membrane fractions</td>
<td>(Dixon et al. 2011)</td>
</tr>
</tbody>
</table>

Table 1. Large-scale proteomics studies for the identification of PI binding proteins by MS
from pig leukocyte cytosol, 16 proteins were identified as PtdIns(3,4,5)P³ and 5 as PtdIns(3,4)P² binding proteins by mass spectrometry (Krugmann et al. 2002). One of these proteins, ARAP3, a GTPase-activating protein, was further characterized as a functional PtdIns(3,4,5)P³ effector protein (Krugmann et al. 2002). Another study identified 10 known and 11 potentially novel PtdIns(3,4)P² interacting proteins using cleavable biotinylated PI baits (Pasquali et al. 2007). None of these proteins overlapped with the ones identified in the previous study.

In a more comprehensive study, Holmes and colleagues have characterized and compared the interactomes of PtdIns(3,5)P² and PtdIns(4,5)P² (Catimel et al. 2008) as well as PtdIns(3,4,5)P³ (Catimel et al. 2008; Catimel et al. 2009) determined from the cytosolic extracts of colon cancer cells expressing WT PI3 kinase. PIs immobilized either onto beads or incorporated into liposomes were used for protein capture from the cytosolic extracts. This led to the identification of 388 proteins in complex with PtdIns(3,5)P² and/or PtdIns(4,5)P² (Catimel et al. 2008) and 282 proteins in complex with PtdIns(3,4,5)P³ (Catimel et al. 2009). A fraction of these were found to form complexes only with PtdIns(3,5)P² (69), PtdIns(4,5)P² (146) or PtdIns(3,4,5)P³ (141). In addition significant overlaps were observed for these interactions, consistent with the promiscuous properties of some of these interactions. These studies represent the first comprehensive datasets of potential cytosolic PI-interacting proteins. In addition, the computational analyses of the molecular functions of proteins found in complex with cytosolic PI interactomes have highlighted roles in the regulation of GTPases, in transport/trafficking, cytoskeletal remodelling, phosphorylation-mediated post-translational modifications.

The first organellar PI interactome was deciphered from secretary granules. Secretary granules were isolated from PC12 cells and 5 PtdIns(4,5)P² binding proteins were identified by affinity lipid pull down and mass spectrometry. These interactions were all validated by lipid pull down and Western immunoblotting.

PIs are also found in the nucleus (Irvine 2003; Hammond et al. 2004; Ye & Ahn 2008; Keune et al. 2011) and we have established a quantitative and proteomic method to identify PtdIns(4,5)P² interacting proteins to gain insight into the PI-mediated nuclear functions (Lewis et al. 2011). The workflow of the method is schematised in Figure 3. The nuclear PtdIns(4,5)P² interactome was characterized using PI-conjugated beads incubated in neomycin-extracted nuclear proteins mixtures and quantitative mass spectrometry using isotopic labeling of cells. Neomycin is known to bind to PIs with high affinity (Schacht 1978; Gabev et al. 1989) and we predicted that neomycin would compete for PIs in complex with proteins. Incubation of intact nuclei with neomycin resulted in the specific displacement of 168 nuclear proteins harbouring a PI binding domain. Using neomycin extracts, 34 proteins were shown to interact with PtdIns(4,5)P² in quantitative affinity purification using specific lipid conjugated matrices. Neomycin extraction of proteins represented an ideal preparation from which to affinity-purify PI-effector proteins using specific lipid conjugated matrices, avoiding the issues of sample complexity and dynamic range. Functional classification and enrichment analyses of the identified PtdIns(4,5)P²-interacting proteins pointed to roles in mRNA transcriptional regulation, mRNA splicing and protein folding.

Dixon and colleagues have recently developed a three phase affinity enrichment method to quantitatively identify PtdIns(3,4)P² effector proteins targeted to membranes (Dixon et al. 2011). 1321N1 astrocytoma cells labelled with either light or heavy isotope were stimulated with bpV, a vanadate analogue, which induces high levels of PtdIns(3,4)P², and in the presence or absence of wortmannin, an inhibitor of the PI3K pathway. After the isolation of
membranes, proteins specifically recruited to membrane fractions following bpV stimulation, were eluted with Ins(1,3,4)P3. Eluted proteins were subjected to ion-exchange chromatography, affinity capture with streptavidin beads pre-coupled to PtdIns(3,4)P2, followed by SDS-PAGE, LC-MS/MS and quantitative assessment of PtdIns(3,4)P2 effector proteins. Previously established PtdIns(3,4)P2-binding proteins, such as TAPP1 and Akt1-3, were identified, providing a strong proof of principle of the method. Overall 80-85 potential proteins were identified and this study provided the first quantitative MS-based identification of PtdIns(3,4)P2 effector proteins. Many but not all proteins harboured lipid binding domains. The binding characteristics of a novel binding protein, IQGAP1, to PtdIns(3,4)P2 were determined, demonstrating the existence of an atypical PI binding domain.

Overall, studies based on affinity capture combined with mass spectrometry serve as useful resources and have the advantage to give a global view of the biological functions of proteins regulated by PIs in different cellular compartments. However a main drawback remains in the inability to discriminate between direct and indirect interactions through
associated proteins. Such analyses should be complemented by biochemical approaches analysing direct interactions for individual proteins. In addition, the potential existence of indirect protein complex networks can be assessed using known data for protein-protein interaction networks for the corresponding cell line or tissue explored.

3.2.2 Lipid-protein interactomes mapping by protein immobilization or affinity purification and identification of lipids by MS
Mass spectrometry has allowed the identification of lipids interacting with proteins both in targeted and large-scale systematic analyses. Different methods have been developed to affinity capture proteins followed by the extraction of bound lipids and their identification by tandem mass spectrometry.

3.2.2.1 Targeted identification of ligands for nuclear receptors
Several studies have focused on developing methods to identify physiological ligands for orphan nuclear receptors. Nuclear receptors represent a family of transcription factors that are activated by binding to specific small molecules to regulate the expression of specific genes.

Saghatelian and colleagues have developed methods to identify indiscriminately the metabolites bound to recombinant proteins (Tagore et al. 2008). A protein of interest is purified from bacteria, immobilized on a solid support via a 6xHis or GST tag and incubated with a lipid extract obtained from cells known to express the corresponding protein. Eluted metabolites are analysed by LC-MS and the metabolite chromatogram profiles are compared computationally to control samples obtained from solid support alone. This strategy was applied for the nuclear receptors, peroxisome proliferator-activated receptors (PPAR)\(\alpha\) and \(\gamma\) (Kim et al. 2011) involved in lipid metabolism. Free fatty acids (FFA) such as arachidonic (C20:4), linoleic (C18:2) and oleic (C18:1) acids were identified as endogenous ligands for both nuclear receptors. Palmitoleic acid (C16:1) was also identified as a ligand for PPAR\(\gamma\).

In an alternative method, a physiological ligand was discovered for PPAR\(\alpha\) by isolating the receptor from liver nuclear extracts obtained from mice either WT or lacking fatty acid synthase (FAS) (Chakravarthy et al. 2009). FAS is an enzyme that synthesizes saturated FA which was previously shown to synthesize de novo a potential ligand for PPAR\(\alpha\) in liver cells (Chakravarthy et al. 2005). After elution of the receptor, lipids were extracted and subjected to tandem MS, which identified the phospholipid, 1-palmitoyl-2-oleyl-sn-glycerol-3-phosphocholine as a FAS-dependent ligand of PPAR\(\alpha\) (Chakravarthy et al. 2009). This is a compelling approach to decipher endogenous ligand occupancy of orphan nuclear receptor in an in vivo setting.

Sewer and colleagues were able to identify several phospholipids bound to another orphan nuclear receptor, steroidogenic factor 1 (SF-1) (Li et al. 2007) with roles in the regulation of steroidogenic hormones expression. SF-1 was immunoprecipitated from adrenocortical cells which express the receptor endogenously and phospholipids were analysed by LC-MS. Phosphatidic acid was found to be a major lipid bound to SF-1 and to activate the transcriptional activity of the receptor. A similar approach allowed the identification of linoleic acid (C18:2) as a ligand for another orphan nuclear receptor hepatocyte nuclear factor 4 (HNF4) \(\alpha\), affinity purified from mammalian cells (Yuan et al. 2009). Importantly the occupancy of the ligand was dependent upon the physiological condition studied: HNF4\(\alpha\) was bound to linoleic acid when the receptor was isolated from livers of fed mice but not of fasted mice. Additionally the ligand did not have any effect on the transcriptional activity of HNF4\(\alpha\).
3.2.2.2 Large scale identification of metabolite-protein interactions

An affinity purification protocol in yeast was recently established by Snyder and colleagues to identify hydrophobic metabolites bound to 103 protein kinases as well as to a selection of proteins including 21 enzymes involved in the ergosterol biosynthetic pathways (yeast molecular analogue of cholesterol) (Li et al. 2010). In this case proteins were fused to an immunoglobulin binding domain and isolated from yeast extracts by affinity pull down. Metabolites interacting with the affinity purified proteins were extracted and identified by LC-MS. Control samples consisted of a yeast strain extract devoid of the corresponding fused protein. Such systematic analysis revealed that about 70% of the ergosterol biosynthetic enzymes and 20% of all protein kinases analyzed were bound to hydrophobic molecules. Known protein-metabolites interactions were observed but a majority of new interactions were also uncovered. Some interactions were unexpected and suggested important roles for ergosterol in the regulation of not only lipid biosynthetic pathways but also of many kinases, amongst which Ypk1 yeast kinase homologue to the mammalian kinase Akt.

3.2.3 Large scale interactomics to identify lipid-protein interactions

The magnitude of metabolites-protein interactions has been highlighted in large-scale screens in budding yeasts using different approaches using protein or lipid arrays (Zhu et al. 2001; Gallego et al. 2010). Using systematic approaches such as these, a comprehensive set of proteins can be simultaneously assessed for their potential interactions with lipids but also with other small molecules.

Firstly, Snyder and colleagues developed protein chips for the yeast proteome of Saccharomyces cerevisiae. These were the first protein arrays for any organism to be engineered (Zhu et al. 2001). The yeast proteome array contained 5800 proteins fused to GST-6xHis and was screened for PI interactions using PIs assembled in liposomes containing phosphatidylycholine (PC) and an additional biotinylated lipid. PtdIns(3)P₂, PtdIns(4)P₂, PtdIns(3,4)P₂, PtdIns(4,5)P₂ or PtdIns(3,4,5)P₃, containing liposomes were applied to the arrays, followed by an incubation with fluorescently-labeled streptavidin. Following the fluorescence detection of the arrays, 49 proteins were found to interact significantly with PIs compared to PC liposomes, with different affinities and specificities for the different PI molecules. Conventional methods were applied to confirm protein-PI interactions for 3 proteins involved in glucose metabolism that were not previously expected to bind to PIs.

The second screen reported is a large-scale analysis of yeast lipid-proteins interactions that was recently performed by Gavin and colleagues (Gallego et al. 2010). An opposite approach was used and lipid arrays were generated using 56 different lipids spanning the main classes of lipids found in yeast applied on nitrocellulose membranes. These arrays were incubated with cell extracts expressing single tandem affinity purification (TAP)-tagged proteins in S. cerevisiae. Interactions with 172 single TAP-tagged protein containing extracts were assessed for their potential interaction with the lipid array and gave rise to 530 interactions involving 124 proteins and 30 lipids. Amongst the 56 lipids studied, PIs were represented by PtdIns(3)P, PtdIns(4)P, PtdIns(3,4)P₂ or PtdIns(3,4,5)P₃, and 86 proteins were found to bind PIs of which 77% harboured a lipid binding domain. PIs represented indeed the lipid category that interacted with most proteins, which is consistent with the wide range of cellular functions they are reported to take part into. Importantly, this study also assessed the quality of the data generated by the lipid array by comparing data...
retrieved from the literature and from genetic interactions. In addition the identified interactions by lipid overlay were validated for 8 proteins chosen amongst the obtained dataset. Overall, Gavin and colleagues reported that 54% of the identified interactions were validated by additional genetic evidence, making this interactome dataset the most comprehensive resource for lipid biology.

Unfortunately, very little overlap, accounting for about 5%, could be observed for PI-protein interactions between the 2 previously described studies despite an overlap of 88% of the proteins analysed (Zhu et al. 2001; Gallego et al. 2010). These contrasting datasets may be explained by different interaction properties being measured. For example, the protein chips may not allow the access of potential binding domain of all proteins to the phospholipids and therefore prevent potential interactions. In addition not all the proteins are overexpressed and purified as full length proteins or at sufficiently high level for the assay. As for the lipid array experiments, indirect interactions are also possible.

Although extensive datasets are now available from PI interactomes in yeast and mammalian cell lines, the overlap between the PI interactomes determined in these different species is unknown. We have therefore attempted to compare the dataset obtained from PtdIns(4,5)P2 nuclear interactome studies (Lewis et al. 2011) to the datasets obtained in S. cerevisiae. Using InParanoid 7 (Ostlund et al. 2010), 18 yeast orthologues were recovered from the 34 murine proteins reported in the PtdIns(4,5)P2 nuclear interactome dataset. Out of the 18 yeast orthologues, 3 proteins were in common with PI binding proteins identified by protein chip lipid overlay by Snyder and colleagues (Zhu et al. 2001). These proteins are listed in Table 2. Cdc19 and Cct8, were found to interact with PtdIns(4)P in the yeast PI interactomics study. Cam1 was identified as a PtdIns(4,5)P2-interacting protein in the same study, which was consistent with our study. These findings certainly warrant further characterization of these proteins. In contrast, none of the 18 orthologues were found to be common to the dataset obtained from the lipid array screen by Gavin and colleagues (Gallego et al. 2010) and this may be explained by the following points. Firstly, the majority of proteins chosen in the lipid array screen were known to harbour at least one lipid binding domain (LBD) as defined by the online tools for protein domain assignment SMART, Pfam or Superfamily, whereas none of the identified nuclear proteins harboured such domains but rather simple basic amino acid rich patches (Lewis et al. 2011). Secondly, the proportion of proteins annotated to the nucleus compartment analysed in the lipid array screen is not known.

<table>
<thead>
<tr>
<th>ORF (yeast)</th>
<th>Gene name (yeast)</th>
<th>Uniprot ID (mouse)</th>
<th>Gene name (mouse)</th>
<th>Protein description (mammalian)</th>
</tr>
</thead>
<tbody>
<tr>
<td>YAL038W</td>
<td>Cdc19</td>
<td>P52480</td>
<td>Pkm2</td>
<td>Isoform M2 of Pyruvate kinase isozymes M1/M2</td>
</tr>
<tr>
<td>YJL008C</td>
<td>Cct8</td>
<td>P42932</td>
<td>P46671</td>
<td>T-complex protein 1 subunit theta</td>
</tr>
<tr>
<td>YPL048W</td>
<td>Cam1</td>
<td>Q9D8N0</td>
<td>Eef1g</td>
<td>Elongation factor 1-gamma</td>
</tr>
</tbody>
</table>

Table 2. List of genes found in common between the datasets obtained from the mammalian PtdIns(4,5)P2 nuclear interactome (Lewis et al. 2011) and the yeast PI interactome (Zhu et al. 2001). Datasets that were compared include all yeast proteins found to bind PIs in the study by Zhu et al and 18 yeast orthologues retrieved from the 34 murine proteins identified in the PtdIns(4,5)P2 nuclear interactome.
4. Chemical proteomics for lipid pathway drug specificity validation

Chemical proteomics has recently been used to explore the specificity of known drugs. In the case of drugs targeting lipid pathways, this method was used to identify all possible protein targets of a class I PI3K inhibitor, LY294002 (Gharbi et al. 2007). An analogue of the PI3K class I inhibitor LY294002, PI828, was immobilized onto epoxy-activated sepharose beads and used to pull out protein targets from whole cell extracts obtained from a human epithelial cell line (HeLa) and a mouse lymphoma B-cell line (WEHI231). Protein targets were eluted and identified by LC-MS/MS. This study demonstrated that this compound, bound not only to class I PI3Ks and other PI3K-related kinases, but also non-lipid kinases which was consistent with the inhibitory profile previously known for this compound. However novel targets were also identified which were reported to possibly explain some of the off-target cellular effects of this compound. The use of such proteomic approach has the potential to determine the specificity of known or new drugs at the cellular level as well as the potential cellular functions altered by the compound studied.

5. Comparison of PI interactomics data to whole genome genetic screens

Chemical genomics preceded the advent of chemical proteomics due to the completion of the S. cerevisiae Deletion Project which allowed whole genome genetic screens of different drugs. Such screen was performed to reveal new functions of PI metabolism by using wortmannin to identify genes which could confer altered sensitivity to the drug (Zewail et al. 2003). In yeast, wortmannin inhibits PtdIns(4)P kinase, Stt4p, and its inhibitory effects have been reported to be due to the depletion of PtdIns(4,5)P2 (Cutler et al. 1997). This screen allowed the identification of 591 genetic interactions due to wortmannin resistance and provided an overview of the actions of the PI pathway. New functions that were not previously attributed to the PI metabolic pathway were uncovered, namely DNA replication and DNA damage checkpoint, chromatin remodelling and proteasome-mediated protein degradation.

A fraction of protein-protein interaction networks can be correlated to genetic interaction networks in yeast. Since wortmannin has been reported to affect the pool of PtdIns(4,5)P2, we assumed that a fraction of PtdIns(4,5)P2 effector proteins identified in physical screens would coincide with a fraction of the wortmannin genetic interaction screen. We have therefore compared our PtdIns(4,5)P2 interaction networks obtained from mammalian nuclei to the wortmannin genetic screen performed in yeast. We were able to identify 4 genes in common between the physical and the genetic interaction datasets and these are listed in Table 3 and shown in the Venn diagram in Figure 4. In addition, 1 of these genes, Cam1, corresponding to the mammalian orthologue Eef1g (Elongation factor-1 γ), was also found to be common to the PI-protein interactomics study from Snyder and colleagues (Zhu et al. 2001). Interestingly, Cam1 was initially characterised as a possible phospholipid binding protein (Creutz et al. 1991; Kambouris et al. 1993), which would therefore be consistent with both physical and genetic studies.

Moreover, comparing the datasets obtained from the PI interactome study from Zhu et al to the wortmannin genetic screen from Zewail et al identified 18 proteins in common (14%). The overlapping data is presented as a Venn diagram in Figure 4.
Table 3. List of proteins found in common in the PtdIns(4,5)P2 nuclear interactome study (Lewis et al. 2011) and in the wortmannin genetic screen in *S. cerevisiae* (Zewail et al. 2003). Datasets included 18 yeast orthologues retrieved from the 34 murine proteins identified in the PtdIns(4,5)P2 nuclear interactome and all of the genes that conferred wortmannin resistance when deleted individually in yeast.

<table>
<thead>
<tr>
<th>ORF (yeast)</th>
<th>Gene name (yeast)</th>
<th>Uniprot ID (mouse)</th>
<th>Gene name (mouse)</th>
<th>Protein description (mammalian)</th>
</tr>
</thead>
<tbody>
<tr>
<td>YER081W</td>
<td>Cam1</td>
<td>Q9D8N0</td>
<td>Eef1g</td>
<td>Elongation factor 1-gamma</td>
</tr>
<tr>
<td>YDR518W</td>
<td>Eug1</td>
<td>P27773</td>
<td>Pdia3</td>
<td>Protein disulfide-isomerase A3 precursor</td>
</tr>
<tr>
<td>YMR072W</td>
<td>Abf2</td>
<td>P63158 P30681</td>
<td>Hmgb1 Hmgb2</td>
<td>High mobility group protein B1 High mobility group protein B2</td>
</tr>
</tbody>
</table>

Fig. 4. Venn diagram representation of common proteins identified in PI interactome studies (Zhu et al. 2001; Lewis et al. 2011) and chemical genomics screen using wortmannin (Zewail et al. 2003). Datasets that were compared included all yeast proteins found to bind PIs in the study by Zhu *et al.*, the 18 yeast orthologues retrieved from the 34 murine proteins identified in the PtdIns(4,5)P2 nuclear interactome (Lewis *et al.*) and all of the genes that conferred wortmannin resistance when deleted individually in yeast (Zewail *et al.*).

6. **Conclusion: Challenges and future direction in lipid-protein interactomics**

Systematic and unbiased proteomics studies have answered some of the questions regarding lipid-mediated pathway functions. Most studies have focused on mapping PI interactomes from separate cellular compartments while more recent studies have expanded our knowledge to other lipid subclass interaction networks. In addition, the availability of genome wide genetic screen in yeast allows the potential discovery of overlaps with...
physical interactions datasets, thereby strengthening interaction data. Identifying lipid binding proteins has indeed provided some insights into the possible biological functions of the corresponding lipids mainly by inference of protein function. The next challenge is to assign biological functions to each of these interactions.

Several large scale interactome studies have shown little overlap in findings and both false positive and false negative are likely to be generated by these types of methods. Overall a great body of data is now indeed available and studies will still be required to further validate these interactions at the biochemical and cellular level, in vitro and in vivo. The reported interactomes are at present synonymous with a static view of molecular complexes and this raises therefore a number of questions and challenges worthy of further scrutiny. What are the lipid-protein interactomes at the sub-cellular level? How are these interactions regulated in time and space? What are the mechanisms of regulation? What are the modes of interactions? Moreover do these interactions affect other types of interactions mediated by other macromolecules? This last question entails probably a new challenge in systems biology, i.e. the integration of data obtained from lipid-protein interactomes to those obtained from other protein-macromolecule interactomes.

Finally, lipids, in particular signalling lipids such as PIs, but also eicosanoids, sphingolipids and fatty acids are known to control critical cellular functions and the alteration of lipid-mediated pathways is known to contribute to the development of pathologies, such as chronic inflammation, cancer, neurodegenerative and metabolic diseases (Pendaries et al. 2003; Wymann & Schneider 2008; Skwarek & Boulianne 2009). Newly acquired knowledge on lipid-protein interactions may pinpoint potential lipid effectors that may become targets for drug therapies. This may become even more relevant if changes in lipid-protein networks are identified in pathological states.

7. References


Proteomics was thought to be a natural extension after the field of genomics has deposited significant amount of data. However, simply taking a straight verbatim approach to catalog all proteins in all tissues of different organisms is not viable. Researchers may need to focus on the perspectives of proteomics that are essential to the functional outcome of the cells. In Integrative Proteomics, expert researchers contribute both historical perspectives, new developments in sample preparation, gel-based and non-gel-based protein separation and identification using mass spectrometry. Substantial chapters are describing studies of the sub-proteomes such as phosphoproteome or glycoproteomes which are directly related to functional outcomes of the cells. Structural proteomics related to pharmaceutics development is also a perspective of the essence. Bioinformatics tools that can mine proteomics data and lead to pathway analyses become an integral part of proteomics. Integrative proteomics covers both look-backs and look-outs of proteomics. It is an ideal reference for students, new researchers, and experienced scientists who want to get an overview or insights into new development of the proteomics field.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: